Environmental Science: an Evidence-Based Study of Earth's Natural Systems

Environmental Science: an Evidence-Based Study of Earth's Natural Systems

JASON KELSEY

MUHLENBERG COLLEGE
ALLENTOWN, PENNSYLVANIA

Contents

Preface

Welcome!

Environmental Science, an Evidence-Based Study of Earth's Natural Systems, is an introductory textbook intended to provide no-cost yet quality learning materials for college students with little or no background in the subject.

You might wonder: as there is no shortage of introductory environmental science textbooks on the market, why write yet another? That's a good question, one I have wrestled with for some time. First and foremost, no single published volume has met all the needs I have for the course I teach most often. After over two decades of failing to find the right fit, I decided it was time for me to author my own book. I hope it also serves students and instructors beyond my campus. Second and related to the first, I want students to have a book that is firmly grounded in science, one driven by data and evidence rather than advocacy and ideology. In other words, I have tried to stay objective in my approach to a field that can be emotionally charged. Third, I think a book on environmental science can and should be accessible, engaging, and dare I say, entertaining, while maintaining rigor and relevance. Frankly, you must learn a lot of new vocabulary and concepts in your study of such a complex subject, but I have worked to weave these ideas into a coherent and readable narrative. In other words, you should not equate my stated attempt to stay objective with "dry" or "dull". My wish is that some of my love for the story of Earth's natural systems—and all their intricacies—rubs off on and inspires you to learn more about your world.

I wish you well on your educational journey, whether you continue to pursue environmental science or if this will be your only foray into the field. Since our understanding of the natural world advances through good science, scrutiny, and communication, I encourage you to reach out to me to make comments, suggestions, or to ask questions. I would love to hear from you.

Jason W. Kelsey, Ph.D.,
Allentown, Pennsylvania
kelsey@muhlenberg.edu
April, 2023

Textbook features at a glace

1. Accessible introduction to the broad field of environmental science
2. Evidence-based coverage of the science of important and controversial topics
3. Key concepts listed at the beginning of each chapter and questions to promote additional self-guided learning presented at the end of each chapter
4. Extensive cross referencing within and among chapters
5. Glossary which can be searched or accessed through highlighted terms
6. Regular updates and revisions to the text as science advances
7. The book layout has been optimized for web browsers on computer screens; some formatting may be lost on tablets and mobile phones or if the text is converted to pdf.

Acknowledgements

To the thousands of students who have taken courses with me: thank you for your inspiration to share, laugh, teach, and learn.

To the many colleagues at Muhlenberg College and elsewhere who provided reviews and comments: thank you for generously sharing your time, expertise, and insights. The late Don Shive, a long-time Muhlenberg Chemistry Professor and informal mentor to me, was particularly important in providing encouragement, instruction, and suggestions right up to the day he died.

To Tim Clarke and Lora Taub in the Muhlenberg Digital Learning

Center and Kelly Cannon in Muhlenberg's Trexler Library: thank you for all the help, time, and patience.

To Alissa Knopf: thank you for your assistance in the preparation of the Key-concepts and Think-about-it-some-more boxes.

To my children: thank you for enduring the years of writing and, apologies for the world I will leave to you.

To my wife Nicki: thank you for your support, patience, encouragement, and love during all the ups and downs of this project.

A grant from the Pennsylvania Consortium for the Liberal Arts (PCLA), as well as sabbatical leaves provided by Muhlenberg College, enabled me to complete some of the work on this book.

About the author

Jason Kelsey has been a member of the faculty at Muhlenberg College in Allentown, PA, since 2002. He is a professor and director of The Program in Environmental Science. Previously, he taught in the Biology Department at Marymount University in Arlington, VA. He received his Ph.D. and M.S. in Environmental Toxicology from Cornell University and his A.B. in Geology and English from Lafayette College. He has published numerous scholarly papers on the biodegradation and bioaccumulation of soil pollutants.

PART I

1. An Introduction to Interdisciplinary Environmental Science

JASON KELSEY

Environmental science is a term familiar to most modern audiences, but its meaning often is misunderstood. Some people equate it with activism, protests, and rallies about preservation of wild areas and organisms. Others see it as a simple set of rules about how we should separate clear from green glass, turn off lights in vacant rooms, use less water, and buy recycled paper. Such assumptions lead many to conclude that environmental scientists are opponents to business, technology, freedom, and, frankly, good times. It turns out that little about these or related views is accurate, as we will soon see.

Key concepts

After reading Chapter 1, you should understand the following:

- The scope of environmental science
- The distinction between objective environmental science and subjective environmentalism

- How environmental science developed as a field
- The storage, movement, and finite nature of Earth's materials
- How environmental science informs the principles and goals of sustainability
- The meaning and relevance of the term environment
- How organisms interact with their environments
- The importance and roles of the principle of environmental unity, the precautionary principle, and cost-benefit analysis
- The complicated relationship between Earth and humans

Chapter contents

1.1. DEFINITION OF ENVIRONMENTAL SCIENCE

What, then, is environmental science? We begin with a seemingly straight-forward answer: it is the study of Earth's natural systems and the forces affecting them. It is characterized by several important features.

1.1.1. It has a broad scope

If you stop and think about it, you will realize the enormous scope and complexity of this field, because by "Earth's natural systems" we mean nothing short of all the living and non-living entities on this planet. Several sciences and their tools are clearly required if we are to take on this giant task, most notably biology, ecology, chemistry, geology, and physics. Put another way, it is very much a *multidisciplinary* endeavor. In addition, though, since the sciences on our list do not remain separate, rather they are combined in new ways to help address questions unique to our field, environmental science is *interdisciplinary*. By drawing upon and blending many disciplines, environmental scientists can study various issues, for example the ways organisms interact with their surroundings, how critical resources are produced, cycled, and transformed, and how human activity influences Earth and its many systems (including the quality of life for people).

1.1.2. It is objective

Environmental scientists strive to increase our knowledge and understanding of Earth's natural systems. Put succinctly, we observe phenomena, collect evidence, and report findings and data—that is, we are **objective** in our approach. We must not interfere with or offer **subjective**, values-based judgements about what we think *ought* to be true. Note how our science is *not* the same as the similar-sounding **environmentalism;** the later term refers to a type of advocacy driven by a sense of right and wrong and a desire to protect nature from human activity (Figure 1.1 suggests one way environmental scientists and environmentalists differ).

Figure 1.1. *A tree might be hugged by an environmentalist, left, or measured by a scientist, right. Kelsey,* CC BY-NC-SA.

Objectivity is a particularly critical characteristic for us to consider because, unlike fields such as chemistry or physics, many people assume environmental science is a matter of opinion. It is rare to hear chemists accused of bias when they describe the workings of an atom or the amount of energy required to heat a given volume of water, say, but when an environmental scientist speaks about pollution, extinction, or climate, they will often be subject to criticism for what are perceived to be politically motivated conclusions. As we proceed through this book, you should learn to recognize the difference between approaches that are scientific and those that are non-scientific. We will see much more about science as a way of knowing and what distinguishes it in Chapter 2.

1.1.3. It is a relatively new field

People have been paying attention to the world around them for a lot longer than this, but a field known formally as "environmental science" did not really come into being until the 1960s. Biology, chemistry, and the rest mentioned above have arguably been around for centuries, so environmental science represents a novel way to study the Earth. Just what sparked the formation of this new field? The answer is complicated, but it is tied to the birth of what is known as the **modern environmental movement**. Several events, including the industrial revolution, development of assembly lines, urbanization, and large-scale agriculture, made people increasingly conscious of the potential adverse effects humans can have on the Earth and its systems. The widespread use of synthetic **pesticides** (Chapter 9) that began in the 1940s, though, had a particularly profound effect. Put very briefly, some researchers began to worry that these chemicals designed to kill unwanted insects could also affect organisms we value. One of the most important contributors to this growing awareness was a woman named Rachel Carson. Her 1962 book, *Silent Spring*, was a fictionalized tale of a future with no more songbirds (hence the book's title). What was her villain? Indiscriminate use of pesticides such as **DDT**, which through various mechanisms we will explore in Chapter 15, killed birds. The reaction to Carson's work was strong, with many in industry (particularly those who manufactured and sold pesticides) quite critical of both her science and politics. She was publicly vilified and ridiculed by some powerful people, although others embraced her message and began to study the processes about which she warned. She died shortly after the book was published, but she was ultimately vindicated by the work of a presidential scientific commission. Additionally, the formation of the United States **Environmental Protection Agency** in 1970 and the U.S. ban on the use of DDT in 1972 were influenced by Carson's book and her followers.

1.1.4. It is highly relevant and often newsworthy

Environmental scientists have become increasingly visible during the past several decades because they are generally called upon to study and fix environmental damage caused by human activity. Since the 1960s, they have worked on many dramatic and highly publicized incidents involving the exposure of people to poisons. It is by no means an exhaustive list, but some of the most noteworthy are briefly described here.

Cayahoga River, Cleveland, OH, U.S.A., 1969

This river was so polluted by industrial waste that its surface caught fire, capturing the attention of the public and highlighting some of the risks associated with unregulated dumping of hazardous substances into water.

Love Canal, City of Niagara Falls, NY, U.S.A., 1976 – 1977

This event was a defining moment in the environmental movement of the 1970s. Very briefly, hazardous waste was dumped into an unused canal between 1920 and 1953 and then buried. The area was developed into a neighborhood, including a public school, and seemingly no attention was given to the potential hidden risks beneath the surface. All was well until very heavy rains started to expose the buried materials in 1976. The public was horrified by media images of school children fleeing from noxious odors, the melting of tires of cars parked in chemical puddles, and the idea that careless (or worse) management of industrial products could threaten human lives. The effects were exaggerated by some, including reports of multiple deaths that never occurred; nevertheless, a serious health crisis ensued. Shortly thereafter, the place was deemed uninhabitable, and it was purchased by the United States government. An expensive clean-up was undertaken, paid for by a combination of federal and industry funds, and nearly 30 years later most of the neighborhood was reopened (Figure 1.2 is a photo of a barrier and a warning sign posted during the site remediation). The aftermath of this event included protests, suspicion, lobbying, and the enactment of many new laws controlling hazardous waste management (see Table 1.1 for some examples). Environmental science as a field also became noticeably more important.

Figure 1.2. *Love Canal during the clean up. Access was controlled to protect human health. US EPA, Public Domain.*

Table 1.1. A sampling of some important environmental regulations enacted in the United States since 1960.

Name of law	Summary
National Environmental Policy Act	Established guidelines for the construction of fede adverse environmental consequences.
U.S. Environmental Protection Agency (EPA) founded	A federal agency was created to monitor and regul substances into soil, air, and water.
Clean Water Act (amendments)	Established in 1948; amendments of 1972 brought s rules governing the amount and types of substanc waters.
Endangered Species Act	Established criteria for defining organisms at risk f guidelines for their protection.
Clean Air Act	Established rules governing substances released in Subsequence amendments expanded it.
Resource Conservation and Recovery Act (RCRA) AND The Toxic Substances Control Act (TSCA)	Established definitions and guidelines on record k transportation, and disposal of hazardous and othe environmental contamination; outlined roles and r waste generators and handlers.
Comprehensive Environmental Response, Compensation, and Liability Act (Superfund)	Established rules governing disposal and managem created a fund to help clean up polluted sites.
Marine Protection, Research, and Sanctuaries Act	Restricted and regulated dumping of potentially ha world's oceans.
Oil Pollution Act	Strengthened guidelines regarding oil spill clean u
Energy Policy Act	Established rules governing the manufacturing and power sources; set aside funds for development of

Three Mile Island nuclear power plant, near Harrisburg, PA, U.S.A., 1979

This extremely important event continues to exert influence over public policy and opinion in the United States. In short, an accident led to the release of some radioactive material into the surrounding area. Although it could have been much worse, the event was serious, sending shockwaves through the public, bolstering the anti-nuclear movement, and creating an enormous amount of suspicion about such facilities. In fact, no new nuclear power plants were brought on line in the U.S. after the accident (until 2023, as we will see in Chapter 10). Part of the Three Mile Island site was eventually brought back on line, and it continues to generate electricity today. Two other accidents in subsequent years—at **Chernobyl**, former USSR in 1986, and Fukushima, Japan in 2011—were far more destructive, leading to substantial releases of radioactive materials into Earth's atmosphere and waters. The accident in 1986 led directly to about 30 human fatalities, and thousands more are thought to be at risk of death from chronic diseases like cancer. The regions surrounding both sites continue to be sufficiently contaminated to pose health risks (even the 30-year-old one at Chernobyl). In addition, the 2011 accident has been linked to elevated levels of radiation in marine life. Worldwide response to these incidents was intense: some countries using it declared they will phase nuclear power out of their energy grid in the coming years, and others who do not currently use it say they will remain nuclear free indefinitely.

Bhopal, Madhya Pradesh, India, 1984

A chemical plant owned by Union Carbide released large amounts of methyl isocyanate gas (and probably others) that sickened over 500,000 people. On the order of 20,000 deaths have been linked to the event. The cause of the release is subject to some debate, with locals citing poor, even criminal, plant management, whereas corporate officials claim it was caused by sabotage. In any case, much legal action followed in the decades afterward; Union Carbide paid out a large sum of money and a handful of company employees were convicted of crimes. What the long-term health effects will be remains unclear, and environmental scientists continue to study it. Suspicion and animosity persist to this day.

Prince William Sound, Alaska, U.S.A., 1989

The oil tanker Exxon Valdez ran aground and released nearly 11 million gallons of oil into a pristine ocean ecosystem. Images of oiled beaches and animals led to intense public outcry, questions about the wisdom of oil exploration in environmentally sensitive areas, and increased scrutiny of the petroleum industry (Figure 1.3). Several changes in U.S. government policy regarding responses to oil spills were enacted as a result. In addition, environmental scientists took the opportunity to study the effects of oil on organisms and ecosystems as well as new strategies to clean up after such an accident.

Figure 1.3. *Images of the aftermath of the oil spill in Prince William sound led to massive public outcry: birds killed by the oil, left (a) and workers using hot water to clean beaches covered with oil, right (b). US NOAA, Public Domain (a); Exxon Valdez Oil Spill Trustee Council, Public Domain (b).*

In 2010, an explosion on an oil platform in the Gulf of Mexico caused human deaths and an even larger release of oil into the ocean than that seen in 1989, raising additional questions about the responsibility industry has to protect natural environments.

Bozinta Mare, Romania, 2000

A release of cyanide led to the deaths of hundreds of tons of fish in the Somes River; the toxin eventually made its way into the Tisza and Danube Rivers. The fishing industry was adversely affected as a result, and the usual questions and suspicions about industry were raised. Environmental scientists often study these types of events because they affect both natural and human economic systems.

Northern portions of Gulf of Mexico, U.S.A., ongoing

Releases of agricultural waste into the Mississippi River contribute to the formation of a so-called **dead zone** in the Gulf. The presence of certain fertilizers can cause rapid depletion of oxygen in water and the deaths of high numbers of fish (Figure 1.4 shows the results of a typical fish kill). This phenomenon has captured some attention during the past several years, particularly among people who make their living off commercial fishing. Environmental scientists devote a substantial amount of energy to studying this problem, and we will see more about it in Chapters 4 and 5.

Figure 1.4. *Fish die when their water is depleted of oxygen. US FWS, Public Domain.*

Declines in the population sizes of certain organisms, worldwide, ongoing

The previous examples on this list demonstrate how pollution can affect human health and economic wellbeing. A different kind of issue also is within the realm of environmental science, though: factors affecting the extinction of organisms that are *not* human. As we will see in detail in Chapter 6, many forces, both natural and of human design, can threaten the continued success and existence of organisms on Earth. Often, environmental scientists are called upon to study the likelihood that a particular species will go extinct and to make recommendations about how to protect it if it is judged to be endangered.

The potential that DDT has to bring about extinction of certain birds (described above) is one example of the way scientific data were used to protect organisms. In fact, bald eagles and other large predatory birds nearly became extinct in the 1970s, but the steps recommended by environmental scientists helped save them. Many other specific examples of such work could be cited, but they are beyond the scope of this textbook. What is important for us, though, is a common theme that connects most cases involving endangered species, namely, the conflict between protection and human activities (e.g., recreation, farming, manufacturing, development). Quite often a group or groups of people oppose putting an organism on the endangered species list because that action would require them to change their lifestyles in subtle or even dramatic ways. Questions about the relative importance of non-human organisms can lead to heated debates both in- and outside of legislatures. In these instances, environmental scientists may find themselves in an uncomfortable position as they present evidence of the likely adverse effects that a particular extinction could bring about to natural systems. People facing economic hardship or other losses in the name of protection of an animal or plant can be a rather difficult audience to persuade!

1.1.5. Its importance keeps growing

Environmental science is critical to public policy makers, academia, and industry. Among other consequences of the events described in the previous section, the United States government has passed many laws governing toxic and radioactive substances (refer back to Table 1.1). Public opinion certainly played an important role, though scientific data were indispensable to the development of legislation. Beginning in the 1970s, environmental science became an academic major that students could readily study in college, increasing the number of people working in the field. In the 1980s, a large industry grew to help affected companies interpret new and rapidly changing government regulations. The 1990s and 2000s saw additional needs and opportunities as problems related to waste management, energy resources, and air pollution became more and more urgent (we will learn about these and many other topics in upcoming chapters). It would be overly simplistic and inaccurate to suggest that the trend from the middle of 1900s to the current day can be characterized by ever-increasing (or even constant) public and governmental interest in environmental protection. To be sure, there have been ebbs and flows in the amount of energy devoted to these issues. However, it is fair to say that things are dramatically different now than they were sixty years ago, and environmental science has evolved and grown into a mature and highly relevant field.

1.2. FUNDAMENTAL PRINCIPLES IN ENVIRONMENTAL SCIENCE

The field of environmental science, as well as our approach to it, is built upon several principles. We will refer to these throughout this book, so you should become familiar with them early.

1.2.1. Earth is closed with respect to materials

This concept will recur more often than any other on our list. In short, it can be understood to mean that materials will neither enter nor leave Earth's system (for all intents and purposes—trivial amounts of material do come and go occasionally). Note that Earth is *open* with respect to energy (Chapter 4).

Two important consequences

1. **Material resources are finite.** We cannot count on additional inputs of any resources from outside of Earth; instead, we must rely on either what is currently here or on processes that can convert spent materials back into usable forms again. Clearly, many items are important to us, including nutrients, water, soil, breathable air, building materials, fuels—to name just a few—and our continued survival depends on an understanding of the factors controlling their availability as well as our effective management of their supply. This book will focus a great deal of attention on the many natural and human-based ways resources are made accessible to organisms, utilized, replenished, and recycled.
2. **Waste products must be managed.** Contrary to the language we might use when dealing with our trash, it does not go "away" after we pitch it into a can. In fact, unwanted waste can persist long after we are done with it if it is buried under ground. Even incineration, which can change its form, still releases products that must be dealt with. As we will see in detail in Chapter 13, appropriate waste management is critical to environmental quality.

1.2.2. We keep track of materials

Whether or not demands for vital resources can be met is very important to environmental scientists, and knowledge of factors affecting the whereabouts of scarce materials helps answer two fundamental questions.

Where is the material and how does it move?

Reservoirs. As we will learn in upcoming chapters, the materials we value and need can be found in a number of places on this planet. Put into language used by environmental scientists, materials are stored in reservoirs. Note that here we refer only to naturally occurring reservoirs, not tanks, buildings, and others constructed by humans. We will return to this point shortly.

Pathways. Earth is by no means static—materials do not just remain in their reservoirs. Instead, they move around, sometimes undergoing dramatic changes while doing so, via processes referred to as pathways (more in upcoming chapters). Figure 1.5 is an idealized way to visualize the reservoirs and pathways for any material.

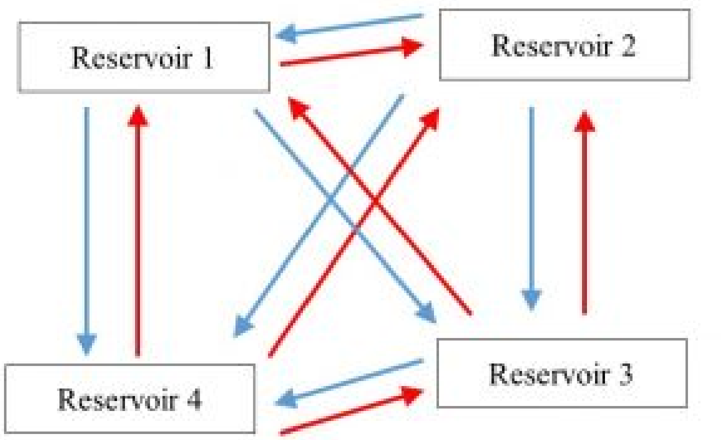


***Figure* 1.5.** *Materials are stored in reservoirs (boxes) and move via pathways (arrows). Kelsey,* CC BY-NC-SA.

Is there enough to go around? We often want to know about reservoirs and pathways because they affect availability of materials we desire. Imagine that you are designing a new housing development and want to know if the amount of water in a local reservoir will meet the demand of all the people who would live nearby. For the sake of argument, say your project is to be constructed near Reservoir 1 in Figure 1.5, above. You could measure how much water is currently stored there, but you need to account for declining availability that would occur after you start to tap into it (Chapters 4 and 11). Whether water demand can be met into the future depends on how quickly it moves along pathways from the other three reservoirs back to Reservoir 1. The wisdom and feasibility of making this investment in new housing would be heavily influenced by the data provided by environmental scientists working at the site.

1.2.3. There are limits on how many organisms Earth can support

Since Earth's physical resources, including space, food, and water, are limited, only a finite number of living things can occupy this planet at any one moment in time. Now, although some of the factors influencing that number are objectively measurable, there is still some disagreement about the upper limit for life on this planet, particularly how many humans can plausibly live here (more below and in Chapter 8). Nevertheless, scientists generally agree that we would experience adverse consequences if human numbers and activities exceeded the Earth's ability to support us. Here we introduce two concepts related to resource availability and what is needed to maintain populations of organisms.

Sustainability

This term is often used to describe phenomena in the realm of environmental science and also has taken on a certain amount of popularity among the general public. It is broadly related to whether a process or activity can continue indefinitely, or endure unchanged, into the future. The concept is often framed in human terms and posed as a question such as: how much of a given material can we use today without changing the access of subsequent human generations to that same material? The exploitation of vital resources such as water, fertile soil, fuels, animals and plants, and land can all be subjected to this assessment. Upcoming chapters will focus on how these and other issues can be viewed through the lens of sustainability.

Carrying capacity

Here we see an application of sustainability that is widely used by environmental scientists. It is defined as the maximum number of individuals of a species (more on the species concept will be presented in Chapter 5) that can live in an **environment** without reducing the ability of that environment to support the same number—with the same quality of life—in the future. Two examples will help illustrate this idea. In the first, we want to know the carrying capacity of white-tailed deer in a certain forest. We could quantify the amount of food, water, and space each individual needs, as well as how much of each of those requirements are met by the environment, and come up with a number. We might do our research, make our calculations, and conclude the carrying capacity in this case is 75 animals. If the size of the current population is higher than 75, say 100 for the sake of argument, the environment will be degraded, food will be eaten faster than it can regrow, and future deer will not have access to the same amount and quality of resources as today's (Figure 1.6).

Figure 1.6. *Spotted deer in India eat, or browse, all the vegetation they can reach and thereby exceed the carrying capacity of their habitat. Krishnappa,* CC BY-SA.

Numbers will therefore dwindle in subsequent generations. In the second we ask what is the carrying capacity of Earth for human beings? It turns out this question is far more difficult to answer than was the one about deer because there are so many unknowns. Although we could certainly determine how much food, air, water, and so forth are required to keep someone alive, we can only speculate about how much a typical person will use *above and beyond* their basic needs. Differences in standards of living across Earth, and just how many resources people will demand in the future, confound our ability to reliably calculate human carrying capacity. In fact, a rather wide range of answers have been offered by various researchers. Some conclude that the upper limit is about 2 billion people if everybody in the world lived a lifestyle akin to that of an average middle class person in a country like the United States, whereas a maximum of 40 billion could be sustained if every person used only enough resources necessary for basic survival (note the current population is about 8 billion people, and according to United Nations projections, is likely to grow to nearly 10 billion by about 2050[1]). Presumably, the answer lies somewhere between those extremes (close to 10 billion is a typical estimate offered these days), but, given the uniqueness of the human race, it is a difficult number to derive. We will explore much more about factors affecting the growth of Earth's human population in Chapter 8.

1.2.4. Organisms are linked to their environments

The title of this book contains the word "Environment" for a reason:

1. United Nations Department of Economic and Social Affairs, Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

it is a central concept, one critical to our understanding of living things. Here we begin to explore how it is important to our studies.

What is an environment?

This word is often misused and misunderstood. Contrary to popular belief, it does not mean all the forests, lakes, oceans, and rest of Earth's outdoor places, nor should it be used to refer to natural resources or other non-human spaces. Its formal definition is composed of two major points.

1. **An environment is a space.** Simply put, it can be thought of as the location in which an organism lives, or at least temporarily resides.

2. **An environment is made up of those things that affect an organism.** In addition to providing the space in which it exists, it influences and shapes an individual. It can be visualized as the set of forces that applies pressure, or stress, to living things.

So, now we know that when someone says "I care about *The Environment*" they should instead say something like "I care about *Earth's natural environments*". The important point is: there is no single environment. Instead, many, many environments can be identified. We will see some specific examples of entities that can be categorized as environments shortly.

What environmental factors influence organisms?

Some of the specific environmental factors of importance to organisms are temperature, presence of sunlight, moisture availability, types and amounts of food sources, altitude, chemical properties, and the presence and activities of other organisms. These and other variables control which organisms are present in a particular place. One of the underlying principles for us is the notion that organisms do what they do in response to their environments—put another way, they **adapt** to the conditions in their habitat. One obvious example would be the strategies fish have adapted to live in water: gills, fins, and so forth allow them to succeed in their environment. Oak trees provide another helpful example. They are adapted to environments with abundant rainfall and rich soil, whereas cacti can tolerate and even thrive in dry deserts, places which would be deadly to the oaks. Lions, seals, snakes, sharks and the rest of the organisms on Earth all have successfully responded to the pressures applied to them—their continued existence here suggests as much. Furthermore, since environments are not constant, the most successful organisms are the ones that can change as necessary. Those that cannot alter their life strategies will fail, likely becoming extinct (more in Chapter 6). A final note is worth considering here: environments play critical roles in shaping the development and defining features of human beings as much as they do other organisms. In our case, environment is quite complex, containing the basic features related to survival listed above, but also includes social forces like family, friends, economic conditions, education, and the like. Who you become, how you appear, and what you ultimately do, are all influenced by those forces acting on you during your life. Of course, innate features affect organisms as well, as we will see in Chapters 5 and 6.

Can organisms affect their environments?

The short answer to the question is: yes. Just by its presence, an organism will change the place in which it lives. Animals eat plants, sometimes take down entire trees, and burrow in soil, and plants remove nutrients and water from soil in which they are rooted. Organisms even affect the composition of Earth's atmosphere, adding biologically usable oxygen gas to it (see Box 1.1 for more about this crucial process and its consequences). And the list goes on, containing both familiar and unfamiliar ways environments are shaped by living things. We will see many examples of this phenomenon as we make our way through this book.

Box 1.1. Enjoying your air?

One of the most fundamental environmental changes brought about by organisms is related to the Earth's atmosphere. As you likely realize, many organisms, including we humans, require **dioxygen gas** (O_2) for basic survival. The fact that breathable oxygen has *not* always been available on Earth might surprise you, though. How do scientists know anything about conditions of the past? Through many studies of evidence stored in rocks, soils, and fossils, they have concluded that the composition of the early atmosphere was different than that of today. Notably, that crucial O_2 gas was not present. Yes, there was oxygen here, but it was stored in reservoirs other than the atmosphere. Now, some organisms, known as **anaerobes**, do not need oxygen to survive—in fact, they are poisoned by it—so they dominated the planet in the early days. Put another way, they were well adapted to their no-oxygen environment. Over two billion years ago, certain ancient organisms that release oxygen gas as a product of their normal activities began to appear (more about these organisms is presented in Chapter 5). It took a very long time, as we will see in Chapters 3 and 6, but eventually the atmosphere was altered such that it contained appreciable amounts of oxygen. Then what happened? Well, organisms adapted to life without oxygen were in trouble. They had to change their strategy, find environments that were free of oxygen (such places still exist today), or go extinct. Those that could use oxygen, **aerobes**, gained a big **advantage** and began to develop from small, relatively simple individual cells, into large, complex, organisms such as animals and plants. The tremendous diversity of living things we see today (Chapter 6) is very much a product of widely available oxygen. Again, organisms affected their environments, and those changes in turn affected other organisms. The relationship between the non-living components of Earth and life is complicated, dynamic, and one that brings about changes on both sides.

Where can environments be found?

They are everywhere. Any entity that meets the criteria we listed above, that is, it surrounds and influences organisms, can be defined as an environment. Size, shape, complexity, or whether it is natural or the result of human actions are irrelevant considerations. In other words, the possibilities are seemingly limitless. Since the space available to us in these pages *is* limited, though, we will consider only a small number of examples

A forest. To many people, this or a similar place comes to mind when they hear or use the term "environment", for it contains all the components commonly associated with the (incorrect!) usage of the term: trees, animals, streams, and natural processes (Figure 1.7). To us, though, it is a *type* of environment with characteristic features and conditions appropriate to support the organisms living there. A list of the specific and influential environmental properties found in a forest in, say, northern New York State (U.S.A.), would include a land-based area (vs. under water), abundant rainfall, temperatures that range from near 38 °C (about 100 °F) in the summer to -20 °C (about -5 °F) or lower in the winter, and sunlight that varies seasonally. The living things present there will be those adapted to the prevailing conditions. Now, if one or more environmental properties change, the organisms that dominate will also change (Chapters 5 and 6).

Figure 1.7. *An example of a forest environment. Kyle R. Burton, CC BY-SA.*

A desert. Although also land based, this environment is very different from a forest. Importantly, a desert receives far less precipitation (Chapter 5). Wide temperature swings between very hot days and cool nights are also likely. Instead of trees, dominant organisms would include cacti, sagebrush, and animals that are able to survive there (Figure 1.8).

Figure 1.8. *An example of a desert. Note how different this land-based environment is from that shown in Figure 1.7. US FWS, Public Domain.*

A lake. Clearly, this water-based environment is fundamentally different from the first two on our list (Figure 1.9). Here organisms will be adapted to obtaining food and oxygen from relatively still, fresh water (not moving, as in a river, or salty, as in an ocean). If this is a northern lake that freezes in the winter, it will be dominated by different organisms than one in, say, Sub-Saharan Africa, a place that does not freeze.

Figure 1.9. *An example of a lake environment. Doris Antony,* CC BY-SA*.*

A classroom. Although not naturally occurring, this space meets our criteria for inclusion here: organisms can live in it (some non-human ones even live there full time) and are influenced by its conditions (Figure 1.10). Properties related to temperature, humidity, odors, light, and food availability could all be described. A room in the basement of an academic building might be damp and warm, encouraging the growth of mold on its ceilings (as evidenced by brown or black splotches). A room on the fifth floor of the same building could be dry and hot, discouraging mold but supporting flies and other insects. Both areas would be teeming with unseen **bacteria** (see Chapter 3 for more about these ubiquitous and important organisms), although the different conditions indicated would likely mean they would not be home to the same bacteria. What about the humans who come and go? Well, at a minimum the conditions are (we hope) appropriate enough for them to survive there for a few hours a week. Proper design would create an environment most conducive to learning. Chalkboards, technology, an instructor, and students also influence humans in the room. All these variables could contribute to different learning outcomes for different people. Put another way, environmental factors play an important role in shaping all organisms, even those spending time in artificially constructed spaces.

Figure 1.10. *A familiar-looking environment? Kelsey*, CC BY-NC-SA.

A person. The final example on our list might surprise you: an individual human being is an environment for large numbers of microscopic organisms (i.e., they are too small to be seen without the aid of a magnifier). Even a completely healthy person houses—in fact, depends on—trillions and trillions of **bacteria** and other **microorganisms** (see Chapter 3) on a regular and on-going basis.

Services performed by microorganisms for us

Digestion of certain foods. The human digestive tract is the home for microorganisms that can help us obtain nutrients from some of the plant material we ingest. In fact, without the help of these tiny partners, we would not be able to digest this food (more about many types of relationships among organisms can be found in Chapter 5).

Protection against some illnesses. Our bodies are attractive places for certain invading microbes, termed pathogens, that make us sick (see Chapter 3). They can find their way inside of us by multiple means, but most of the time enter through one of our permanent openings to the outside (e.g., mouth, nose, urinary / reproductive systems, eyes, ears) or through temporary wounds. Luckily for us, we have developed relationships with *other* microbes that occupy the tissues near some of those openings. Our mouths, for example, are absolutely teeming with these helpful microorganisms. Their presence makes it very difficult (but not impossible) for disease-causing organisms to get very far into our bodies; the locals are so well adapted to the conditions in their environment, as well as sufficiently abundant, that invaders often are simply outcompeted and fail to make us sick.

Factors affecting resident microorganisms

The specific identity of the occupants inside us depends in large part on, you guessed it, the environmental properties of our bodies. Those with the appropriate adaptations can thrive under the temperature and chemical conditions of a human. Unsurprisingly, then, changes would likely affect organisms living in our guts and other places. Keeping these ideas in mind, consider how we respond to a disease such as strep throat. When a certain species of bacteria overcomes our defenses it can cause a reasonably unpleasant ailment. It usually produces a tender, swollen, red throat and a fever. Why do we experience these symptoms? First, bacteria from the outside outcompete our normal resident bacteria, take up space, absorb nutrients, and generally irritate their adopted home. Second, fever is one way your body manipulates environmental conditions to inhibit the growth of harmful and unwelcome invaders. A pathogen adapted to survive at human body temperature (37 °C) will struggle and even die if subjected to a habitat maintained at 38 or 39 °C. Additionally, people sometimes ingest chemical substances known as antibiotics to kill bacteria. In other words, they make a change to the chemical conditions of their bodies to thwart unwanted organisms. Unfortunately, antibiotics can also harm many helpful bacteria, making it difficult to digest food and otherwise maintain good health. As long as those changes are temporary, though, most people are willing to accept the negative consequences (diarrhea is possible, for instance) in the interest of a cure. We will see more about the highly nuanced and fantastically important story of bacteria and **antibiotics** in Box 5.6, Chapter 5.

Everything on Earth is connected: environmental unity

Central to environmental science is the idea that, although we tend to break it up into smaller components for the purposes of study, Earth is one large and complex environment made up of interconnected and interdependent units. We can certainly define and describe the organisms or air or water on the planet, but we understand them to be parts of a greater whole. Furthermore, since everything is connected to everything else, a change in one component—no matter how small—will bring about changes to other, seemingly distant, components within Earth's systems. You should realize that predicting the far-reaching effects of a relatively small-scale action can be quite difficult, and we often are stuck retracing our steps only *after* an unexpected, and very likely undesirable, change has been initiated. We will apply and see examples of this idea, the **principle of environmental unity**, in upcoming chapters.

Decision making is informed by caution and risk assessment

Consequential decisions about our environments and health, such as whether to proceed with the construction of a new highway, approve a new vaccine, use a new pesticide, or allow a new preservative in food, generally are based on two fundamental premises. First, new chemicals, technology, and the like are assumed to pose a risk to natural systems (as well as humans) until shown to be safe. Although this starting point, known as the

precautionary principle, is not embraced by all environmental scientists, public opinion and regulations tend to be driven by it. Importantly, it lays the burden of proof on those who advocate (and are likely to profit from) the use of those new potential hazards. There is some debate over the appropriateness of this principle, as certain people view it as a waste of money and effort. Advocates of it, on the other hand, point out how its conservative approach generally protects public health. Second, some risks are acceptable as long as they are outweighed by the good an action brings about. A **cost-benefit analysis** is critical to the evaluation of the advisability of any initiative or change (see Box 1.2 for a timely application of this tool).

Box 1.2. The COVID-19 vaccine: fun with cost-benefit analysis!

The COVID-19 pandemic of 2020 (and beyond) likely needs no introduction. It brought about much suffering, disruption, and death, but it also led scientists to perform what was believed by many to be impossible: they developed multiple safe and effective vaccines in a matter of months, instead of the many years that is typical for such work. Those who were vaccinated have largely been able to resume normal activities with scant risk of contracting a serious case of the disease (i.e., after vaccination, a small number of people came down with mild illness, but life-threatening cases requiring hospitalization fell to near zero). Those who did *not* get the vaccine continued to be at substantial risk of a debilitating or lethal case of COVID. In fact, after vaccination began, unvaccinated individuals accounted for an increasing percentage of the life-threatening cases treated in hospitals. So, what motivated this second group? Why expose yourself to avoidable danger? The answer is complicated. Of course, some could not tolerate the vaccine for medical reasons, but many eligible people chose to skip the vaccine because they believed, despite a lack of supporting evidence, it would lead to some kind of dire consequences (including, but not limited to, loss of fertility, loss of personal freedom to unscrupulous scientists and government officials, and unforeseen health effects of a novel vaccine). Put another way, they felt its risks outweighed its benefits. Let's interrogate their assumption about risks with data and do a quick cost-benefit analysis

To simplify things, we will look only at the situation in the United States. According to governmental data, there were 33.5 million cases and 600,000 COVID-19 deaths as of June, 2021 (**an average of a bit under 2% of those infected died from the disease**, although factors such as age and pre-existing conditions raised that number to 10% or more in some groups)[2]. To some, that rate seems trivial (families who lost loved ones likely view it differently). Serious long-term yet non-lethal effects of COVID are also possible (so-called 'long COVID'), however, including organ damage, blood clots, chronic fatigue, and ongoing respiratory distress (not to

2. covid.cdc.gov/covid-data-tracker

mention quirkier symptoms like loss of sense of taste and smell that can persist for months). Furthermore, by some estimates, **risk of death in the months following infection rise precipitously**. The benefit of receiving the vaccine can be summed up, then, as substantially reduced risk of death and other adverse consequences.

What are the dangers of the vaccination? If we consult data provided by the U.S. Centers for Disease Control and Prevention (U.S. CDC)[3], we see that the risks of it are quite low. **Somewhere between 0.0002% and 0.0005% of people who got one of the vaccines suffered some kind of anaphylaxis shortly after getting their shot**. About 0.0003% of those receiving the one-shot Johnson and Johnson vaccine were affected by abnormal blood clotting (36 individuals out of 11.7 million vaccinated). **Overall, 0.0017% of those receiving one of the vaccines died afterward** (somewhat more than 4000 out of 117 million). This statistic about mortality is not quite what it seems, though, because rules governing medical facilities require that deaths from *any* cause after vaccination must be reported. So, car accidents, drug overdoses, gunshots, drowning, cancer, and other agents are included in that number. In other words, there is no reason to believe that the vaccine was responsible for all or even the bulk of the deaths. Famously, many people did experience short-term

3. cdc.gov/coronavirus/2019-ncov/vaccines/safety/adverse-events.html

discomfort following their shot, a problem that is more troubling to some than others (in any case, the symptoms were very rarely linked to real degradation in health status).

To summarize (doing a little math): being very conservative and cautious, the risk of an unvaccinated person dying after contracting COVID-19 is **at least 1000 times higher** than the risk of dying after receiving the vaccine. You decide which course of action is likely to cost less.

1.2.5. Humans and Earth operate on different time scales

As we will see in detail in Chapter 3, scientific evidence suggests that Earth is on the order of 4.6 billion years old. Primitive human-like beings, on the other hand, arrived on the planet relatively recently and have been around less than 3 million years. Modern humans have only been present for some 200,000 years. Depending on how you want to define it, evidence of the first recognizable civilization dates back in the thousands of years, with the modern era beginning a century or so ago. Of course, an individual human generally lives less than a hundred years, and much of the business of living is measured in days, hours, and minutes. Given the constraints on our perspective, then, it can be difficult to comprehend the vast amounts of time necessary for many natural forces to shape Earth's environments. In upcoming chapters we will study, among other topics, the formation of continents, oceans, rocks, and oil as well as the development and evolution of living things, even though they generally happen so slowly we cannot observe them in real time. Human existence is just too limited to directly experience much of what occurs on Earth.

The story gets more complicated, though. Despite the short history of our species and the very brief time each of us is alive, humans have become a powerful agent of change on Earth. The pace at which we modify our environments, use and move materials, synthesize novel substances, and generate waste products can be extremely rapid, dramatically outpacing analogous natural processes that have been active throughout the planet's long history. Environmental scientists try to assess the potential short- and long-term consequences associated with human actions in the context of the disproportionately large ability we have to quickly bring about changes.

THE CHAPTER ESSENCE IN BRIEF [4]

Environmental scientists study Earth's systems and the forces affecting them. It is objective, relevant, and founded on essential principles about the interconnectedness of living and non-living phenomena. Chapter 1 lays the important foundation for the rest of this textbook.

---

4. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 1 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

1.3. MOVING AHEAD

We end Chapter 1 with a preview of things to come as well as some suggestions about how to maximize the usefulness of this textbook.

1.3.1. What to expect next

This book is divided into three parts. The first consists of only this and the next chapter and is included to provide an overview of the scope of environmental science (you just read that part) and some tools we will use throughout our study (Chapter 2). The second part (Chapters 3 - 6) describes the basic processes that are active on Earth, with little attention given to any sources of stress (human or otherwise) affecting those systems. The third and final part examines many forces that can disrupt the basic processes described in Part Two. The majority of these later chapters focus on human activities, although some coverage of natural sources of stress are given in Chapter 7.

1.3.2. What you can do get the most out of your reading

This textbook is designed to present environmental science as a coherent story, and you should approach it with that in mind. It is certainly not a work of fiction, but you should still try to follow it more like it were a narrative than as a collection of disjointed facts. To be sure, you will encounter a lot of terms, definitions, and descriptions of concepts new to you. These are presented to enable you to learn the language of environmental science so you can both understand the story as well as get to a point where you can tell the story yourself. You are encouraged to look for connections among the topics within and among chapters; use the cross references provided to help you do this. For example, when you are reading Chapter 5 you will see extensive connections back to Chapter 2. If you look at them together you will both remind yourself of important concepts as well as broaden your understanding of the subject matter. Environmental science is fairly hierarchical, that is, early chapters present foundational concepts upon which later chapters depend. You could jump around and read things in a different order, but if you keep the intended organization in mind while you do so, you will enhance your ability to make sense of things.

Think about it some more...[5]

Are environmental advocacy groups best categorized as environmental scientists or environmentalists? Is it possible they could be both? Neither?

Why do environmental scientists keep such careful track

5. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

of the reservoirs and pathways of important materials? What fundamental principle informs this work?

What is the nature of the relationship between organisms and their environments? Could that relationship change with time?

Could the deaths of thousands of sea birds in Alaska be connected to the construction of a new housing development near, say, Atlanta, GA? Think about the principle of environmental unity as you ponder your answer.

2. Science and Systems Analysis

JASON KELSEY

As the words "science" and "systems" are part of the title of this book, it seems reasonable for us to examine them before we proceed any further. Why are they important enough to earn such prominent placement on the cover? In short, **science** and **systems analysis** provide a framework and tools that we will use extensively in our exploration of Earth's natural systems. These concepts also are useful in the study of many subjects outside of the sciences. In other words, knowledge of them has broad, inherent value.

Key concepts

After reading Chapter 2, you should understand the following:

- How science is a distinctive way of knowing
- How evidence-based conclusions differ from those derived from opinions, values, and other subjective claims
- What can and cannot be assessed with science
- How the scientific method is used to objectively collect data

- Scientific conclusions, theories, and laws
- The sources and expression of measurement uncertainty
- The distinctions among accuracy, precision, and sensitivity
- How systems are defined and used to study many kinds of phenomena, including environmental processes
- The nature and importance of feedback in the functioning of systems
- How input-output analyses are important to environmental scientists

Chapter contents

2.1. SCIENCE

2.1.1. Science is a distinctive way of knowing

Put simply, **science** is a way of knowing, *one* approach to a study of the universe (or, as scientists often say, **the natural world**). There are other ways of knowing, and science should not be held up as the only legitimate one. However, it is important to recognize the ways science is distinguished from those others and, critically, when it is the proper choice as a tool for analysis.

Our formal discussion of science will begin with a description of what it is *not*. First, although it is common for people to cite "scientific proof" when they attempt to make a convincing argument, **science is not actually able to prove anything**. Instead, it presents evidence and describes how certain we are about conclusions and ideas. Degrees of certainty are vital to the practice and understanding of science, as we will see in more detail shortly. Second, **science cannot be used to assess subjective ideas** or evaluate notions about what is good or bad and right or wrong. It is not values or faith based and cannot (and should not!) be used to try to somehow answer questions about the meaning of life, the quality of a musical composition, the morality of cutting down trees, or even the existence of God. It is not the right tool to answer aesthetics questions such as "do these pants look nice on me?" or "is this painting good?". We also should not ask scientists to comment on environmental questions such as whether an endangered species or ecosystem has a right to exist, rather, we can draw upon science to predict the likely consequences to natural systems if a particular organism disappears and use it to inform our decision making. In other words, although science is very powerful, it is also limited.

What then, *is* science? What *can* it do? In short, science is an evidence-based approach used to answer testable questions. Its fundamental goal is to reveal generalizations—the rules, if you will—governing the universe. It is distinguished from other ways of knowing and evaluation by some important characteristics.

2.1.2. Science is objective

Scientists seek to increase our understanding of the natural world. They design experiments and record observations but do not interfere with the way experiments unfold. Crucially, personal bias, values, or preconceived notions about what *should* be true, what the answer to questions of interest *ought* to be, are irrelevant. Ideally, such emotional involvement is completely absent. Since science is carried out by humans, **subjective** judgments can affect the way real-world science is conducted. Accordingly, potential conflicts of interest or bias should be acknowledged and described in all scientific studies.

2.1.3. Science requires evidence

Unless data can be objectively collected, science is not the proper tool to study a particular phenomenon or question. In other words, explanations that require inference, intuition, faith, or indirect connections between cause and effect are not in the realm of science (see Correlation vs. Causation, below, for more about the relationship between cause and effect). Science is only appropriate if an idea can be **tested through observation**, that is, some kind of experiment that allows for the collection of empirical data is required. Now, experiments vary widely in their nature and ability to mimic reality, but they are still a necessary part of scientific inquiry. Chemists, physicists, geologists, biologists, and ecologists collect data under vastly different circumstances using experiments and tools that are unique to their disciplines. Some primarily observe existing natural systems, whereas others conduct controlled experiments inside laboratories. Some scientists rely entirely on computer models. In all cases, though, they use objective methods and what is commonly referred to as the **scientific method** to gather evidence. We will consider more about what constitutes an experiment and how experimentation is a part of the scientific method shortly.

> ***Empirical*** *refers to data that were obtained through*

direct observation. Any conclusion derived from such data would therefore be based on what we call ***empirical evidence****, not mere conjecture or speculation.*

2.1.4. Science uses both inductive and deductive reasoning

As we learned above, science is used to study the natural world, but there are multiple ways to carry out such inquiries. Scientists may examine representative cases—they collect individual data points—and then attempt to establish a plausible and widely applicable explanation for what has been observed. In other words, science often employs **inductive reasoning**: it uses specific observations to establish new generations. Imagine a very simple example in which we study the behavior of falling objects. To accomplish this, we could set up an experiment in which we throw various things out of a seventh-floor window and record what happens. We might begin with a computer, and then follow with a chair, a piece of chalk, a coffee mug, and a melon. Assume we drop twenty objects and they all smash pleasingly on the pavement below. After a while we would likely tire of this exercise (or simply run out of objects to hurl) and decide that we have enough data to try to establish a general rule about all falling objects based on our limited number of specific observations. We could then develop ideas about gravity and so forth that we would assume are broadly applicable. Note that the degree to which the new generalization describes every possible case will depend on how representative our specific observations are. In other words, if the rule we establish is to be valid, the objects we drop from the seventh floor need to behave like anything else that falls to the ground. If we chose live birds, helium-filled balloons, soap bubbles, or flat sheets of paper, our observations would not apply to all relevant situations and could lead to potentially disastrous conclusions. We will come back to this and other potential problems affecting experiments when we discuss uncertainty, below. On the other hand, **deductive reasoning** is the opposite of inductive reasoning in that it is used to test if specific cases follow pre-determined rules (that is, it progresses from the general to the specific). For example, the rules of geometry tell us if a given unknown shape is a triangle, square, circle, etc. Scientists could also use deductive reasoning to test whether a hypothesis or prediction is supported by experimental evidence, that is, specific cases that are analyzed in an experiment. In this latter instance, we assume the generalizations are true until shown to be otherwise.

2.1.5. How it is done: the scientific method

The scientific method is a systematic approach that has been used for centuries in countless studies of the natural world. Classically and formally, it consists of several stages, although it is worth noting that the scientific method as described here represents an idealized version of the way science is conducted. All experiments do not necessarily follow each and every step in the order shown, but they broadly use the approach outlined here.

1. Observation

A scientist notices a phenomenon and decides to study (and, they hope, explain) it. Typically, this is a moment when a question that starts with "how does…" or "why does…" or "what is…" is formulated. It can be a question about anything, but it is likely something that seems important or interesting to the scientist asking it.

2. Hypothesis

The scientist develops a plausible answer to the question posed in step 1, above. In other words, this is the step in which a hypothesis is formulated. Put simply, a hypothesis is a proposed explanation for the observed phenomenon of interest. People sometimes refer to this as a "guess", although such a characterization is not entirely appropriate. Usually, a hypothesis is put forth only after careful thought and review of what is currently known about related phenomena.

3. Experiment

Experiments are designed and conducted to test the validity of the hypothesis. As we saw above, there are many different types of experiments, but they all involve observation and data collection. Scientists examine data and draw a conclusion based on them. The conclusion reached is an answer to the question posed in step 1. At this stage, the answer suggested by the experimental data is compared to the hypothesis, and the hypothesis is either supported or refuted. Remember that science cannot provide absolute proof that anything is true, instead, it can make a case that a proposed explanation is appropriate by supporting a hypothesis multiple times under differing conditions (interestingly, when experiments do *not* support a hypothesis, we typically say something like "the hypothesis is *disproven*"). For example, one might measure how the amount of sunlight received can influence the growth of a certain plant. A scientist would start out with a hypothesis that is something like "plant growth will depend on sunlight received; both insufficient and excess light will adversely affect the height a plant reaches within two months." The experiment would involve groups of plants exposed to different amounts of sunlight (say, nine groups ranging from complete darkness to continuous light, with ten plants in each group—i.e., light is the **variable** we assess). The plants would otherwise be treated identically (same amount of water, soil, etc.). The scientist would measure the heights of the plants each week for two months. At the end of the experiment, a conclusion about the relationship between light received and plant height would be formulated. The validity of the initial hypothesis would also be evaluated given the data collected (Table 2.1). Note that an important part of our conclusion is that different amounts of sunlight *caused* the observed differences in growth; the two were not simply *correlated* (for more, see Correlation vs. Causation, below).

Table 2.1. *Effect of sunlight on plant height*

Light (hours)	Height (cm)
0	0.0
3	4.1
6	7.3
9	15.2
12	17.8
15	21.6
18	18.5
21	14.6
24	9.2

4. Retest or refine a hypothesis

If the initial hypothesis is supported by the data collected, the experiment should be conducted again. An experiment must be repeated multiple times, and data must be compiled from many studies before scientists can begin to make generalizations about the phenomenon of interest. In the plant experiment imagined in step 3 we could analyze the data and conclude that there is an optimal amount of light exposure for plant growth (15 hours a day in this case, as shown in Table 2.1). In other words, our hypothesis seemed to be valid. The next step for us could be to carry out the whole experiment again and compare the findings from the two studies. We might decide to repeat this process many times. If, however, the hypothesis is refuted, or only partially supported, we would have to revise it.

5. Communication and review of results

Once scientists have repeated an experiment and reached some conclusions about the phenomena under investigation, they must then communicate that information to the scientific community. Such sharing enables other scientists to scrutinize and repeat the experiments on their own to independently test the conclusions drawn from them. Results can be shared informally—among colleagues at the same or different research institutions—or formally, through a published paper or oral presentation. An important part of formal communication is what is known as the **peer review process**. Researchers wishing to publish in a scientific journal, for example, generally submit their manuscript to an editor. Scientists with the appropriate education and experience then review the submission to determine if it should be accepted for publication (generally, they insist on a number of revisions to the manuscript in any case) or rejected. Oral presentations are vetted to varying degrees, but a scientist presenting data to an audience will be subjected to questions and expected to justify and defend the work most of the time. Science is therefore scrutinized and challenged, requiring researchers to objectively and carefully study the natural world. Put another way, whether as formal reviewers or simply readers of scientific journals, scientists are expected to be skeptical of conclusions presented to them—they should not accept any new idea until they have studied and challenged it thoroughly. Of course, all this rigorous checking and review mean that our understanding of the universe advances gradually. Contrary to the way they are often portrayed in popular culture, scientists very rarely exclaim "Eureka!" or the like. Sudden breakthroughs and discoveries are the exception, not the rule.

A ***manuscript*** *describing methods used, background information, results, and the relevance of the findings is submitted to a journal editor who then asks scientists in the field to evaluate it. If it gets published, it is an* ***article****. Newspaper stories, opinion pieces, and scientific reports that are not rigorously reviewed should be read with caution.*

2.1.6. Correlation and causation are not the same

When we are conducting experiments to test the effect of some variable on the behavior of objects or organisms it is important to distinguish between correlation and causation. **Correlation** refers to the way two events occur at the same time but suggests nothing about the ways one of the events influences the other. **Causation** implies that the two events under investigation are connected and that one led directly to the other. Although it can be tempting to assume a stressor brought about observed changes just because two events are simultaneous (or sequential), science strives to actually establish a link between cause and effect. Imagine we decided to study the relationship between the presence of a cigarette lighter in one's pants pocket and the development of lung cancer. The data would show that cancer occurs in people who carry lighters more often than in those who do not (see Figure 2.1).

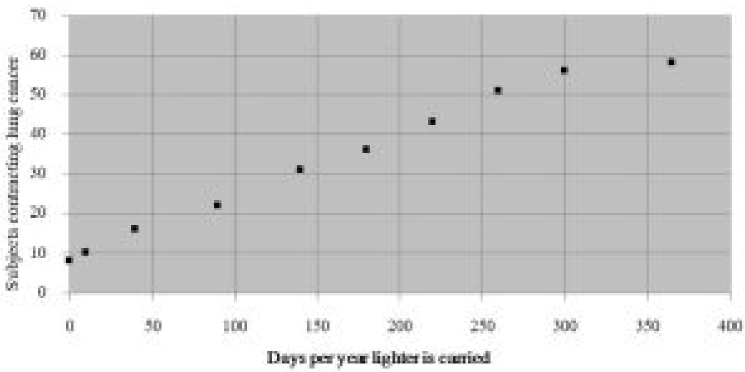


Figure 2.1. *Hypothetical relationship between frequency of carrying a lighter and likelihood of developing lung cancer. The data only suggest correlation, not causation. Kelsey*, CC BY-NC-SA.

We can see that lung cancer is correlated with the carrying of a lighter. It is likely not surprising to you, however, that lung cancer is not *caused* by the presence of a cigarette lighter, rather, lung cancer is directly linked to cigarette smoking (which is tied to carrying a lighter). A less fanciful example can be seen in the debate about the relationship between childhood vaccinations and autism. Although some children start to exhibit signs of autism shortly after receiving the MMR vaccination, there is no evidence that the vaccine brings it on. In fact, numerous studies conducted to test whether or not vaccines cause autism have repeatedly shown that although sometimes correlated in time, there is no mechanistic link between the two. See Box 2.1 for more on autism and vaccination.

Box 2.1. Correlation is not causation: the case of vaccines

Whether phenomena are linked by causation or only correlation is an important consideration, as a misunderstanding could lead to the misallocation of resources to treat a neurological difference such as autism. In other words: because of efforts by a small yet vocal group of people who are adamant that childhood vaccination is responsible for autism, money, energy, and time are diverted away from research that could reveal its actual cause. Furthermore, and arguably more problematic, some people who think autism is brought about by vaccination refuse to have their children immunized against a wide range of preventable illnesses. The result can be a public-health crisis. For example, an outbreak of measles in some California public schools (and Disney theme parks) in early 2015 was driven by low vaccination rates. Many of the parents of unvaccinated children decided against inoculation because they believed it posed health risks. Objective interpretation of data, though, has led all but a few scientists and physicians to conclude that the risk associated with contracting measles is far higher than that of receiving the MMR vaccine. Review Box 1.2 for more about cost-benefit analysis.

2.1.7. Experimental vs. observational science

In addition to helping uncover the rules of the universe, the scientific method can be used to classify and identify components of the natural world. These two uses of science are referred to as **experimental** and **observational**, respectively.

Experimental science

Here, scientists assess the effects of stimuli on test subjects (objects, organisms, etc.). The experiment described above involving the relationship between light received and plant growth is a good example of experimental science. Note that the responses of test subjects must be compared to **control** subjects, those that go untreated. A group of animals that does not receive an experimental drug, say, but otherwise is identical to those that do, would be a good example of the use of **controls**. As we have seen, new generalizations about the ways the universe functions are developed with experimental science.

Observational science

In this case, scientists collect information about a test subject at a moment in time. The effects of changing conditions are NOT assessed in observational science, neither are the subjects influenced by or exposed to stressors during these types of studies. Observational science can be applied to many situations. For example, we could study the long-term effects of smoking on human health by surveying a group of people to assess any connections between smoking (how much per day, number of years a person smoked, etc.) and the occurrence of diseases such as cancer and heart attack. Remember, the test subjects would not be given cigarettes and then observed for years afterwards, instead, we would study their health status today in light of the smoking they did previously (you might imagine that such a study is fraught with difficulty and uncertainty, in part because we must rely on test subjects to report their smoking habits). Observational science could also be used to determine public opinion (e.g., people are polled about how they will vote), or to classify birds, rocks, and other objects into appropriate groups. In all cases, test subjects must be evaluated (through questioning or other means) in an objective way and should not be manipulated or influenced such that their present condition is altered.

2.1.8. Conclusions, theories, laws: just how confident are we?

As we have seen, experimentation leads to notions about the rules governing the universe. You should recognize that our understanding evolves as more and more research is conducted, and our level of confidence about this understanding evolves similarly. For clarity, scientists use three basic terms to indicate the level of certainty associated with a scientific idea. The first, **conclusion**, is the most tentative of the three. It is derived from the results of a small number of experiments and is expected to change as additional data are collected. Conclusions can be applied only narrowly. With more time and study, and repeated challenges and revisions, a conclusion can be elevated to the second level, that of theory. **Theories** are widely accepted explanations of important phenomena. Two of the most well-known theories are that of **plate tectonics** and **biological evolution** (Chapters 3 and 5, respectively). Both have been developed, studied, and scrutinized by many generations of scientists, are supported by a lot of data, and are thought by most to plausibly describe the development of Earth's surface and its life, respectively. These and other theories have withstood over a century (or more) of close examination and attempts to refute them and yet continue to be supported by data. Despite their longevity and apparent validity, scientists do not stop studying and revising theories, however. In fact, theories could be shown to be invalid, or at least change, with enough data. Finally, after even more time and many more experiments designed to test a theory, some ideas become so well entrenched that they are elevated to the highest level, that of **scientific law**. Because the criteria that must be met are so rigorous, relatively few theories have become laws. The **laws of thermodynamics** (Chapter 4) and gravity are among the important examples of the types of generalizations thought to be worthy of this categorization. Few think they will be refuted, although scientists must always allow for the possibility that our understanding of even these most foundational principles could change with continued study.

2.1.9. Scientific measurement and uncertainty

Measurements in experimentation

Tools, or **instruments**, are used in science to quantify (assign a number to) properties including length, temperature, mass, volume, and chemical composition of the objects they study. For instance, one could respond to the question “How long is this tree branch?” with an answer like “The branch is 3.2 meters (m) in length.” Notice that the length is expressed with both a number and unit. The unit given is universally accepted and recognized by all (or nearly all) scientists. For the length of interest in this case, the appropriate tool would likely be a meter stick. As we will see shortly, a meter stick is not the best device with which to measure the length of all objects (see Accuracy, Precision, and Sensitivity, below). Thermometers, balances (essentially the same as scales), and other instruments assess various properties. Although a range of tools can be used to quantify many different characteristics, all measurements have something important in common: they represent approximations of the actual values of interest. In other words, every measurement has some amount of **uncertainty** associated with it. Even the most experienced scientist working under the best circumstances must still contend with the reality of scientific **uncertainty**.

Uncertainty happens!

Imagine that you are hosting a dinner party and serving wine to your guests out of a glass carafe (you secretly combined the remains of several opened bottles of wine you had lying around, some of which are less than spectacular). Before pouring the last round, you want to determine the volume of wine remaining. In the kitchen you transfer the wine to a cup measure and conclude that there are a little under 4 cups (945 mL) in the pitcher. You are pretty sure this amount is sufficient to just fill the glasses, but because the cup measure is not fancy enough to use in front of your guests, you dump the wine back into the carafe and hurry back into the dining room. To your surprise (and embarrassment), you are not able to top off all the glasses evenly, and the last person in line for a refill gets visibly less wine—by a fair amount—than everybody else. Somehow, it seems you do not have the 945 mL of wine you measured in the kitchen. After your guests leave, you decide to figure out what happened and recreate the awkward situation from earlier in the evening. You add 945 mL of water (your guests drank the last of your wine) to the cup measure and then use it to fill the five glasses still sitting at the table. Remarkably, this time there is sufficient water to fill the glasses adequately. You are puzzled and want to pursue this minor mystery. So, you repeat the steps you took, noting that there are different amounts of water left to fill the final glass each of the six times you carry out the test. As a next step you choose to dump the water back into the cup measure to determine the total amount present in the five glasses *after* they have been filled. You write down the volume left after each trial and come up with the following six numbers: 940, 938, 937, 935, 929, and 917 (all mL). How could this be? What happened? Why do you get six different numbers? The answer to these questions is linked to experimental error and scientific uncertainty, and we will return to it shortly.

Experimental errors

Outside of science, the word "error" generally refers to a mistake or an accident. If you drop the glass pitcher of wine (from our scenario, above) and it shatters on the floor, you might use the word to describe your action (with or without other, less refined, words added for emphasis). In a more scientifically relevant example, if you intend to determine the height of something but instead measure its width, you will have made a mistake. Importantly, that mistake can be corrected if you redo the work. However, the word means something quite different when used in the context of measurement. Put simply, **experimental error** is defined as the difference between the *measured* value of a property and the *actual* value of that property. For example, if the temperature of a liquid is 25.6 °C but you measure it as 25.1 °C, you are off by 0.5 °C, that is, your error is 0.5 °C. Unlike the mistakes described above (dropping the glass pitcher, etc.), not all experimental errors can be corrected simply by repetition. In fact, although they can be minimized, they cannot be eliminated—errors are an inherent part of every measurement and must be quantified and reported in any experiment.

Types of experimental errors.

Systematic errors (= determinate errors)

These affect **accuracy**, how close a reading is to the actual value (see the section below, Accuracy, Precision, and Sensitivity, for more information). Because of these types of errors, every one of the readings of the same phenomenon is wrong in the same direction: they are all either higher than or lower than the actual value.

Systematic errors tend to arise from faulty or uncalibrated instruments or human error (more about **calibration** is described below in the section Accuracy, Precision, and Sensitivity). For example, imagine you are using a cloth tape measure to determine the length of a snake you found under your porch. Unfortunately, since it has been stretched out from years of heavy use (yes, you live in a snake-filled swamp), it consistently yields readings that are too short (i.e., although the tape "says" it is measuring something at 1 m, the object is actually 0.9 m long). You might take twenty measurements of the snake, but each will be wrong—too long—because the instrument you have is faulty. Using a consistently bad technique to make a measurement will also generate systematic errors (e.g., a scientist might not know how to properly read an instrument). Importantly, no matter how many times you repeat a measurement, it will remain inaccurate. Strategies that *do* work to reduce systematic errors include calibration of instruments and training of scientists doing the measuring.

Random errors (= indeterminate errors)

These errors affect **precision**, how well multiple readings of the same phenomenon agree with each other (again, see Accuracy, Precision, and Sensitivity for more information). Unlike systematic errors, random errors lead to a group of measurements that is off in both directions: some will be above and some below the actual value. Random errors arise from variables that are difficult or impossible to control. Consider how a hypothetical experiment conducted in an outdoor field to assess the effects of fertilizer amount on plant growth will be affected by weather. Among other problems, each day, conditions such as temperature, sun exposure, and moisture will change. It is possible that differences in growth are linked to climatic factors and not fertilizer availability at all. Variation in responses among different test organisms, termed **biological variability**, is another common source of these types of error. For example, one laboratory rat will likely respond differently to the same dose of a drug undergoing testing than will a second rat. Scientists can also introduce random errors by the way they take and read measurements. Imagine you are trying to determine the volume of water in a glass container. You read the markings on the wall of the container (called **graduations**) and notice that the top of the solvent lies between the 17.5 and 17.6 milliliter (mL) marks (see Figure 2.2).

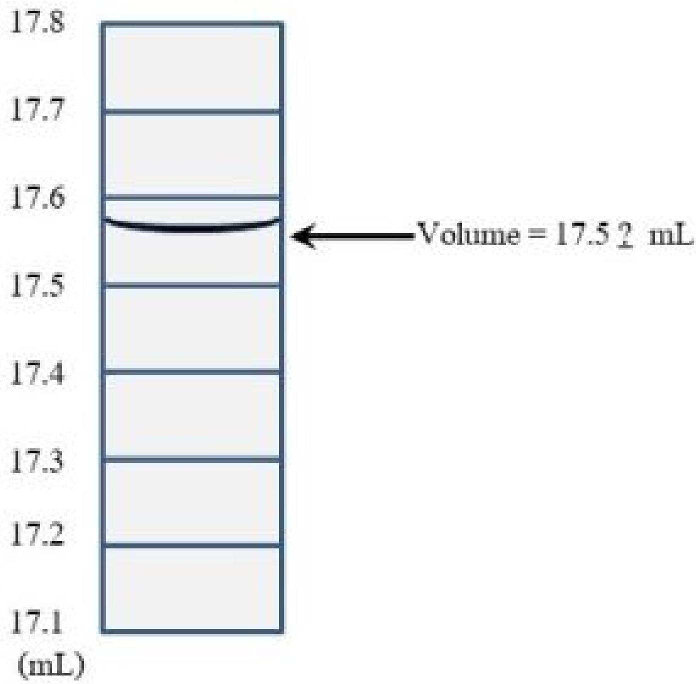


Figure 2.2. *Measurement uncertainty due to interpolation: opinions will vary on the value of that fourth digit.*

Generally, researchers will estimate the last digit using interpolation. In our case, you would look at the liquid and make a reasonable guess about what number should go in the 100^{ths} place. After staring at it for a while, you might decide that 17.56 is the best answer. A different person might say 17.58, a third 17.54, and so forth. No matter what you do, there is some subjectivity associated with this measurement, and, as a result, such a data set will contain points that are clustered both above and below the actual answer. Unlike systematic errors, repetition can reduce the size of the random error. If one takes 5, 10, or even 100 measurements of the same phenomenon, error will be reduced (which can be demonstrated using the tools of statistics, something that is beyond the scope of this text). In any case, a single measurement would be considered insufficient to draw any useful conclusions—more data points are always better than fewer. Taking steps to minimize variables can also reduce these types of errors. In the field experiment suggested above, we might control weather conditions by moving all the plants to a greenhouse. In this way, we can isolate the variable we would like to study, the effect of different amounts of fertilizer on plant growth, and keep everything else constant (i.e., the same for all plants). In drug testing, the use of large numbers of animals that are as identical to each other as possible and the development of standardized procedures will reduce random errors as well (see Chapter 15 for more about drug testing).

Determining and reporting experimental errors. A well-designed experiment includes multiple readings of the same phenomenon. In other words, we would not make any conclusions about plant growth (our experiment above) by looking at a single plant. We would instead treat several plants in the same way and try to determine some representative answer based on observations of the group. Remember that biological variability will cause individual organisms, even those that are very closely related, to respond differently to the same stressors. So, we will likely get multiple responses, perhaps as many as ten, in the group of ten plants provided with the same amount of nutrients. How do we resolve one question with ten different answers? Although there are several ways to handle such a situation, a commonly used strategy is to calculate the mean (i.e., average) and standard deviation of each group to express an answer. In somewhat oversimplified (yet reasonable) terms, **standard deviation** represents our level of uncertainty, or, put another way, the range within which the actual answer is likely to be found, and the **mean** is used to suggest the most likely answer in that range (i.e., it is an appropriate representation of the actual answer). Consider two data sets (given as mean ± standard deviation): 10.2 ± 2.6 and 7.5 ± 0.7. For the first, we can approximate the range as 7.6 – 12.8 and for the second 6.8 – 8.2. Importantly, a bigger standard deviation indicates a larger amount of error and lower confidence in the data, which may be more easily visualized graphically (Figure 2.3).

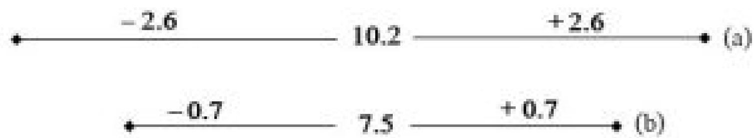


Figure 2.3. *A larger error term (a) suggests a wider possible range of answers than does a smaller error term (b). Not to scale.*

Progress despite errors. If every measurement is only an approximation, how do we ever learn anything about the natural world? How can scientists retain their sanity if they are constantly dealing with error? Might it be preferable to just close down their laboratories and open up ice cream shops? It turns out that, despite uncertainty, there are ways to interpret the ranges indicated by means and standard deviations to draw conclusions and keep science moving forward. The details of the mathematics involved in analyzing data are not the point of this textbook, but in short, it is possible to determine relationships among groups receiving different treatments (i.e., exposed to different conditions) in an experiment. Through the mathematical equivalent of comparing the ranges of the groups and looking for likely overlaps, a scientist can determine the effects of a changing variable on the subject under study. A fundamental yet critical concern is whether groups of test subjects receiving different treatments are different from each other, that is, whether the variable of interest has a real effect that cannot be explained away because of uncertainty. Recall the experiment in which we studied the effect of different amounts of fertilizer on the growth of plants. Among other things, we want to determine if the groups are different from each other—if and how nutrient availability affects growth. To simplify our discussion, we could compare the size of the plants at only the highest and lowest nutrient conditions. The measured masses from the ten different plants in each group are presented in Table 2.2.

Table 2.2. *Effect of fertilizer on plant growth. "±" denotes "plus or minus", how far the error term goes above and below the mean.*

Mass of Plant at Harvest (grams)	
Low Fertilizer	High Fertilizer
35.2	44.4
34.3	47.1
29.1	48.3
27.5	50.2
37.2	39.7
33.0	42.1
32.6	43.5
30.9	46.7
26.4	41.6
38.8	49.2
Mean	
32.5	45.3
Error	
± 4.1	± 3.5

Now, it might seem odd to even wonder such a thing, but we want to know if the two groups are different from each other. Although 45.3 is surely a higher number than 32.5, differences in the amount of growth of individual plants receiving the same amount of fertilizer complicate a seemingly simple comparison and leave open the possibility that the two numbers need to be thought of as the same. There are many responses that we must consider when we try to make our conclusion: plants receiving less fertilizer produced between 28.4 and 36.6 g of tissue whereas plants receiving more produced between 41.8 and 48.8 g. Again, keeping things fairly simple, we can scrutinize the ranges to determine how likely it is that 45.3 really is higher than 32.5. As before, we could use formal statistics to confirm our conclusion, but a visual inspection reveals what the math would tell us, namely, that there is no overlap between the two groups (see Figure 2.4a). Even if the low-fertilizer growth is as high is it could possibly be (36.6 g), it is still less than the smallest amount of growth under high-fertilizer conditions (41.8 g). We can be reasonably confident, therefore, that the two groups are different and that plants grow better with the higher level of fertilizer. However, what if our error terms were substantially larger, say 10.3 for the low group and 9.6 for the high group? Our ranges would become 22.2 to 42.8 and 35.7 to 54.9 for the two groups, respectively. As unsatisfying and perhaps illogical as it might seem, we would be forced to conclude that 45.3 and 32.5 are not necessarily different from each other—it is possible, given uncertainty, that growth under low-fertilizer conditions is the same as or even high than that under high-fertilizer conditions. In this case we could either report those findings (i.e., amount of fertilizer does not affect growth) or, in all likelihood, redesign and repeat the experiment to try to reduce uncertainty (see above for ways to shrink the error). Figure 2.4b shows the overlapping ranges for our second case.

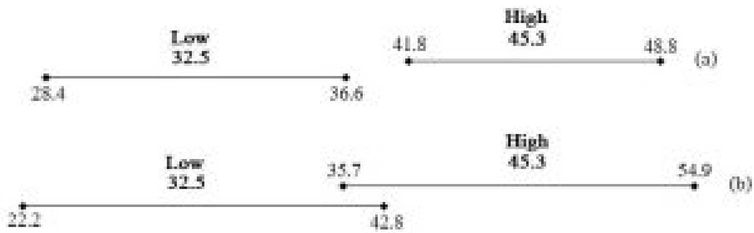


Figure 2.4. *Plant growth (grams of tissue) under low- and high-fertilizer conditions. Ranges do not overlap with a small error term (a) but the groups are assumed to be the same with a large error term (b). Not to scale.*

The inclusion of an error term is necessary for the proper understanding of data both in- and outside of the sciences. You likely have seen how public opinion polls designed to predict the outcome of upcoming elections tend to provide an indication of uncertainty when they are reported. In a hypothetical race for U.S. President, one candidate might be polling at 47% and the other at 44% with a margin of error of ± 5%. Numerous factors could account for the uncertainty, including the facts that only a small subset of people is used to represent the entire voting population and poll questions could be misinterpreted by respondents. In any case, because the ranges overlap, the most reasonable conclusion is the race is currently a tie, or "too close to call." Of course the answer revealed on election day can act as a kind of check on the quality of the polling—how well voters' behavior was predicted by today's survey can inform the reliability of future polls done by a particular organization. Put another way, election results can be used to **calibrate** a poll (more about calibration is described below). We can only evaluate results of experiments, and come to whatever conclusions they suggest, if uncertainty is rigorously calculated and described.

Accuracy, Precision, and Sensitivity

Although they are sometimes used interchangeably in casual settings, these terms have three distinct and important meanings to scientists.

Accuracy. This term refers to correctness, how close a measurement is to the actual value of the property we are assessing. So, if we determine the mass of an object to be 10.3 g when the true value is 10.6 g, we are inaccurate by 0.3 g. It is a relatively simple matter to determine just how wrong we are if we happen to know the true mass of the object in question. But what if we do not know the actual value? How then can we judge accuracy? Remember the example of the polling data from above. We can try to predict the outcome of an election by sampling the opinions of voters, but our answer will have some level of uncertainty associated with it and will likely be wrong by some amount. Just how good our polling methods are can be tested by comparing our prediction to what happened in the election. We have the luxury to learn the right answer to the question. When the actual value is not available to us (which generally is the case in science) we instead rely on **calibration**, that is, we use an object with a known value for the property in question (mass, length, volume, etc.) to test our instrument. When determining the mass of an object, for example, we ought to consider whether the balance we are using is working correctly. To be flip about it, does the balance know what it is doing? Does it have any idea what a gram even feels like? To answer these questions we can acquire a calibration weight, an object with a known mass, and measure it on our balance. If the object we know to be 1.0000 g is assigned a mass of 0.9986 g on our balance, we need to reset, or calibrate, the instrument with the 1.0000-g mass to improve the accuracy of future measurements. The same must be done regularly with meter sticks (length), graduated cylinders (volume), and other laboratory instruments. See Box 2.2 for more on calibration.

Box 2.2. How do we know what 1 g *really* is?

Before you attempt to use a balance, you should check the instrument with a carefully manufactured object of a known mass. If the mass reported by the instrument to be used is either more or less than that of the test weight, then you know that the balance is wrong, and you must calibrate before proceeding. We might think of this as telling the balance what the right answer is—or the standard against which it should compare—before we try to measure an unknown. But, what about that expensive calibration weight we purchased? How do we know that *its* mass is correct? The simple answer to this question may surprise you: there is an actual object stored at the International Bureau of Weights and Measures (near Paris, France) which, by international agreement, is considered to be the prototype of the kilogram (1000 g). All balances are ultimately compared to and calibrated with that universally accepted standard mass. Since the object is kept under glass inside a secure vault and only taken out occasionally, it is obviously not used directly. Copies of it, and copies of those copies (and so forth) are used to calibrate the world's balances. The standard unit for length was similarly based on a single prototype, the meter, until the 1980s. Since then, a different way to define the meter has been used, although the idea of a universally accepted standard persists for it and other units.

Precision. This term refers to how well separate measurements of the same object (termed replicate measurements, or **replicates**) agree with each other. In other words, if you measure the volume of a liquid multiple times, how similar to each other are your replicates? Imagine one experiment in which we test the temperature of a liquid four times to yield the answers 35, 41, 29, and 45 °C, and in a second experiment we end up with 34, 36, 32, and 30 °C. As we saw previously when considering the effect of fertilizer on plant growth, there is a range of answers here. We could quantify our precision and our level of uncertainty using the tools of statistics, but in simple terms an examination of the two data sets reveals that there is a larger range (i.e., *lower* precision) in the first experiment relative to that in the second experiment. Like accuracy, precision is a critical property—but the two terms refer to very different considerations. In fact, as explored in Box 2.3, a set of data could be very precise yet still inaccurate.

Box 2.3. Accuracy vs. Precision, OR: close only counts in horseshoes and hand grenades….what about science???

Just because you can get nearly the same number each of the dozen times you measure something, you should not assume your answer is accurate. You might generate a data set of readings that are tightly clustered together and record a mean with a small error term. Unfortunately, every one of the readings will be wrong, perhaps by a lot, if your balance has not been properly calibrated. We can use the somewhat old-fashioned lawn game known as Horseshoes to think a bit more about the distinction between accuracy and precision. Keeping things simple, in Horseshoes one throws an object—that's right, a horseshoe—12 meters toward a vertical stake in the ground. If the horseshoe completely encircles the stake, the shot is called a ringer and 3 points are awarded. A throw that is not a ringer but lands within 15 centimeters (6 inches) of the stake is close enough to score 1 point. So, how are all these rules related to accuracy and precision? The highest-scoring and most-difficult shot, a ringer, can be equated with the right answer, or high accuracy. A less accurate shot is worth only 1 point. Now, it is possible that all the horseshoes you throw in an inning land close to or even on top of each other but still far from your intended target. Clearly you have made consistent throws if the shoes land together, but you still get 0 points if they sit 2 meters from the stake. In this case, you are both beautifully precise and horribly inaccurate (see Figure 2.5)! Although repeatability is often thought to be confirmation of correctness, you should be cautious and remember that precision and accuracy are not the same things.

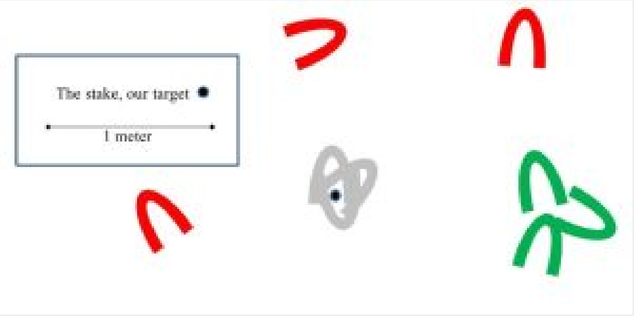


Figure 2.5. *Precision vs. accuracy. Three grey shoes are clustered on the stake with good accuracy and precision. The green ones have good precision but poor accuracy, and the red ones are neither precise nor accurate (map view).* Kelsey, CC BY-NC-SA.

Sensitivity. This final term is generally applied to the instrument used to make a measurement more than the data themselves, and, although it is distinct from both accuracy and precision, it can influence those two other properties. Simply put, sensitivity refers to the smallest measurement an instrument can make, or sometimes we say it is the minimum change an instrument can detect. The sensitivity of an instrument is a crucial consideration when one is making a measurement. For example, if you wanted to determine the size of a bacterium, a meter stick would not be the right tool to use. If you consider that bacterial cells are about 1 millionth the length of a meter stick (on the order of 0.000001 m, or 1 micron, long) the problem should become immediately clear. If you were somehow able to see a single bacterium lined up next to a meter stick (a significant challenge since **bacteria** are microscopic—consult Figure 2.6), you would notice that the cell is about a thousand times smaller than the smallest unit marked on the stick (millimeters, mm, or 0.001 m).

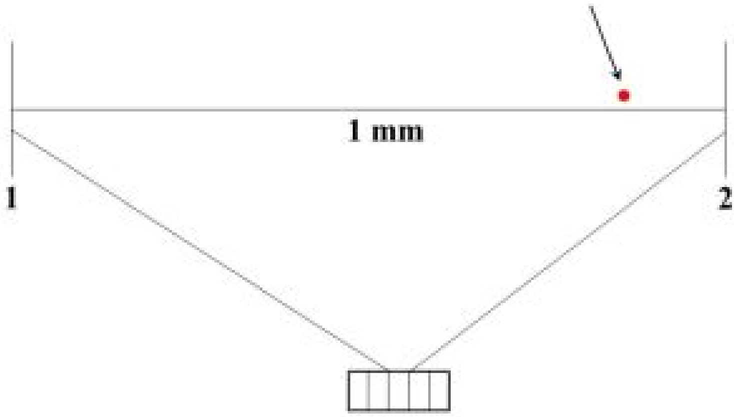


Figure 2.6. *A bacterial cell relative to a millimeter. The cell is the red dot beneath the arrow. A much more sensitive tool would be needed to make an accurate measurement of length in this case. Kelsey,* CC BY-NC-SA.

Even if it is perfectly calibrated, the best answer you can get with a meter stick is something unhelpful like "bacteria are a lot smaller than 1 mm". The tool is clearly not *sensitive* enough to determine the length of a very small object. Similarly, it would be foolish to try to measure something with a mass of 0.0045 g using a balance that cannot detect changes smaller than 0.00 g. The balance would continue to read 0.00 before and after the object was placed on it. Keep in mind that, although accuracy and precision are not the same as sensitivity, they can still be adversely affected if an instrument does not measure small enough quantities. As you might imagine, it would be very difficult to determine the true value of the length of the bacterium we mentioned above with a meter stick. Getting replicate answers to agree with each other would also be nearly impossible under these circumstances.

Uncertainty happens, revisited

Now that we have considered sources of error and uncertainty, we are ready to return to the problem of insufficient wine at our dinner party. You will recall that, when we conducted our experiment to track down an explanation, we ended up with different amounts of water each of the six times we measured. How could this be? Well, many potential sources of error exist, including: the measuring cup might not have been properly calibrated, we could have read the volume from a different angle each time (leading to the different answers), and we did not account for water spillage during transfers back and forth (note that the volume left in the glass declined after each replicate, 940, 938, 937, 935, 929, 917 mL). Errors associated with the dinner party itself, such as our rush to get back to the dining room and our assumption about how much wine was required to fill the glasses, could have compounded the problem. Of course, the addition of alcohol into our bloodstream should not be discounted either. Other systematic and random errors surely could have affected the outcome in our situation. In short, even a problem as simple as pouring wine into multiple glasses is fraught with uncertainty. Imagine how many potential errors can affect more complicated environmental experiments! As we progress through this textbook, we will identify and manage sources of uncertainty yet still make reasonable conclusions about the natural world. This is the way of science.

What is the point of all this? What good is science anyway?

As we know, experiments are designed to objectively collect data. But to what end is science conducted? Does science serve any useful purposes? Although scientists tend to be drawn to their fields because they find them fascinating and simply enjoy learning how the world works, science is used extensively to address most of the practical questions important to society as well. Medicine, agriculture, nutrition, food safety, development of new materials, construction, cosmetics, pharmaceuticals, power generation, drinking water availability, air pollution, and waste management are but a sampling of the fields and products that depend on scientific inquiry. In fact, it is difficult to name many areas of modern life that are *not* directly informed and improved by science.

Before we leave our introduction to science, we must address one final and critical question, namely, who decides what should be studied? The answer to that question is tightly linked to another important, albeit somewhat crass, one: who pays for scientific research? Conducting science can be very expensive, and there is a finite amount of money available to do so. Much, but not all, of the funding for research is provided by government programs that tend to reflect the values and needs of society. So that means scientists must apply for money to support their work. Since the process associated with securing funding is generally extremely competitive, the questions to be asked, as well as the design of the experiments proposed to answer them, must be very well thought out and deemed relevant to society before funding is granted. Like the peer-review process for potential scientific papers described above, grant proposals undergo intense scrutiny by experts in the appropriate field before money is released. It is fair to say that there are many more losers than winners when it comes to competitive grant writing.

2.2. SYSTEMS ANALYSIS

Like the scientific method, systems analysis will serve us often and well during our exploration of Earth's environments. In addition, you will soon see that the tools we learn here are helpful in the studies of many subjects, events, and processes far beyond environmental science.

2.2.1. A whole greater than the sum: the system definition

In very simple and general terms, an entity can be classified as a **system** if it meets certain criteria.

It is made up of component parts

These subunits are connected to each other somehow, affect each other, and broadly speaking, work together. The old expression "the whole is greater than the sum of the parts" is apt here because, through their connections and interdependence, the combined pieces can accomplish outcomes that are different than would be possible if they all worked on their own.

It can be distinguished from its surroundings

A system has characteristics and a discrete identity that set it apart in obvious ways from other entities. Since some systems are embedded within or firmly attached to other systems, this separation may be only conceptual, not physical. For example, the human digestive system is an entity in its own right, but clearly it cannot easily be disconnected and taken away from the other systems in the body without some unfortunate results for the person involved. A single human is itself a larger system that is physically and obviously distinct from other systems.

It receives inputs and produces outputs

Many kinds of **factors** (materials, energy, etc.) can enter a system and undergo processing by it. Those factors can be released by the system unchanged relative to the form they took when they entered, or they can be altered inside the system and exit as a part of new products. Figure 2.7 is a generalized way to visualize any system, including the relationship between inputs and outputs.

Figure 2.7. *The basic design of any system. This symbology can (and will!) be used to analyze many environmental processes we encounter in this book. Be sure to get comfortable with it. Kelsey, CC BY-NC-SA.*

This model will appear throughout this book as it is an appropriate, not to mention very useful, way to help understand many environmental processes. Keep in mind that what a system produces is a function of what it receives: if inputs change, outputs will change as well. In other words, **systems are responsive.** We will consider some of the inputs and outputs that affect and are affected by systems shortly.

It interacts with other systems

Although systems can be defined as separate from their surroundings, rarely is a system completely isolated from all others. In fact, systems generally affect each other, and we categorize them based on the ways they interact. A system is **closed** with respect to a particular factor if it does not exchange that factor with other systems and **open** with respect to any factors it does exchange. Note that a system can be closed with respect to one factor and open with respect to others, so it is not correct to simply refer to a system as either "open" or "closed"—we should specify which factors are shared and which are not. Earth, for example, is closed with respect to materials and open with respect to energy (Chapter 1).

Size is irrelevant

They are made of component parts, but systems can range in size from microscopic to as big as the universe.

Many familiar phenomena and objects lend themselves to systems analysis

If you recall our basic definition from the beginning of this section, you will start to notice systems all around you. Before we proceed any further with our general discussion, it might be helpful to consider a few specific entities that are good examples of systems.

A person. An individual human being can be viewed through the lens of system analysis because it meets the basic criteria. First, it is made up of components that work together. The circulatory, respiratory, and digestive systems are but three of the many subunits within a body. We can also identify organs such as the heart and lungs, tissues such as fat and bone, as well as each of the trillions of tiny cells that make up everything in the body. Second, it receives inputs such as food, heat, water, education, love (or otherwise), and so forth. These inputs are processed and released as outputs such as waste, social skills, children, information, and anything else produced or given by a person during its lifetime. An individual human being is by no means completely independent and self sufficient, it is open to materials, energy, ideas, and feelings. Finally, it can be easily distinguished and physically isolated from its surroundings.

A computer. This familiar object readily fits our three criteria: it is made of components working together, it is easily isolated from its surroundings, and various factors enter and are produced by a computer. This last characteristic provides an excellent example of the relationship between inputs and outputs so will get a little additional attention here. Computers receive many inputs, including programming, numbers from measurements, and the design of the computer hardware itself. What is produced by a computer is very tightly and clearly linked to the inputs it receives. Word processing documents, calculations, diagrams, musical compositions, actions of valves, doors, and vehicles—the list of possible outputs from computers could go on and on—are all highly dependent on specific instructions, data, and ideas they receive. Clearly, a person who types incorrect numbers into a spreadsheet formula will receive faulty answers to whatever question is being asked. A change in the input values will lead to a change in the output as is true for any system. Arguably, a computer is open with respect to materials, as parts can be added and subtracted, along with energy and information.

A pond. Depending upon how broadly we care to define it, the component parts of this small system include water (and all the subunits that make *it* up), sediments (like water, composed of many smaller components), plants, animals, and microscopic organisms. Generally speaking, a pond is easy to distinguish from its surroundings, as it is an isolated and stagnant body of water. We could decide to include some amount of the land surrounding the water, but just how we delineate it would be influenced by what exactly we want to study. What about inputs and outputs? As we will see in more detail later in Chapter 5, individuals and groups of organisms need certain nutrients to survive: these materials and energy must be provided by the system in which they live. The pond receives inputs such as light, oxygen, and various other factors that organisms depend on. Some of the products released by the pond organisms are released from the system, some largely stay put. Certain gases and heat, for example, readily exit the system as outputs (the system is open with respect to them). On the other hand, materials such as physical waste and remains of dead organisms will likely remain inside the system (the system is closed with respect to them, at least in the relatively short term), perhaps settling to the bottom and serving as food sources for other organisms. An analysis of inputs and outputs would help us to predict the future health of this system, as changes in anything will likely affect all the organisms living there. If certain factors accumulate because they enter at rates higher than those responsible for their removal, for instance, the entire pond could undergo dramatic changes. We will revisit this example when we discuss nutrient cycling in Chapter 4.

2.2.2. Feedback: a system responds to the output it produces

Earlier we said that systems are responsive: what they produce is a function of what they receive. However, we did not pay much attention to the origin of those inputs. Although a second system is often the source of factors entering one system, it is not impossible for a system to provide factors to itself—something produced by a system can actually re-enter and affect that same system. This phenomenon in which a system responds to its own outputs is referred to as **feedback** (see Figure 2.8), and the connection between what goes out and what goes in is sometimes referred to as a **feedback loop.** Feedback can bring about one of two very different outcomes, both of which play critical roles in the behavior of many systems (environmental and otherwise).

Figure 2.8. *A model system and feedback. Kelsey,* CC BY-NC-SA.

Negative feedback

With respect to a particular factor or factors, this first type of response will keep a system from changing much relative to its **initial conditions** (i.e., its starting point). How does negative feedback work? In short, output released now becomes new input later and *counteracts* actions of a system brought about by previous inputs. We can think of negative feedback as a mechanism by which current system response can cancel out previous system responses, or, put another way, regulate or reverse any movement away from its initial conditions. This somewhat murky concept becomes clearer if we consider a familiar system that is regulated through negative feedback. Climate control of interior spaces provides an excellent example of the principles of negative feedback. First, consider the goal of such a system: we want it to keep the temperature of the air in a room at some level we deem to be appropriate. Second, imagine how we might go about achieving our goal. Somehow, whenever there is a change away from the temperature we desire, often referred to as the **set point**, we would like the system to respond such that the change is reversed. If it gets too cold, we want the room to get warmer, and if it gets too warm, we want it to get cooler. Think about how you would design a classroom to maximize comfort (presumably, to minimize distraction from learning, not to induce sleep). Among other properties, we could decide that the ideal temperature is 20 °C (68 °F). So we include a climate control system that has the ability to both cool (i.e., air conditioning) and warm (i.e., heat) the space as needed. In this case let us assume our thermostat detects that the air it draws inside itself is 19 °C. That input, air temperature, triggers a response, and the heater is activated. Eventually, the temperature of the air in the room will rise to 20 and then 21 °C. At that moment, the changed input (i.e., air that is warmer than the set point) brings about a new response: the heater is turned off and the air conditioner is turned on. That system response will continue until the air is cooled below the set point and the heat will come back on once again. The result is a room that is maintained at a stable temperature indefinitely—there is little or no change. Figure 2.9 shows how our system model helps to visualize climate control.

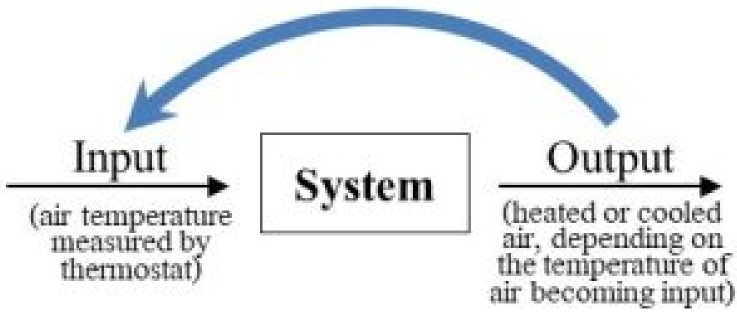


Figure 2.9. *A climate control system is an excellent example of negative feedback. Air temperature in a room is stabilized as a result. Kelsey,* CC BY-NC-SA.

Human body temperature is also stabilized through negative feedback. People will sweat if they get too hot (say, while running a race) and will shiver if they get too cold (say, while swimming in a chilly lake or after prolonged sweating). Whenever the body's internal thermostat detects a deviation from optimal conditions, 37 °C (98.6 °F), its response is to reverse the change. Negative feedback controls a large number of processes relevant to environmental science, as we will see in upcoming chapters.

Positive feedback

In this second case, the system response to its own outputs leads to changes with respect to a particular factor. Importantly, as feedback continues, the system moves farther and farther from initial conditions at an ever-increasing rate. Unlike negative feedback, current output becomes new input later that *accentuates* actions brought about by previous inputs. Again, examples will help us better understand this complex idea. We will start outside of the sciences with an adage: "it takes money to make money." It turns out this expression is a nice way to summarize how positive feedback can increase the amount of money in a bank account. Say you deposit $1000 into a bank that pays 5% annual interest on your balance. Importantly, the interest is compounded, meaning that the amount earned in interest is added to the principle, and that combined amount is then used to calculate how much future interest is paid. Keeping things simple, a 5% return on your initial investment would yield $50 after a year. In the second year, though, the bank would pay interest on $1050 (the principle plus the interest earned in the first year), which is $52.20, so after two years you would have $1102.20. After three years, the account value would be $1157.62, and so forth. The more money you have, the more money you will make, and the faster your savings will grow. A second example closer to the subject of this book is how positive feedback can cause forest fires to spread. Visualize what could happen to a slightly damp forest if a seemingly inconsequential ignition source (maybe a lit cigarette) is released into it: a rather small area will start to burn. After a short time, combustible material in the vicinity will get dried out due to the heat from the fire, and it too will ignite. Now a larger area than we saw initially is burning. That larger fire will heat and dry additional material, but at this stage the amount dried and prepped for a fire is larger because the size of the fire has increased. As the fire grows, it will dry material at an ever-increasing rate and therefore grow at a similarly accelerated rate (Figure 2.10). Put into casual language, the bigger it gets the bigger it *will* get.

Figure 2.10. *The effect of positive feedback on the growth of a hypothetical forest fire. The fire (orange) dries out a larger and larger area with time (light). As a result, the fire grows at an ever-increasing rate. In map view.* *Kelsey*, CC BY-NC-SA.

Population growth is another important phenomenon fueled by positive feedback. We will see this in more detail in Chapter 8, but for now consider how a small population can rapidly increase in size as it grows. In principle, the actions of just a few parents can ultimately lead to a very large number of descendants. For the sake of argument, we imagine a population consisting of three breeding pairs (six total individuals) that each produces four offspring. After one generation, the number of individuals will be 12. Keeping things simple, those offspring will eventually mature and pair up to produce their own offspring. If they behave like their ancestors, each of the six couples will give rise to four offspring, and the resulting population will consist of 24 new individuals (plus some survivors from previous generations). Under idealized conditions of infinite resource availability (not realistic, of course), the same pattern will repeat, and positive feedback will produce a constant rate of increase known as **exponential growth**. Put another way, the number of individuals in every generation can be multiplied by the same number to predict the size of the population in the future. A population that grows at 1% per year will add more and more individuals as the population gets larger and larger. Growth is a constant 1% of an ever-increasing number (e.g., 1% of 1000 the first year, 1% of 1010 the second year, and so on). See Figure 2.11 for a graphical representation.

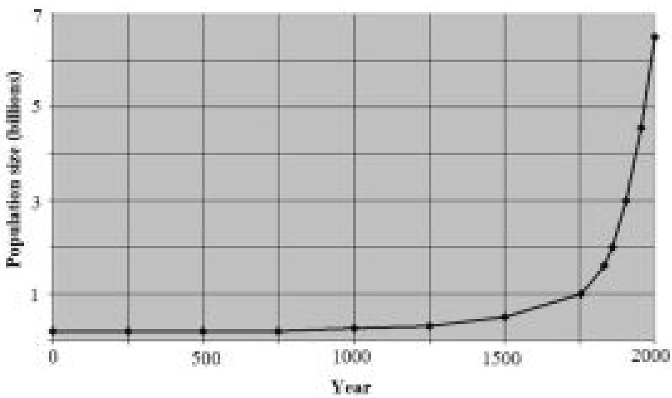


Figure 2.11. *Exponential growth of the human population due to positive feedback. Kelsey,* CC BY-NC-SA.

The same forces that drive compound interest (above) are active here: as the number of individuals capable of producing offspring grows, the number of individuals present increases at a faster and faster rate. As with negative feedback, we will see how positive feedback plays a role in many important environmental issues.

Three final points about feedback

Rate of change matters. Although it is tempting to view it as a phenomenon that always increases the size of a factor, note that positive feedback leads to an ever-increasing rate of *change* away from initial conditions. The observed change could yield either more *or* less output in the future. Consider the spread of a deadly disease, for instance. As more and more people are infected, the size of the population will *decrease* at an exponential rate. The same can be true of a melting glacier—as it gets smaller, the rate of shrinkage will increase because the low-temperature glacier keeps ice in its vicinity cool. Less ice leads to warmer temperatures near the glacier and even less ice in the future (Chapter 14). So, positive and negative feedback differ in that the former leads to instability and change whereas the latter brings stability.

Output must affect the system. It is critical to understand that the relationship between outputs and future inputs is a causative one. That is, a system moves back to (negative) or away from (positive) initial conditions as a *direct result* of output it just produced. To revisit two of our examples, a human body will cool itself *because* it was heated, and a forest fire gets bigger *because* of the drying action of an earlier fire. As we saw near the beginning of this chapter, correlation does not guarantee that two phenomena are mechanistically linked.

This is not about bad and good. The words "negative" and "positive" should not be read as value judgments. Negative feedback is so named because the output cancels out or negates the action of a system. The result could be either beneficial or harmful. Positive feedback leads to change, a result that could be a good or bad thing. In fact, stability is often good for living systems, and change can be deadly. We will see many examples of the effects of feedback on environmental systems in upcoming chapters, some of which we will interpret to be helpful and some we will likely decide are detrimental.

2.2.3. Relationship between rate in and rate out

The relative rates of input and output determine if the amount of a factor inside a system increases, decreases, or stays constant with time. This is an important consideration because, as we saw in Chapter 1, environmental scientists track the movement and storage of important materials on Earth. They often use the tools of systems analysis to predict the availability of resources vital to both humans and non-human organisms alike. We can identify three distinct possible relationships between the rate at which factors enter and those at which they exit from a system (Figure 2.12 and the text that follows).

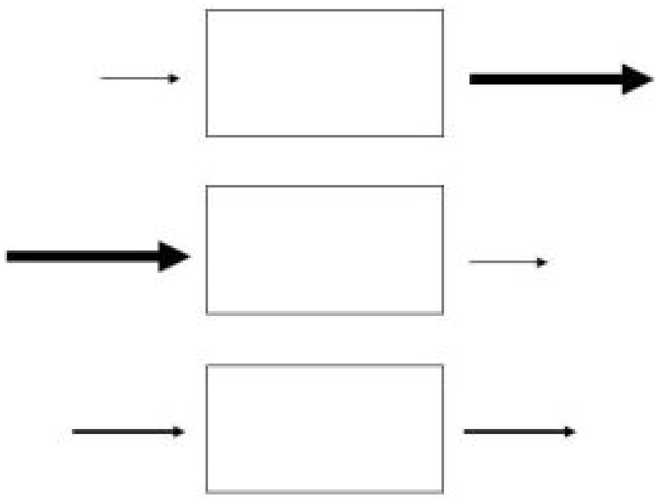


Figure 2.12. *Relative rates of input and output of a factor determine if depletion (top), accumulation (middle), or steady state occur (bottom). Size of arrow is proportional to rate. Kelsey,* CC BY-NC-SA.

Rate of output could be higher than rate of input (top model in Figure 2.12)

In this case, we would observe **depletion**, a measurable decrease in the amount of material *inside the system*. Again, we will begin with a financial example. Investment accounts typically grow while the people holding them are employed and earning a salary. Workers contribute to them and their values increases. However, upon retirement, these accounts act as sources of income for those who created them. In all likelihood, the rate at which money is withdrawn will exceed the rate at which it is added. We will see depletion on a number of occasions during our study of Earth's natural systems. Dwindling reservoirs of water, fuel, and fertile land are only three examples we will consider.

Rate of input could exceed rate of output (middle model in Figure 2.12)

Unlike the outcome of our first case, the factor under investigation **accumulates**, meaning there is an appreciable increase in the amount of the material *inside the system* with time. A familiar example from outside of the sciences is a savings or investment account. In short, we wish to see an increase in the amount of money inside the system—your portfolio, in this case—with time. This goal can be met by simply adding money faster than you withdraw it. Of course, the compound interest and positive feedback we considered previously will only increase the rate of input over that of output. Environmental scientists would be interested in this model as well, particularly if accumulation leads to damage of a system. A small pond could receive pollution far faster than it is able to process and discharge it, for instance, so the toxic material will accumulate.

Rates of input and output could be equal (bottom model in Figure 2.12)

In this case, there is flow through a system, but the net amount of a material *inside the system* does not change with time. Whatever exits is exactly replaced by new inputs. The term **steady state** is often used to describe this case. Is there a relevant example from the financial world to illustrate this case? When might we be interested in maintaining the amount of money in an account with no net increases or decreases? Checking accounts generally are designed to discourage both accumulation and depletion in the long run. Since little or no interest is paid on these types of accounts, there is no incentive to keep more money in there than is necessary to cover expenses (say, all the withdrawals made in a month). Depletion is also undesirable, though, as the inconvenient and costly phenomenon known as overdraft is the result. If your balance drops below the amount needed, your bank probably will assess a fee to you as a penalty. The steady-state model is important to environmental scientists because it is linked to the concept of sustainability we encountered in Chapter 1. Depending upon how they are defined and how long they are observed, some systems can be appropriately classified as steady state. For example, water in one small region might seem to be accumulating while it is depleted in another, but there is no change in the total amount of water on Earth. Practices such as tree farming in which harvesting is exactly offset by replanting could be viewed as steady state for a relatively short period of time. As we will see later, though, systems in which we cultivate plants of any type are subject to a reduction in how much growth can be supported in the long run (Chapter 9). Changes can also occur in the short run because materials are removed from systems at rates that are either higher or lower than those responsible for replenishing them. A related concern is the concept of **residence time**, that is, how long a material will persist in a system (see Box 2.4). We will return to input-output analysis and how it informs our understanding of resource availability throughout this book.

Box 2.4. Residence time: how long will stuff stay in a system?

Environmental scientists often wonder just how long a factor of interest will be present inside a system at steady state. If the material is required by organisms, we want to know if it will be present for sufficient time to support growth and survival. Sometimes the material is toxic, however, and the longer it stays inside a system the more opportunities it will have to cause harm. If we know the size of the pool, that is how much material is in a system, and the rate at which the material in question passes through that system (at steady state, either the rate of input or output provides the answer), we can use the following formula to determine residence time:

Amount in reservoir ÷ rate of outflow (or inflow) = residence time

Let's use a simple and familiar example to get a sense of how this works: the length of time it takes a typical undergraduate to earn a Bachelor's degree, or the residence time of a college student on a campus. Imagine a hypothetical college with a total of 2000 undergraduates and a graduating class of 500 each year. Residence time is therefore

2000 students ÷ 500 students / year = 4 years.

An environmental scientist might ask a seemingly more urgent question such as, what is the residence time of a harmful pesticide in a pond? If the size of the pond is 100,000 gallons (378,541 liters) and the rate of input is 1000 gal (3785 L) per month, then residence time is

100,000 gal ÷ 1000 gal / month = 100 months.

How is this information useful? If we know residence time of a toxic substance we will have an idea of how long it will take to clean up the affected system.

Pollutants with shorter residence times are likely of lesser concern than those that persist for long periods.

THE CHAPTER ESSENCE IN BRIEF [1]

The scientific method provides a framework to collect evidence and increase our understanding of the natural world. It is objective, collaborative, and holds itself to high standards. Additionally, the simple set of tools of systems analysis allows us to assess the movement and storage of materials as well as the relative stability of many different phenomena.

---

1. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 2 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[2]

Why is the scientific method useful? What's the point of such a complex set of steps?

How can you distinguish between an objective observation and a subjective one?

How would you react if someone says "your opinion is wrong!"?

Is scientific uncertainty an indication that scientists do not know what they are talking about?

Is it possible to obtain highly precise measurements if you use a scale that is poorly calibrated?

How could pollution of a pond be a good example of an imbalance between inputs and outputs as well as positive feedback?

2. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

PART II

3. Fundamentals of Geology and Microbiology

JASON KELSEY

In Chapter 1 we learned that environmental science is a relatively new, interdisciplinary, field concerned with the interactions among the living and non-living systems on Earth. The task is broad, and it requires insights and tools from many of the older, disciplinary sciences. We will review basic principles in chemistry, biology, and physics, for example, as they become necessary. Given the amount of critical context they will provide for us throughout this textbook, and the relative unfamiliarity people tend to have with them, geology and microbiology will receive special consideration here.

Key concepts

After reading Chapter 3, you should understand the following:

GEOLOGY

- The internal structure of the Earth and what it tells us about our planet's history
- The theory of plate tectonics and how moving plates shape the living and non-living components of Earth

- How rocks are formed, changed, and destroyed by the many processes of the rock cycle
- How geologists study Earth's history
- The nature of geologic time
- How knowledge of geology is important to environmental scientists

MICROBIOLOGY

- The basic structures of different cell types
- The major characteristics and roles of bacteria, fungi, algae, protozoa, and viruses
- How humans can harness microbiology for our own use
- How knowledge of microbiology is important to environmental scientists

Chapter contents

3.1. GEOLOGY

Geologists primarily study Earth's physical structures and history. If you recall the scope of environmental science from your reading of Chapter 1, the utility of geology to us should be, therefore, clear. First, since organisms are inextricably linked to their physical environments, geology helps us understand the stressors shaping the biosphere (we will learn more about this and other spheres in Chapter 4; for now, **biosphere** refers to all the living things on Earth). Some of these are sudden and short-lived events, for example earthquakes and volcanic eruptions (Chapter 7), whereas others require longer periods of time, such as climate change (Chapter 14), as well as the creation and destruction of landforms and bodies of water (this chapter, below). Second, geology provides evidence of Earth's tremendous age and explains how slow and steady processes can dramatically alter physical features that might appear stagnant when viewed on a human time scale. Finally, knowledge of geologic forces and history can help uncover naturally occurring resources such as water, fuels, and minerals. As we will see in part two of this textbook, locating and managing these and other materials is of great concern to environmental scientists.

Here we will only briefly consider the principles of geology most relevant to our study of Earth's natural systems. You should keep in mind, however, it is a very wide-ranging, and this short primer cannot possibly do justice to all its subdisciplines and complexities.

3.1.1. A dynamic Earth

It can be tempting to view our planet as static and, for all intents and purposes, permanent. In fact, the prevailing thought throughout much of human history was that Earth's surface has been the same since its beginning. Only relatively recently did people start to seriously consider the possibility that landforms and bodies of water are temporary, even if long lived. Using several important principles, including the two included below, current-day geologists understand the Earth to be ever changing.

The theory of plate tectonics and the tectonic cycle

This is one of the most foundational and influential ideas developed by geologists during the past century. It describes a plausible mechanism by which the surface of the planet is shaped and reveals that, contrary to long-held opinions, large-scale changes to the size and shape of mountains, continents, and oceans have occurred during the past and will continue into the future. The main points of the theory are as follows.

Earth is neither solid nor homogeneous.

Chemistry changes with depth

As shown in Figure 3.1, the Earth is divided into three layers based largely on chemical composition (more about Chemistry is presented in Chapter 4).

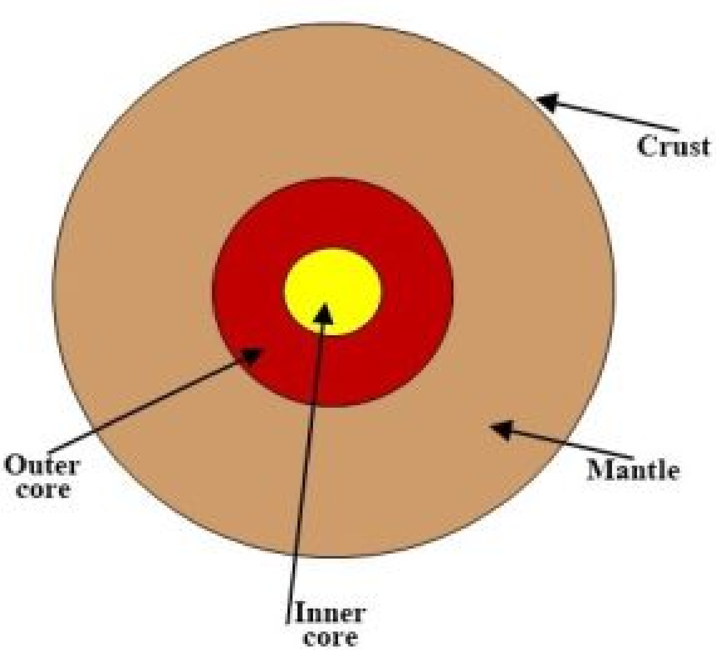


Figure 3.1. *Diagram of the Earth and its inner zones. Not to scale, in cross section. Kelsey, CC BY-NC-SA.*

Crust. This zone ranges from 0 – 50 km deep (its dimensions vary and can be as thin as 10 km in some places). It is the exterior shell covering the entire outside of the globe. It is mostly comprised of low-density materials such as oxygen, silicon, and aluminum. Substances like iron and magnesium, which are relatively dense, are present in small amounts. Temperatures in this layer range from what can be found at the surface (down to about 0 °C) to near 400 °C in its deepest sections.

Mantle. Ranging from 40 – 2900 km deep, this much thicker section lies between the crust and the core and makes up the bulk of Earth's volume. It too contains oxygen and silicon, but magnesium and iron are present in higher amounts here than in the crust. Temperatures in this layer range from about 400 °C in the upper portions to about 4000 °C near the core (next item).

Core. The center of the Earth (2900 – 5150 km depth) is thought to contain two portions, an inner solid core and an outer molten core. Dense iron and nickel dominate, and only small amounts of oxygen and silicon are present here. Temperatures in this layer are likely on the order of about 5000 °C.

Physical properties change with depth

Materials beneath the surface also can be categorized according to their physical behavior, resulting in three fundamental layers.

Lithosphere. As we will see in Chapter 4, this is one of the four spheres of concern to environmental scientists. In simple terms, the lithosphere is made up of the crust and upper portions of the mantle and extends to depths of about 100 km. Materials in this layer are hard and brittle, that is, they tend to break when stress is applied to them.

Asthenosphere. This zone is made up of deeper portions of the mantle, down to about 200 km. Here, materials are plastic instead of brittle. So, when stress is applied to them, they do not break; instead, they move a bit like a very viscous liquid, changing shape without fracturing.

Mesosphere. This encompasses the lower portion of the mantle and extends to the core.

The lithosphere floats on the asthenosphere

Since the lithosphere is comprised of the relatively low-density crust and upper mantle, it sits atop the higher-density materials in the asthenosphere. Figure 3.2 shows the position of these two layers.

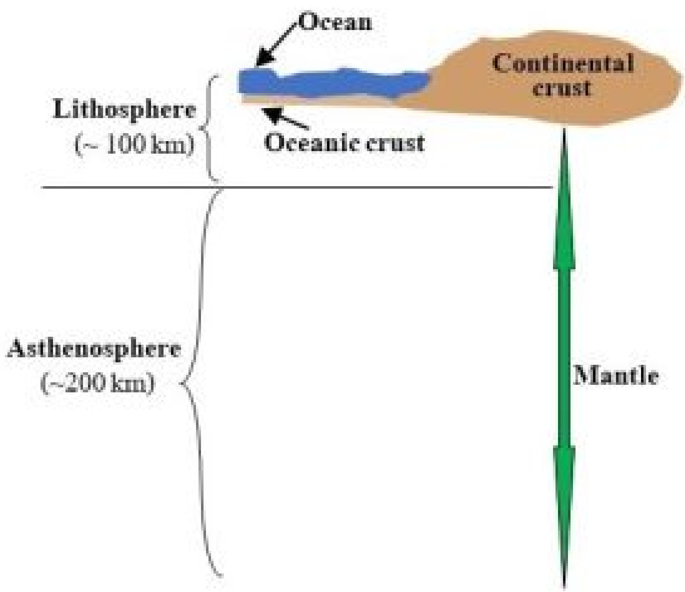


Figure 3.2. *An idealized diagram of the lithosphere and asthenosphere. Not to scale, in cross section. Kelsey,* CC BY-NC-SA.

The lithosphere is made up of large pieces

The brittle lithosphere is not one continuous layer but instead is broken into irregular units referred to as **tectonic plates**. These plates move slowly along the surface of the plastic asthenosphere. Figure 3.3 shows a map view of the 15 major plates making up the bulk of the lithosphere.

Figure 3.3. *The different colors represent major tectonic plates; arrows indicate direction of movement of each plate. Map view. USGS, Public Domain.*

Moving plates interact. On average, plates move about 2 cm each year. Since they are sliding along the surface of the same globe, all plates share boundaries with other plates. Several important processes are active at these boundaries, including production and destruction of new crust, formation of mountains and volcanoes, and induction of earthquakes (more in Chapter 7). Three basic types of interactions are possible (Figure 3.4).

Divergent (Fig. 3.4a)

At this boundary type, plates are moving away from each other. As a result, a long thin opening is formed in the lithosphere. New crust is produced as molten rock rises from the mantle to the surface where temperatures are low enough for solid rocks to form (more about rock formation is described later in the chapter). The crack running approximately 16,000 km north-south in the middle of the Atlantic Ocean is a good example of a divergent boundary. This structure, the **Mid Atlantic Ridge**, is confined to the ocean floor except for a small area that makes up Iceland.

Convergent (Fig. 3.4b)

Here, plates move toward each other, colliding at their boundary. Two broad outcomes are possible. In the first, one of the plates dives below the other in a process known as **subduction**. If this occurs, old crust sinks into the mantle, melts, and is absorbed into the asthenosphere. Some of that melted material rises toward the surface and can lead to volcanic activity (Chapter 7). The Andes Mountains in Western South America provide a good example of the result of this type of plate interaction. Second, there is no subduction. Instead, the two plates produce tall mountains because the only direction the colliding rocks can move is upward. The Himalaya Mountains in Asia illustrate this phenomenon. Keep in mind that the lateral pressure associated with plate movements can ultimately cause **uplift**, a process whereby rocks on the ocean floor or in other low places are pushed to higher elevations. We will see that uplift plays a critical role in the recycling of many of Earth's materials (Chapter 4).

Transform (*Fig. 3.4c*)

These boundaries are different from the others because plates slide along next to each other, grinding and scraping as they move. It is fair to say that crust is *displaced* in these zones rather than created or destroyed, and earthquakes (Chapter 7) are a typical result. Arguably, the most famous of these is present between the North American and Pacific plates and is responsible for the San Andreas Fault in Southern California (U.S.A.).

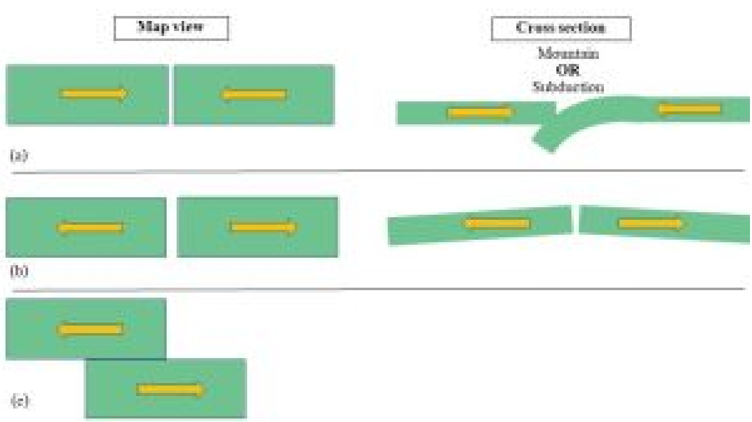


Figure 3.4. *Three plate boundary types in map view and cross section: a, convergent; b, divergent; c, transform (cross-sectional view is not relevant for c, transform). Kelsey,* CC BY-NC-SA.

Plate movements matter. It would be difficult to overstate the importance of plate movements. Their constant motion subjects the Earth's surface to unrelenting change. As we saw above, plate interactions alter landforms and trigger earthquakes and volcanic eruptions. Moreover, continents and ocean basins are slowly created, shaped, and destroyed due to tectonic forces. One well-studied example of the power of moving plates is seen in **Pangaea,** an extinct supercontinent that was once a single landmass made up of all Earth's modern continents. About 200 million years ago, plate movements led to the break-up of Pangaea, creating isolated landmasses, including North America and Africa, and bodies of water, such as The Atlantic Ocean (See Figure 3.5).

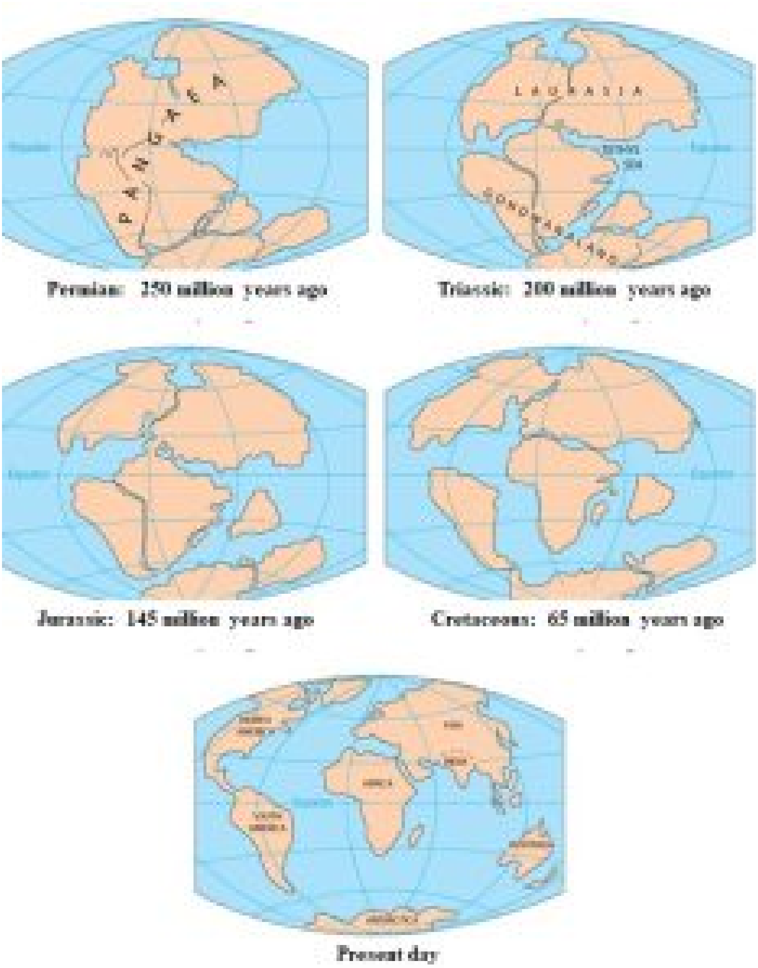


Figure 3.5. *Reconstruction of 250 million years of continental drift. The names given in each panel refer to units of time known as geologic periods. See Table 3.1., below, for more. USGS, Public Domain.*

What drives plate movements? Much evidence supports the premise that lithospheric plates are in motion and carry Earth's crust with them (see below for more on the development of the theory of plate tectonics). But, just *how* giant lithospheric plates move across the surface has been the subject of study and debate for some time. Although no definitive and universally accepted answer has yet been found, a few possible mechanisms have been proposed. The dominant idea for several decades held that plates were propelled by convection cells in the mantle: circular motion is initiated as hot material rises at divergent boundaries and cold material sinks at convergent ones. As happens so often in science, though, recent data have caused us to revise our understanding. Additional proposals have been put forward, including the possibility that gravity plays a large role. More research is clearly needed to work out this problem.

The evidence for plate movements is abundant. As we learned in Chapter 2, the scientific method involves several steps and helps us study the rules that govern the Universe. Importantly, generalizations posited by scientists must be supported by objective evidence if they are to be taken seriously. The story of plate tectonics provides a great example of the way science was used to slowly develop a widely accepted theory. It began with an observation made in the 17th Century that many of the coastlines of the world's continents appeared to be complementary, like puzzle pieces. Was this mere coincidence or had these edges once been connected to each other? As outrageous as it seemed to many people, the following hypothesis was offered: the Earth's continents are not stationary, but instead, move. A single, giant landmass of the past somehow broke apart, resulting in the current arrangement. The notion of **continental drift** was born. In subsequent years and centuries many scientists collected data to test the hypothesis. Along with other pieces of evidence, the following five have elevated this idea to a theory.

1. The puzzle fit

Although coastlines are complimentary, the fit is even better if continental shelves (the coastal edges submerged by ocean) are considered. See Figure 3.6.a.

Figure 3.6.a. *Illustration of the so-called jigsaw fit of the continental shelves of Africa and South America. Woudloper, CC BY-SA.*

2. Fossil evidence

Fossils (preserved remains of ancient life forms) of the same land-based organisms are present on continents separated by wide oceans (Figure 3.6.b).

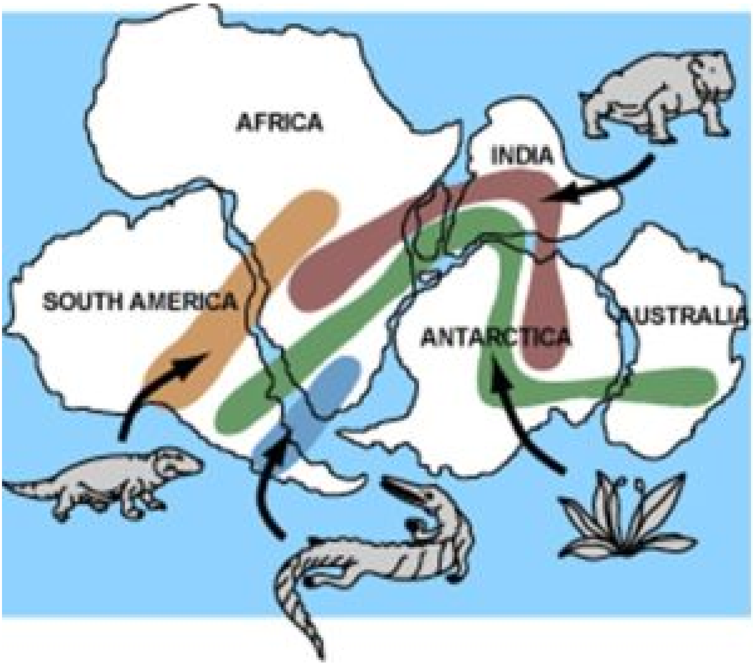


Figure 3.6.b. *Reconstructed single landmass from fossil evidence. Fossils of the same organisms are found on continents currently separated by oceans.* USGS, *Public Domain.*

Of the three most likely explanations offered by science, (c) is the one best supported by evidence.

(a) The same organism, with the same adaptations and traits, would have developed at the same time under different environmental conditions, for example in Australia, Africa, and South America. As we learned in Chapter 1, organisms are shaped by the specific stressors acting on them, and it is hard to imagine how *identical* organisms could develop in response to *different* temperature, light, nutrition, moisture, etc., conditions. We will return to these ideas in Chapters 5 and 6.

(b) The organisms would have migrated by swimming across thousands of kilometers of oceans, despite their apparent inability to do so (i.e., they lacked proper adaptations).

(c) The organisms developed, lived, and died on a large landmass made up of modern-day continents. After their remains were fossilized, the landmass broke up. Again, geologists have embraced this idea as the most probable, particularly when combined with all the other evidence supporting the idea of continental drift.

3. There's an ocean in the middle of a mountain range

Rocks and mountains of the same age, and with the same properties, can be seen on the western and eastern edges of the Atlantic Ocean. So, during a trip north through the Appalachian Mountains (USA) to the coast of Nova Scotia, you would notice the mountains end where the ocean begins. If you then boarded a plane and flew east, you would see the Caledonian Mountains begin at the coast of England and continue into Scandinavia (Figure 3.6.c).

Figure 3.6.c. *The Appalachian Mountains in North America and the Caledonians in Europe (approximate locations are shown in orange) contain identical rocks, structures, and fossils, suggesting continents were once part of the same landmass. Not to scale. Tentowtwo, CC BY-SA (modified by Kelsey).*

A study of the two mountain ranges would reveal something you might consider startling: they appear to have the same features, rocks, and fossils in them. Geologists interpret these findings to mean a single mountain range formed while the continents were part of Pangaea. After the continents moved away from each other, two separate yet identical landforms persisted on either side of a large ocean. Most scientists find the alternative explanation, that the mountains formed simultaneously in different locations, to be far less plausible.

4. Glacial remains in tropical areas

Glacial evidence found on the separate landmasses is inconsistent with stationary continents for the following reasons (Figure 3.6.d).

(a) Remains of past glacial events are present in areas that were too warm to support year-round ice (climate conditions have been estimated by a study of rocks of the same age).

(b) The patterns left behind by moving **glacier**s have been well studied and can only be reasonably explained if a single landmass was covered by ice before it broke apart. Then, after the continents drifted, the remains of the ancient glaciers were left in present day South America, Africa, Asia, Australia, and Antarctica. Otherwise, the past glacial movements do not make any sense.

Figure 3.6.d. *Glacial deposits found on modern Earth suggest the continents were once part of the same landmass. The figure shows a portion of Pangea with the approximate positions of glacial deposits visible today inside the white areas (modern Antarctica and parts of South America, Africa, and Australia). Kieff, CC BY (modified by Kelsey).*

5. Ocean-floor symmetry

Ocean-floor features on either side of the Mid Atlantic Ridge (i.e., the divergent plate boundary that divides the Atlantic Ocean) are symmetrically arranged, suggesting a spreading center. Chemical tests used to calculate the age of rocks yielded a crucial finding: materials in the center of the ocean, in and around the volcanic crack, are youngest, and ages increase as one travels toward North America and Africa. Notably, the age arrangement is symmetrical around the ridge, and the ages of the rocks on the coast of each continent are the same, about 200 million years old (Figure 3.6.e).

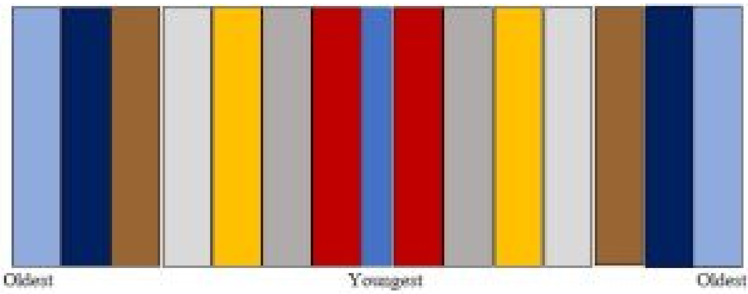


Figure 3.6.e. *Schematic diagram of the ages of rocks on the floor of the Atlantic Ocean. Bands of the same color are the same age (of course, the actual ocean floor is not marked like this!!). Map view. Kelsey,* CC BY-NC-SA.

Geologists conclude from these data that North America and Africa started to pull away from each other (roughly west and east) around 200 million years ago. The boundary between them was (and continues to be) a linear opening through which molten materials rise and then cool; the resulting rocks form new crust. In other words, the Atlantic basin has been increasing in size since **Pangaea** began to split. Other analyses have revealed similar arrangements of certain chemical properties in other oceans, further supporting the continental drift hypothesis.

How is the theory of plate tectonics useful? The emergence of this theory profoundly affected the way we view Earth. Geologists use their understanding of plate tectonics to help reconstruct the planet's long history and aid in their study of important phenomena. For example, the theory provides an explanation for continental drift and informs the study of the distribution and evolution of past and present organisms. It also provides some capacity to predict earthquake and volcanic activity (Chapter 7).

The rock cycle

Like mountains and other landforms, **rocks** can persist for very long periods of time, millions or even billions of years. Ultimately, though, they are temporary structures. The rock cycle describes the processes responsible for the formation, modification, and destruction of rocks. It is important to environmental scientists because it can be used to track the recycling of materials vital to many of Earth's systems (see Chapter 4 for more about the movement and cycling of materials). The major processes of the cycle are briefly summarized below and in Figure 3.7. Note it is not linear but represents multiple possible trajectories and outcomes.

Crystallization. At very high temperatures, in the range of 600 – 1300 °C, rock melts to a liquid (called **magma** if below the Earth's surface and **lava** if on the surface). This molten material typically moves upward from hot zones into cooler ones until it **crystallizes**, or turns into solid **igneous rocks.**

Uplift, weathering, and erosion. Tectonic forces can move rocks from low elevations to high (i.e., the process of **uplift**). In addition, soils and rocks under which existing rocks are buried can be removed. Rocks affected by uplift or uncovering can then undergo **weathering**, a collection of processes that fragment or decompose them. For example, chemical reactions can break bonds between atoms found inside rocks (Chapter 4), or physical disintegration can occur when water freezes and expands within cracks. The products of weathering, known as **sediments**, are subject to **erosion**, that is, they are transported away in water, ice, or wind (so, although related, weathering and erosion are *not* the same). Some of these sediments become new rocks (see the next item on the list), whereas others are incorporated into **soil**, a material that has a profound effect on farming (Chapter 9).

Deposition and lithification. Sediments produced from weathering move downhill until they are dumped, or **deposited**, by the agent transporting them (e.g., moving water). Very often, these materials settle in the bottoms of lakes, rivers, and oceans. As layers of sediments accumulate on the floor of a given waterway, they are squeezed together and buried. Under the proper conditions, these materials are bound together in what is known as **lithification**, to become solid **sedimentary rocks**.

Burial of existing rocks. Rocks can be covered again by soils, sediments, and other rocks. If temperature and pressure increase sufficiently, physical and chemical alterations can produce **metamorphic rocks**. Note that temperatures are not high enough to melt the rocks, so only solid-state changes occur. Existing metamorphic rocks can continue to change as they are buried. Eventually, rocks become hot enough to melt, and the process can be said to repeat itself.

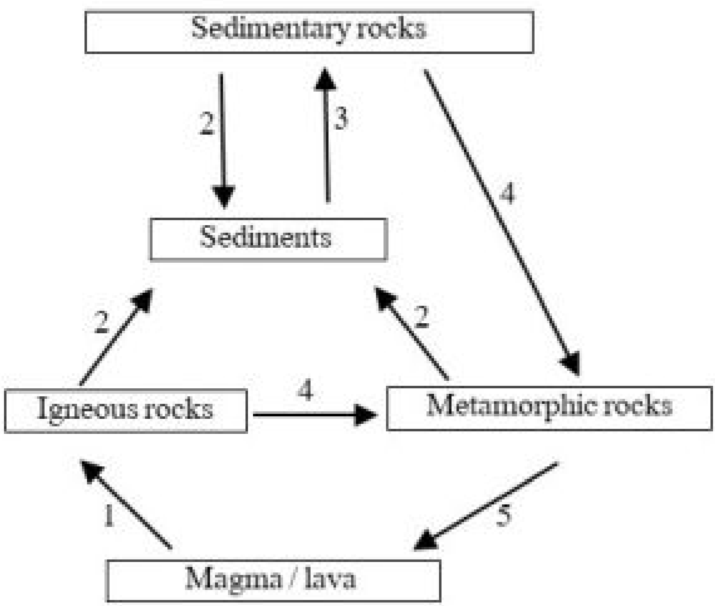


***Figure* 3.7.** *Schematic diagram of the rock cycle. 1. Cooling and crystallization of magma produces igneous rocks. 2. Weathering and erosion of existing rocks produce sediments. 3. Lithification produces sedimentary rocks. 4. Burial of existing rocks produces metamorphic rocks. 5. Burial and melting of existing rocks produces molten magma / lava.* Kelsey, CC BY-NC-SA.

3.1.2. Geologic time

Earth has a long history

Data collected by geologists can be used to make a very good case that Earth is ancient, some **4.6 billion years old** (see Box 3.1 for a brief explanation of the tools used to date the Earth). It will likely be unsurprising to you that describing such a vast stretch of time poses some substantial challenges. To help manage this enormous task, geologists have developed what is known as **the geologic time scale** (Table 1), a way to summarize major events that occurred during the past. It is divided into units based largely on **fossils** found in sedimentary rocks.

Table 3.1. *Abbreviated Geologic Time Scale*[1]. *Note that about 90% of geologic time is compressed into the era known simply as Precambrian, i.e., the relative sizes of the eras indicated by the colors are not to scale. Ages are approximate. The term vertebrate refers to organisms with backbones, invertebrates lack backbones.*

1. Modelled on USGS and Geological Society of America time scales.

Era	Period	Years ago (millions)	Notable events
Cenozoic	Quaternary	0-2	Modern humans appear
	Tertiary	2-65	Mammals begin to dominate
Mesozoic	Cretaceous	65-144	Dinosaurs become extinct
	Jurassic	144-213	Dinosaurs dominate
	Triassic	213-248	First dinosaurs, mammals, birds appear
Paleozoic	Permian	248-286	Modern insects appear
	Carboniferious	286-360	First amphibians and reptiles appear
	Devonian	360-408	Plants with seeds and boney fish appear
	Silurian	408-438	First land plants appear
	Ordovician	438-505	First vertebrates appear
	Cambrian	505-570	Invertebrates dominate oceans
Precambrian		570-3800	Prokaryotes dominate
		4600	Earth formed as molten ball

Table 3.1 is one way to represent the time scale. We will not concern ourselves with the details, but you should note some important highlights (years given are approximate).

1. Earth formed as a molten ball of material about 4.6 billion years ago.
2. We have found no evidence of any organisms living before 3.8

billion years ago.

3. Single-celled organisms known as **prokaryotes** were the only life forms present until 600 million years ago. More about these and other microbes can be found later in this chapter.
4. Complex organisms made up of more than one cell(i.e., they are **multicellular**) did not appear until Earth was 4.0 billion years old.
5. The ancient plants that became today's coal died about 300 million years ago.
6. Dinosaurs were present between 245 and 65 million years ago.
7. Mammals (the group that includes humans) did not become relevant until after the extinction of the dinosaurs (more about extinction is presented in Chapters 5 and 6).
8. Human-like organisms appeared 2 million years ago, although modern humans probably did not begin to develop until 200,000 years ago.

Box 3.1. How geologists come up with those extreme ages

There are two approaches used to establish the age of Earth's materials.

1. The first and older one is referred to as **relative age dating**. Put simply, geologists piece together the most likely series of events to account for an existing set of geologist structures. Such efforts are informed by a complex set of rules and principles, the details of which are beyond the scope of our need to know. The basic strategy is informed by the following assumptions (Figure 3.8).

(a) In a sequence of sedimentary rocks, materials get progressively older with depth;

(b) Features that cut across existing layers and structures are younger than the existing materials they disrupt; and,

(c) Rocks containing the same fossils are of the same age. The original materials making up plants, animals, or microbes have been replaced by non-living minerals and rocks, but they retain the shape of a long-extinct organism. So, dinosaur footprints left in mud 70 million years ago and impressions of ferns from 200 million years ago were preserved in rocks and are classified as fossils today. Although 350-year-old coal no longer looks like the materials from which it was derived, even it is called a **fossil fuel** because it too came from ancient plants, (Chapter 10). Fossils provide a lot of useful information to geologists, including the nature of past environments. They also can help establish the relative age of rocks.

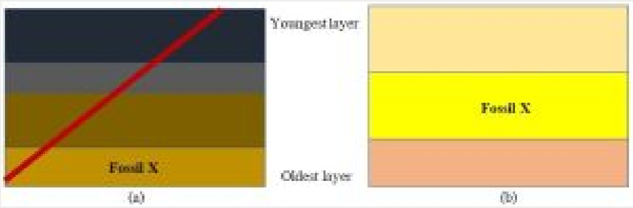


Figure 3.8. *Diagrams of two rock sequences to illustrate relative age dating techniques. Each color band represents a different rock type. The diagonal structure (red) occurred most recently, cutting across existing layers of rocks, a. Rock sequences a and b are separated by many kilometers, but the bottom layer of a is the same age as the middle layer of b because the same fossils (hypothetical organism X) are present in each. Note that rocks get progressively older with depth in both sequences. In cross section. Kelsey, CC BY-NC-SA.*

2. The second method evolved more recently and uses **radiometric age dating** to assign an age, in years, to geologic materials. This strategy employs chemical analyses and calculations, the details of which are not necessary for our purposes. The essential steps involved in this technique are briefly summarized below.

(a) Certain atoms are said to be radioactive, which means they change, or **decay**, with time (atoms and other chemical terms are defined in Chapter 4). We call the radioactive starting materials **parent atoms** and the products **daughter atoms**.

(b) Results of previous studies tell us how quickly different atoms will decay and what atoms they will become after they change. For example, certain types of radioactive potassium atoms are converted into non-radioactive, or **stable**, argon atoms. This transformation is very, very slow: in a given sample of radioactive potassium atoms, 50% of them will decay to argon in 1.25 billion years. Put formally, the **half-life** of potassium is said to be 1.25 billion years (half-lives of other atoms range from less than a second to billions of years).

(c) The relative amount of parent and daughter atoms present in a rock can be determined; and,

(d) Information about decay products and half-life are used to calculate the amount of time that must have elapsed to yield the current parent-daughter ratio in a sample.

The relevance of time

Modern geology is underlain by one critical principle: Earth is extremely old. In fact, unless one allows for the billions of years of history supported by geologic evidence, it is essentially impossible to use science to explain the processes shaping both interior and surface features. Consider our discussion of plate tectonics, above. Given the slow rate of movement—on the order of 2 centimeters each year—hundreds of millions of years were required to form the Atlantic Ocean after Pangaea started to break up. Weathering and erosion of tall mountains also take an enormous amount of time. Depending on several variables, the rate of removal of materials is in the range of a millimeter per year. So, a structure that is many kilometers tall, like Mt. Everest in Asia, would take millions of years to be worn down.

The nature of past geologic processes

Generations of people have debated whether the physical features we observe today were primarily formed through sudden, rare events or as the result of slow, regular processes. One dominant and long-standing assumption that Earth is young, approximately 10,000 years old, allows for just the former explanation; that is, given the very limited amount of time available, dramatic and short-lived upheavals are the only possibility. Geologists of the past, therefore, had to somehow reconcile considerable evidence that Earth is shaped very gradually with the constraint that history was not sufficiently long—by many, many millions of years—to allow such obviously time-consuming mechanisms to produce mountains, oceans, and the like. Modern geologists find themselves in a different position, as described in the previous paragraph: they assume Earth is old enough for even weak yet steady processes to both form and alter surface structures. So, to return to a point made above, given enough time, just 1 millimeter per year of weathering *can* completely obliterate the tallest mountain range. Even with this expanded view of time, different proposals about Earth's history persist.

Uniformitarianism. In short, this principle holds that the rules governing today's world have always been in force. So, since we do not currently observe instantaneous formation of mountains, oceans, rivers, or valleys, we cannot assume that they sprung up quickly in the past either. We must instead draw upon the slow and steady forces working now to explain historical events. For example, if tectonic plates move slowly in the modern world, they must have done so in the past. Clearly, one of the major assumptions upon which Uniformitarianism is founded is that Earth is very ancient. Uniformitarianism is often summarized very succinctly in the following way: **"The present is the key to the past."**

Catastrophism. Unlike Uniformitarianism, this idea assumes that the past must have been characterized by a much more active planet than we see today, one that featured many powerful forces working to quickly and profoundly shape the surface. For example, rather than requiring 200 million years to form, the Atlantic Ocean might have opened up overnight due to some mechanism that is no longer active. Importantly, catastrophism does not require an ancient Earth: everything currently present, and all the evidence suggesting billions of years of time, could have been produced in a few thousand years.

Geologic actualism. Our current model is a hybrid of the first two and holds that Earth goes through extended periods of relative quiet and inactivity, but short bursts of intense activity can quickly transform the surface. For example, as we see today, phenomena such as earthquakes and volcanic eruptions can bring about sudden changes, but day-to-day forces work slowly.

3.2. MICROBIOLOGY

We are absolutely surrounded by organisms we never see. Like it or not, bacteria, protozoa, fungi, and algae are in or on our food, water, air, skin, mouth, intestines, clothes, furniture, utensils, railings, toilets—the list is very, very long. It would be safe to say that very few places on this planet lack them. They tend to go unnoticed by most people, however, because of their extremely small size: without the aid of a magnifier (e.g., microscope), these organisms are invisible. Whether most humans are conscious of them, the enormity of their ongoing role cannot be overstated. A basic understanding of microbiology is therefore indispensable to a study of environmental science. Like our approach to geology in the first part of this chapter, we will only take a brief look at this broad and complex field.

3.2.1. Cells

All organisms, including those of interest to microbiologists, are made up of at least one microscopic unit known as a **cell**. The properties and ideas about cells most important to us are as follows.

A small system

A cell can be thought of as a living system, one that receives inputs such as nutrients and information and releases outputs such as waste products. Sizes range from around 1 μm (a micron, or 0.0000001 meter) in prokaryotes to 10 μm in eukaryotes.

Distinct from its environment

Cells live within an environment (e.g., soil or water) but are identified as distinct entities by their **cell membrane** (also called the plasma membrane), a flexible outer boundary that controls much of what passes in and out of them. Some cells have an additional rigid boundary known as a **cell wall** (Figure 3.9).

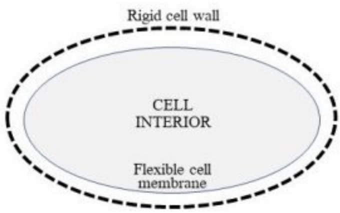


Figure 3.9. *A simple diagram of a single cell. All living cells are bounded by a membrane, those in plants, algae, and most bacteria also have a cell wall (animals and protozoa lack cell walls). Kelsey, CC BY-NC-SA.*

Component parts

The interior of a cell (i.e., the space enclosed within the cell membrane) contains water and several other sub-cellular structures. **Eukaryotic cells** have well-defined interior components, or **organelles**, surrounded by their own membranes. Possibly the most famous of these is the nucleus, the site of information storage. There are others, though, including **mitochondria**, organelles responsible for maintaining appropriate cellular energy levels. On the other hand, **prokaryotic cells** lack distinct membrane-bound organelles. Many of the same tasks (and more) carried out by eukaryotic organelles can be accomplished within prokaryotic cells, but the interior structures are less obvious and well defined in the latter. See Figure 3.10, a simple and generalized diagram of different cell types.

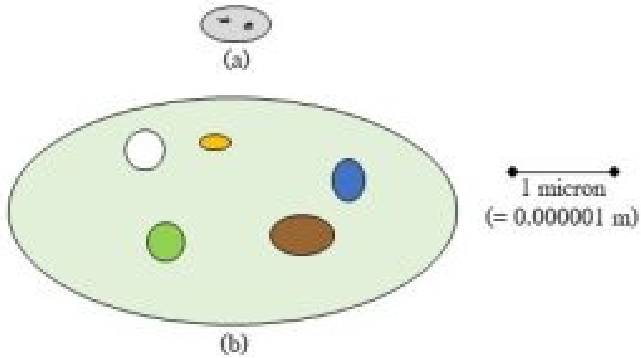


Figure 3.10. *Simplified models of basic cell types: prokaryotic (a); eukaryotic (b) (see the main text for more). The relative sizes of these cells are as suggested by the diagram. Note the approximate scale, to the right of the figure.* Kelsey, CC BY-NC-SA.

Reproduction

Cells can copy themselves and produce offspring through a variety of mechanisms (Chapters 5 and 6).

How many cells per organism?

Most microorganisms are **unicellular**, that is, made up of just one independent cell. Some organisms are **multicellular**: they are made up of many cells (in the case of humans, something like 10^{13} native cells!) that interact to form a larger living system. **Colonial** is a third type of growth, and involves many individual cells that occur together, interact, and probably communicate, but their interdependence and specialization are not as developed as they are for true multicellular organisms.

3.2.2. The microorganisms

We learned that these are too small to be seen with the naked eye. Other than their size, though, it would be wrong to assume they are all the same—this is an immensely diverse assemblage. To simplify things, microorganisms are subdivided into five groups, each with several important defining characteristics.

Bacteria

Relative simplicity. They are prokaryotic, unicellular, short lived, and fast growing.

Adaptability. Having fast growth and relative simplicity, they can quickly and effectively respond to stressors (Box 3.2). This trait enables their long-term success (Chapters 1, 5, and 6) but also can create some substantial human health concerns such as the development of antibiotic-resistant bacteria (Chapter 5).

> **Box 3.2. Simple and fast: advantage bacteria**
> Two important traits give bacteria a remarkable ability to quickly adapt to changing environmental conditions. First, their **DNA** is very simple compared to that found in eukaryotes. Why does this matter? DNA is the molecule in which all the information needed to make a copy of a cell is kept (also called its **genome**). It can be thought of as a code that is translated into physical traits such as size, shape, and chemical properties. If a change or **mutation** to the code occurs (a common event), then the dependent traits are likely to be altered as well. Often, the result is bad for the cell, but sometimes the new traits provide an advantage in

descendent cells (Chapter 6). DNA can be classified as simple, as it is in bacteria, if it only carries information that is converted into tangible characteristics. For the most part, bacterial genomes contain no extraneous information—all the code is put to use. In humans and other eukaryotes, the genome is far more complicated. Importantly, although our DNA certainly carries information that is decoded to construct physical structures and dictate basic characteristics (e.g., eye, hair, and skin color) it also contains sections that go largely unused. Thus, a mutation to our DNA is not guaranteed to affect any observed traits, particularly if the change occurs to one of those untranslated sections (Chapter 6). There are other important differences, but the bottom line is: mutations are much more likely to bring about fast and obvious changes as genomes get simpler. Second, bacteria reproduce extremely rapidly. In some cases, a new generation of cells will appear every 10 minutes, a far cry from the decades typically needed for the human population to double. This enormous time difference is relevant because substantial changes to the characteristics of bacterial offspring could become evident within a few hours, allowing for rapid adaptation to new stressors (See also Box 5.3). Many millennia (or much more) are likely to elapse before mutations to animal genomes are relevant, if they ever manifest at all.

Distribution. Members of this group can be found in nearly every environment on Earth, from deep oceans to mountain tops to deserts to forests to soils to other organisms. They can survive a range of conditions relative to oxygen and light availability, in some cases even thriving in frozen, boiling, acidic or other extreme places. Not surprisingly, the word *ubiquitous* is often used to describe bacteria.

Nutrient sources. As we will see in Chapter 5, although all organisms need to fulfill the same basic requirements if they are to survive, a variety of possible strategies is used within the biosphere. Bacteria are by far the widest ranging, with everything from the remains of dead organisms to sulfur-rich gases emitted from vents on the ocean floor, to sunlight, to **pesticides**, drugs, munitions, industrial chemicals, and other synthetic products serving as food for different members of this group. We will explore these and other nutrient sources in Chapters 4 and 5.

Roles. Bacteria are integral to many of the processes we will consider in upcoming chapters, including recycling of Earth's natural materials (Chapter 4), the functioning of ecological systems (Chapter 5), human health (Chapter 5), farming (Chapter 9), water pollution (Chapter 11), and waste management (Chapters 13 and 15). Some even aid in the production of food such as cheese and yogurt. Note that not all bacteria have the capacity to do everything listed; these traits and roles can be found *within* this diverse group of organisms.

Physical form. Viewed under a microscope, bacteria with wildly different survival strategies, food sources, and functions often appear to be the same size and shape: most are about 1 **micron** in diameter and round, oval, or spiral (Figure 3.11). As such, they provide a good lesson in the folly of using physical characteristics to judge degree of relatedness among organisms. Far more important are differences and similarities among genetic characteristics (we will return to this topic in Chapter 6 when we consider biodiversity). Finally, bacteria tend to possess the ability to move (i.e., they are **motile)**.

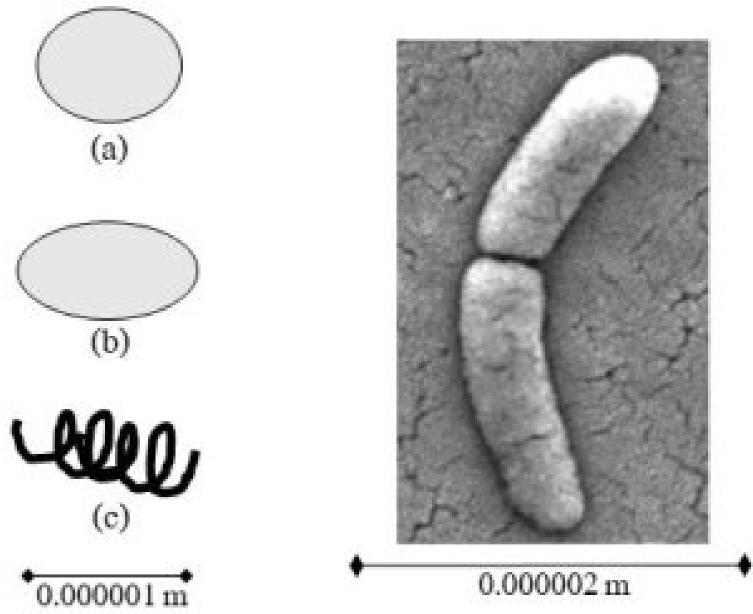


Figure 3.11. *Three common bacterial shapes, left (a) and an image of two bacterial cells, right (b) Note the different scales, below the figures. Kelsey, CC BY-NC-SA (a); US CDC, Public Domain (b).*

Relevance to humans. Their extensive roles in Earth's systems make bacteria as vital to us as to any other organism. Additionally, though, they interact directly with humans in important ways: some can harm us by causing diseases such as strep throat, gonorrhea, cholera, and tuberculosis, whereas some provide benefits through aiding in digestion and helping to keep harmful microorganisms out of our bodies (see Chapter 1 for more on these two phenomena).

Fungi

Moderate complexity. These are **eukaryotic**, **colonial** or

unicellular, fast growing and short lived. Some forms are motile, others are largely stationary.

Adaptability. Fungi can **adapt** to changing surroundings, but their more complex cell structure makes them slower to do so than are bacteria.

Distribution. It would not be appropriate to call fungi ubiquitous (just one way they are distinct from bacteria). However, they are found in many environments and can survive harsh conditions like low availability of water and food. Most require oxygen, but a few, namely yeasts (famously responsible for the fermentation of grain to alcohol) do not. Fungi can be found in soil, air, and water, as well as in or on other organisms.

Nutrient sources. Different fungi have the capacity to use different food types, but the list of potential nutrient sources is far shorter here than it is for bacteria. Most of them digest the remains of dead organisms, and some will also break down synthetic pollutants (Chapter 15). Importantly, certain members of this group can manage very large and unwieldy materials: they release substances into their surroundings that begin the process of digestion *outside* of their cells. When the materials have been reduced in size, they can enter the fungal cells and provide needed nutrients. Unlike certain bacteria (and algae, which we will explore next), fungi *cannot* obtain energy from the sun.

Roles. In most systems, fungi are **decomposers.** In short, they break down waste products and remains of other organisms; this action helps to recycle important nutrients (we will see more about **decomposition** in Chapters 4 and 5).

Physical forms. As a group, fungi are physically diverse relative to bacteria. Fungi can take on many different shapes, sizes, and colors. Additionally, fungi include unicellular forms like anaerobic yeast as well as large multicellular forms like mushrooms and bread mold. Individual cells are about ten times larger than those of **bacteria** (about 10 microns, 0.00001 meters), but still microscopic. Figure 3.12 shows two of the many forms fungi can take (see Box 3.3

for some thoughts on how fungi like the large ones in the figure can be classified as microorganisms).

Figure 3.12. *Two common forms of fungi. Left (a) Peter Stevens, CC BY; right (b) User:Piotrus, CC BY-SA.*

Box 3.3. But…mushrooms are not microscopic!

You might be wondering why a group that includes the clearly visible mushroom is classified as a microorganism. The explanation is that the mushroom is only the above-ground portion of a large network of largely microscopic cells that grow in soil. Also, because of certain other traits, many organisms visible during some part of their life cycle are simply grouped here. Be all that as it may, it still is a bit confusing, and there is not 100% agreement that mushrooms should be on this list. The common characterization of algae (below) as microorganisms results from similar considerations.

Algae

Moderate complexity. Like fungi, these are unicellular or colonial eukaryotic organisms. Algae can be mobile or stationary.

Adaptability. As **eukaryotes**, they are less readily adaptable than **prokaryotes**.

Distribution. Algae require light to grow (as described below), so they rarely are seen in dark environments. They also tend to require water, so they are common in rivers, lakes, oceans, and wet surface soil.

Nutrient sources. As we will see in Chapters 4 and 5, some organisms, including algae, can use the sun as an energy source and convert important building blocks like carbon atoms from unusable to usable forms. So, unlike fungi, they do not consume remains of other organisms

Roles. Their ability to convert light energy and carbon into usable forms means algae can provide nutrients to organisms that cannot carry out those important conversions. Put another way, they act as food for other organisms. Again, Chapters 4 and 5 will expand on this topic.

Physical forms. Like fungi, this group includes truly microscopic forms such as diatoms, as well as giant kelp, organisms that can be two or more meters long (recall Box 3.3). Individual cells are similar in size to those of fungi, about 10 microns (Figure 3.13).

Relevance to humans? Algae play very important roles in many systems on Earth, so they are obviously indirectly relevant to us (Chapter 4 and 5). Otherwise, we interact with them less than we do with bacteria and fungi. Human diseases caused by algae are rare, and although some are edible, algae play only a minor role in meeting the nutritional needs of most people.

Figure 3.13. *Forms of algae: right (a) in water; bottom (b) on wet rocks . Swarnyk, CC BY-SA (a); Hodnett, CC-SA (b).*

Protozoa

Moderate complexity. These organisms are generally **unicellular** (a few are colonial), eukaryotic organisms. Protozoa are mobile.

Adaptability. They are less readily adaptable than are prokaryotes, although their **DNA** is simple relative to that in animals.

Distribution. Protozoa can be found in soils, water, and inside organisms, but not in places with extreme conditions of temperature or chemical properties.

Nutrient sources. By and large, protozoa are like fungi in their need for nutrients. Then tend to eat bacteria and other microorganisms or free-floating debris that originated from organisms.

Roles. Protozoa both eat smaller organisms and serve as food for larger organisms. They are therefore an integral part of the recycling of materials in natural systems.

Physical forms. Compared to bacteria, this group includes organisms with a range of diverse forms. Like fungi and algae, individuals are also approximately 10 microns in size (Figure 3.14).

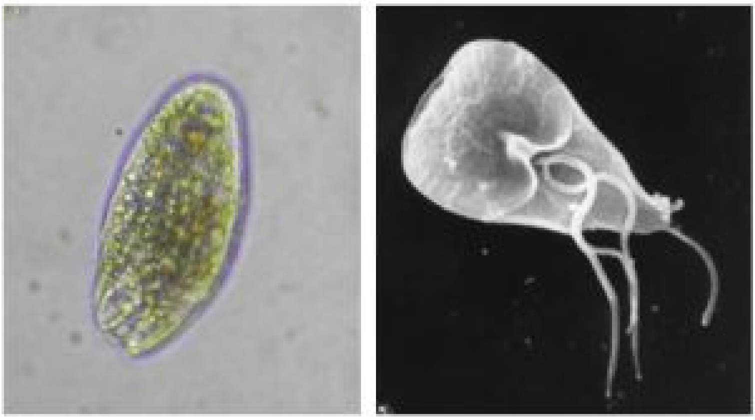


Figure 3.14*.* *Two common protozoa: left, paramecium (a); right, the water-borne intestinal pathogen, Giardia lamblia (b). Hobern, CC BY (a); US CDC, Public Domain (b).*

Relevance to humans? Protozoa are directly important to us largely because of the diseases they can cause. As unlucky hikers and campers can tell you, extreme intestinal distress can be the result if an organism called *Giardia lamblia* ends up in your gut. It can be consumed through contaminated water and has the potential to be lethal due to the massive dehydration it brings about.

Viruses

These acellular (i.e., not, or without, cells) entities are not proper organisms like the others on this list, but they are clearly biological and relevant to us. The details of their existence are beyond the scope of this textbook, but a few important properties and roles are worth noting. Put simply, viruses invade the cells of other organisms: microorganisms, plants, and animals are all susceptible to attack. Once inside a cell, they can use the machinery of their host to make copies of themselves. In fact, viruses depend on their hosts for reproduction and transportation. A possible consequence of viral infection is disease, that is, the host organism is damaged. A very short list of human diseases caused by different viruses includes COVID-19, herpes, AIDS, Ebola, and the common cold. Finally, they are extremely small, on the order of 0.01 – 0.001 microns. Put another way, bacteria are 100 to 1000 times larger than a typical virus. See Box 3.4 for more about microbial sizes.

Box 3.4. Just how small is microscopic?

By now you should realize that, although all are invisible to the naked eye, microscopic organisms are far from uniform in size. We have seen an enormous range, from tiny viruses that are 0.000000001 meters (0.001 microns) to giant protozoa that can sometimes top out at nearly 0.001 meters (1000 microns). Keep in mind that there is at least as much of a size difference among microbes as we see in our macroscopic world (say, a whale vs. an insect). If we include even smaller, non-living entities like atoms and other particles (Chapter 4), the range within the invisible world increases by at least another 1000 fold or so. Clearly, "small" is a term fraught with uncertainty.

3.2.3. Earth's oldest organisms

Most scientists accept that prokaryotes have dominated this planet for most of its history and are likely to do so into the future. Recall from our look at geologic time, above, that evidence of single-celled lifeforms is 3.8 billion years old, whereas multicellular organisms did not arrive until 600 million years ago. Interestingly, although an examination of data strongly suggests that the earliest bacteria specialized, evolved, and eventually gave rise to the rest of the biosphere, many modern-day prokaryotes still seem to closely resemble their ancient ancestors. So, microbiologists have been able to learn a lot about ancient organisms by studying current ones. Additionally, information gleaned from studies of bacteria has provided a lot of data about the development of eukaryotes. Knowledge about them also informs our study of the development of Earth's physical environments. As we know, the geologic time scale employs biological criteria to establish its organization and chronology. We will look more at the development of living things and the mutual effect they and their environments have on each other in Chapters 5 and 6.

3.2.4. Applications of microbiology

Knowledge of microbiology is applied to many areas we value as both scientists and citizens of the modern world.

Medicine and public health

As noted above, different kinds of microorganisms can make us sick, with consequences ranging from mild to severe. Any organism that can cause disease is termed a **pathogen**. Although most microorganisms are *not* harmful to us, those that are receive a great deal of attention. Notably, we use information about factors affecting their survival to help improve public health. Box 3.5 describes how science was used historically to understand the ways microorganisms bring on diseases and other undesirable effects.

Box 3.5. Disease, food spoilage, and magically appearing mice: get me a scientist!

We saw in Chapter 2 that the scientific method can be used to better understand the rules governing the universe and to test explanations about important phenomena. Widely held assumptions, even those not supported by solid evidence, can be difficult to discredit, and the efforts of multiple scientists working during decades or centuries are often necessary to alter entrenched theories. For example, several ideas about the causes of illness and various kinds of contamination were evaluated and ultimately rejected only after many years of scrutiny. Two important instances of paradigm shifts are worth a few minutes of our attention here because they demonstrate how knowledge evolves as well as the way science can be used to challenge even long-standing beliefs.

1. What causes disease? Most modern people know that specific entities—often referred to as **germs**—can move among animate and inanimate objects to induce infectious diseases. People of the past, however, did not have access to such information and lived in a world dominated by microscopic forces they did not understand. One of the more vivid historical examples of the way ignorance led to suffering occurred in Europe of the middle 1800s, near the start of the Industrial Revolution. During this period, The Thames River in London, England, became increasingly polluted due to the rapid growth of the human population in the city. Since indoor plumbing was not yet available, and the

river was a moving and ever-present receptacle, people dumped their excrement into its water. Other trash ended up in there as well. Unfortunately, The Thames also served as the major source of drinking water for many Londoners. Apart from the aesthetic reasons that are likely obvious, drinking human sewage is a very bad idea because it contains a high density of microscopic pathogens. Perhaps not surprisingly, an epidemic of water-borne cholera devastated the population, causing the deaths of many thousands in a short period of time. It may seem inconceivable that people would consume their own waste, but subsequent scientific inquiry has revealed to us two things the citizens of 1850s London did not realize: infectious diseases are caused by discrete biological entities, and human feces contains a lot of pathogens. It was a study of this very cholera outbreak that led to some important new knowledge about the transmission of human diseases. A physician named John Snow noticed that rates of cholera infections were far higher among people who drank from the river than in those who used different sources of water. He hypothesized that something carried in the water was responsible for the disease, not so-called bad vapors (known as miasma), the prevailing explanation at the time. How did Snow test his hypothesis? He convinced the city leaders to shut off access to the river and force people to haul drinking water from inconvenient wells located some distance away. The result: cholera rates dropped dramatically, and public health microbiology took a giant leap forward. We will return to the methods we currently use to slow the spread of diseases like cholera, notably the ways we minimize the amount of sewage released into our drinking water, in Chapter 11.

2. Who put the bacteria in the potato salad? The reason foods like milk, meat, and various mysterious leftovers can go bad is generally known to have something to do with the action of microbes. Like we saw above with diseases, though, people of the past did not understand the mechanism by which food becomes rotten and inedible. One of the most persistent historical explanations for food spoilage (and contamination in a broader sense) was called **spontaneous generation.** For centuries, people assumed that the organisms responsible for spoilage arose from *within* food, not from without. We know today that the mold and other agents growing on a loaf of bread after a few weeks got there due to contamination, that is, fungal spores, bacterial cells, or even insect eggs were transferred in from a source *external* to the food. Even more remarkable to a modern audience, the appearance of rodents inside piles of dirty clothes or food pantries was typically attributed to spontaneous generation rather than the movement of such pests into places they found attractive and habitable: yes, instant, fully grown mice! Why did this explanation of contamination seem reasonable enough to dominate for so long? In short, scientists made simple observations of materials before and after they rotted and concluded that the newly appearing organisms had to be produced *by* those materials. Although maggots could be seen with the naked eye, the eggs (which eventually hatched into

those maggots) laid by flies landing on the food were not so obvious. In some ways, the development of the microscope in the 17th Century caused even more confusion because it allowed scientists to see clear differences between unspoiled and spoiled food, with the former lacking microorganisms and the latter teeming with them. Unfortunately, it did not reveal the cells landing on fresh food and their early stages of growth. Only the end result, obvious spoilage, was observable. People who did not accept the prevailing paradigm, as is so often the case, used the scientific method to both challenge spontaneous generation and test hypotheses about alternate explanations. As early as the 1600s, experiments were conducted to assess whether covering food would prevent contamination. Sure enough, when all other variables were equal, exposed meat ended up infested with maggots, whereas meat sealed in a container did not. Despite these findings, spontaneous generation was not generally rejected until Louis Pasteur conducted a rather famous experiment in the 1800s. In essence, he showed that sterile liquid remained so until and unless it was exposed to non-sterile air. The belief that organisms could pop out from the center of non-living material was finally debunked—after centuries of tests!

*Diseases that can be transmitted among people, for example the common cold, are known as **infectious**. Diseases such as cancer, multiple sclerosis, and diabetes* *do not move from person to person but are caused by some other mechanism; they are* ***non-infectious.***

Food and drink

Humans have learned how to harness microorganisms to produce foods like cheese, yogurt, and bread, as well as beer and wine (noted above). The science of microbiology, therefore, can profoundly affect nutrition and economic well-being.

Waste clean-up

As we will see in Chapters 11, 13, and 15, some microorganisms are able to break down our unwanted materials. A lot of research is done to optimize the growth and activity of the bacteria and fungi that provide such valuable services to us.

Agriculture

The actions of certain microorganisms are critical to the recycling of nutrients used by crops. We will return to this topic in Chapters 4 and 9.

THE CHAPTER ESSENCE IN BRIEF [2]

The science of geology allows us to study Earth's past and present as well as its future trajectory. It provides knowledge about the physical forces shaping organisms and materials vital to humans. An understanding of microbiology gives us insights into the largely unseen organisms without which our lives would be impossible.

Think about it some more... [3]

Why do environmental scientists care about Earth's

2. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 3 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

3. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

history? What difference does it make to those of us living today?

Do you think the theory of plate tectonics is appropriately known as a "theory"? Would you prefer to call it a "conclusion" or a "law" (as those terms are described in Chapter 2)? Why?

Would you assume that bacteria and other organisms too small to be seen with the naked eye exist if you could not see them directly? What evidence do we have that they are all around us?

ADDITIONAL RESOURCE

See also a link to the time scale as endorsed by the USGS: http://pubs.usgs.gov/fs/2007/3015/fs2007-3015.pdf

4. Earth's Environments, Cycles, and Energy

JASON KELSEY

In Chapter 1 we were introduced to topics studied by environmental scientists, including factors influencing the availability of materials on Earth. There is good reason for our interest in this area: both human and natural systems must have ready access to necessary nutrients if they are to function. Here we will examine more details about the storage, movement, and recycling of some vital materials. We will also take our first look at the fundamentals of energy, a topic important to natural and human systems and one that will appear again in upcoming chapters.

Key concepts

After reading Chapter 4, you should understand the following:

- The characteristics of Earth's four spheres as identified and studied by environmental scientists
- How Earth is a system that is closed with respect to materials and open with respect to energy
- The chemical, physical, and biological processes responsible for the movement and recycling of Earth's

atoms among its four spheres

- The major reservoirs and pathways of the carbon, nitrogen, and phosphorus cycles and why these cycles are important to both Earth's systems and environmental scientists
- How both insufficient and excess nutrients can disrupt a system
- The reservoirs and pathways of the water cycle
- How atoms in the lithosphere are moved and recycled through the tectonic and rock cycles
- The sources of energy for Earth's systems
- How energy is converted from high- to low-grade forms
- How the first and second laws of thermodynamics limit all of Earth's processes and systems

Chapter contents

4.1. EARTH'S SPHERES

Materials move within and among Earth's many systems. Before proceeding with a detailed study of cycles, we will look at the four basic zones or **spheres** within which materials of interest to us are found.

4.1.1. Atmosphere

Put very succinctly, this sphere is the envelope of gases surrounding the Earth. It extends from the surface outward, up to approximately 10,000 km into space in some places. The atmosphere is subdivided into several horizontal layers (also known as spheres), of which only the lower two will receive much attention from us. The altitudes listed are approximate[1]. See Figure 4.1 for a schematic diagram of the atmosphere.

1. Altitudes in the Atmosphere according to data provided by NASA (NASA.gov).

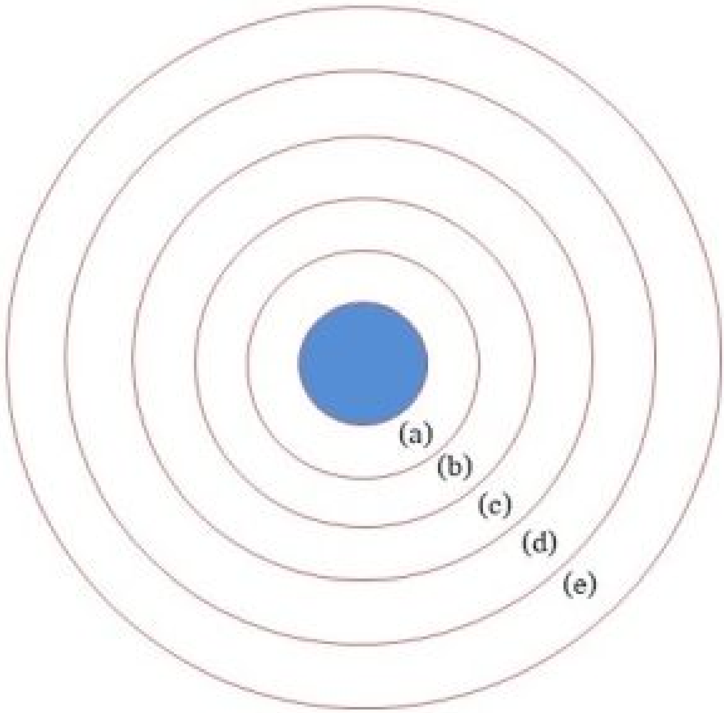


Figure 4.1. *Idealized view of Earth (center) and its atmosphere (open circles). Not to scale, in cross section.*
(a) Troposphere (b) Stratosphere
(c) Mesosphere (d) Thermosphere (e) Exosphere
Note: the ionosphere overlaps spheres c – e and is not shown. See main text for more details. Kelsey, CC BY-NC-SA.

Troposphere: 0 – 15 km

This is the part of the atmosphere with which we have direct contact and the place in which weather events occur. The bulk of atmospheric gases are found here, including dinitrogen, N_2 (makes up about 78 %), dioxygen, O_2 (21 %), argon, Ar, and carbon dioxide, CO_2 (together, these last two make up a little less than 1 %), and many other gases present in very small amounts. Note that the list in the previous sentence reflects the *current* atmosphere. A lot of evidence found in rocks and ice indicates that the composition of tropospheric gases has changed throughout history. Notably, the O_2 that organisms like us need for survival only appeared about 2.5 billion years ago (about 2 billion years after Earth was formed, as we saw in Chapters 1 and 3). So, the troposphere is a very dynamic layer, one which interacts with and is influenced by living and non-living systems at the surface. The outer limit of this zone, that is, the boundary between it and the next sphere, is called the tropopause. We will see more about the **troposphere** when we study air **pollution** in Chapter 14.

Stratosphere: 15 – 50 km

Unlike the troposphere, which gets colder as one travels outward, here temperature increases with altitude. This difference is due in part to chemical changes. Most importantly for our study, atmospheric ozone (O_3) is concentrated in this layer (Chapter 14).

Mesosphere: 50 – 90 km

This zone is important to life on Earth because it is where the bulk of meteorites entering the atmosphere burn up before they strike the surface.

Thermosphere: 90 – 500 (up to 1000) km

This is often thought of as the boundary between Earth's atmosphere and outer space.

Exosphere: from the top of the thermosphere (500 – 1000 km) to 10,000 km

Human-launched satellites orbit within this zone.

Ionosphere: 50 – 1000 km

This zone overlaps the mesosphere, thermosphere, and exosphere. It is characterized by its charged nature—radiation from the sun generates a lot of ions and electrons (more about these particles is presented below). These particles reflect radio waves, enabling long-distance communication.

4.1.2. Hydrosphere

Formally, the hydrosphere contains all the water on Earth. Now, as will become increasingly clear, water is not found exclusively in this area. In fact, none of these definitions should be taken as inflexible delineations because the four spheres interact quite a bit. For now, though, hydrosphere is a convenient and useful term, one that we will explore in detail shortly.

4.1.3. Lithosphere

As we learned in Chapter 3, the lithosphere is the brittle, outermost layer of the solid Earth and is comprised of the **crust** and the upper **mantle**. The shape of the planet's surface features, as well as the cycling of its materials, are linked to the movement of lithospheric plates across the top of the asthenosphere (see Figures 3.1 and 3.2 in the previous chapter).

4.1.4. Biosphere

The fourth sphere simply refers to all living matter. Many subdivisions can be employed to define and organize the vast diversity of organisms found on this planet. For our purposes, though, only broad categories are necessary to understand the ways this sphere interacts with the others and the important roles it plays in shaping Earth's systems.

Land and water organisms

Members of the biosphere are found in many different environments. We will explore this distribution in more detail in upcoming chapters, but for now you should recognize two major categories. **Terrestrial** organisms live largely on land, and include plants such as trees, microorganisms such as bacteria, and animals such as rats, lizards, tigers, and humans. **Aquatic** organisms live in water, so whales, sharks, lobsters, jellyfish, algae, and non-terrestrial bacteria are grouped here. We can further divide this second group into those that live in freshwater environments like rivers and lakes and those that are found in saltwater, or marine, environments such as oceans.

Microscopic organisms (review Chapter 3)

Microbes are things too small to be viewed with the naked eye—a magnifier, like a microscope, is required to see them. Despite their practical invisibility, *living* microbes, generally called **microorganisms**, are immensely important to the functioning of Earth's natural systems. This very diverse group includes **bacteria**, **protozoa**, **algae**, and **fungi** (although not exactly alive, we also categorize **viruses** here).

Macroscopic organisms

Organisms in this group are large enough to be viewed without the aid of a microscope. These appear to dominate Earth's systems, but they are dependent on microorganisms for their survival.

Plants. An exhaustive description of these organisms is not necessary for this book, but a short overview of their most important characteristics will help us to understand the roles they play. Plants are multicellular, eukaryotic (Chapter 3), immobile organisms that extract water and many of the nutrients they need from the soil in which they are rooted. They are distinguished from animals, the other major macroscopic organisms on our list, in large part by their ability to use the sun for energy and **CO_2** in the atmosphere as their source of carbon (more in the discussion of the carbon cycle, shortly). In so doing, plants often form the foundation for land-based ecological systems. We will return to a discussion of the ways different kinds of organisms obtain necessary nutrients, including energy and carbon, as well as interactions among members of the biosphere, in Chapter 5.

Animals. These are also eukaryotic and multicellular. Unlike plants, though, animals tend to be mobile and use neither the sun nor CO_2 directly to meet their needs. Instead, they depend on plants (or plant-like organisms) for biologically useable forms of energy and carbon, that is, animals consume what plants produce. We will revisit these ideas later (notably, Chapters 5 and 9)

4.2. CYCLING OF MATERIALS

As we learned in Chapter 1, the assumption that Earth is closed with respect to materials is fundamental to our study of environmental science. Just as critical, though, is the fact that Earth is open with respect to **energy**.

We will learn more about the differences between materials and energy later in this chapter, but for now it is important to realize that **materials are recycled, whereas energy is not** (Figure 4.2 provides a basic model to distinguishes their behavior).

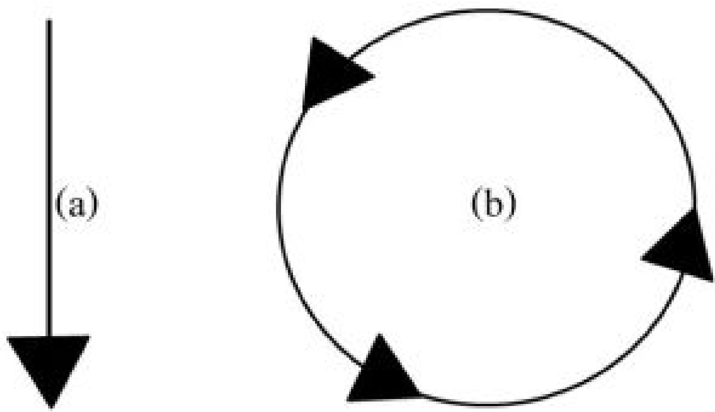


Figure 4.2. *Difference between energy and materials on Earth. a: energy passes through the system and is not recycled; b: materials remain within the system and are recycled. Kelsey*, CC BY-NC-SA.

4.2.1. A few words about chemistry

An exploration of Earth's materials requires a brief introduction to some essential terms and principles.

Chemistry

This is the science that investigates **matter** and changes it can undergo (informally, matter is just stuff—physical entities that take up space). Of particular interest are the fundamental composition and properties of the building blocks of everything in the universe and the ways these building blocks interact with (also *react with*; see below) each other. Often, chemists study what happens to matter on a very small scale.

Atoms

These are the tiny building blocks of all matter, the smallest units into which any element (see below) can be divided.

An atom is made up of even smaller components, subatomic particles such as **protons** and **neutrons** in its nucleus and **electrons** in clouds in its exterior. A neutral atom had an equal number of protons and electrons. Figure 4.3 is a simplified model of an atom.

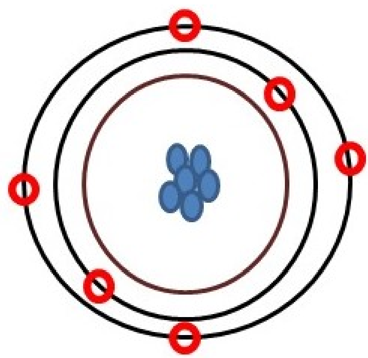


Figure 4.3. *Idealized atom. The inner portion (brown circle) is the nucleus; positively charged protons (blue) and neutrons (not shown) are here. Negatively charged electrons (open red circles) surround the nucleus. Not to scale: the radius of a typical atom is around 0.0000000001 meters. Kelsey, CC BY-NC-SA.*

Ions

An atom has the same number of positively charged particles (protons) as negatively charged ones (electrons). Atoms, therefore, have no net charge. For various reasons, the number of electrons in an atom may change, though, yielding an entity called an ion. Atoms are said to be **reduced** when they gain electrons and become negatively charged, whereas those that lose electrons are **oxidized** to take on a positive charge. Keep in mind that ions behave differently than neutral atoms, generally dissolving more easily in water and binding to other charged particles.

Element

This term refers to a substance that cannot be broken down into simpler units by ordinary means. It is related to a previous term, above, in that *atoms* with the same basic chemical properties are placed into categories called *elements*. As of 2022, there were 118 known elements, although some elements are more relevant to us, as both living beings and environmental scientists, than others. Note that the terms "atoms" and "elements" are sometimes used interchangeably, but they do not refer to the same things. One casual way to think about it: atoms with the same chemical properties are members of the same club, that is, they are grouped together as one element. For example, all the atoms in the universe having 6 protons in their nucleus are classified as the element carbon (C).

Bond

Individual atoms can be connected or attached to each other through various attractive forces or bonds. We use capital letters to represent individual atoms and short lines to represent bonds. For example, H – H is shorthand for a molecule made of two hydrogen atoms connected by a single bond, and O = O shows two oxygen atoms held together via a double bond.

Molecule

When two or more atoms undergo bonding, they form a molecule. For example, under the right conditions 2 hydrogen atoms and 1 oxygen atom can be brought together to form a water molecule, H_2O. See Figure 4.4.

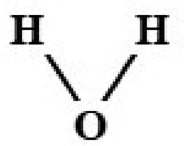


Figure 4.4. *Diagram of a water molecule. Letters are single atoms, lines are bonds.*

Compound

This is a molecule with more than one type of element in it. So, O_2 is classified as a molecule but not a compound, whereas H_2O is both a molecule and a compound.

Reaction

Two or more atoms brought close together can undergo changes as a result of chemical reactions. For example, atoms and electrons within molecules can be added, removed, or rearranged, and existing bonds can be broken and new bonds formed. Some of the energy released during certain reactions can be captured and used to power both living and non-living systems.

> *The entities changed in reactions are known as* ***reactants****, and the new entities that result are called* ***products****.*

4.2.2. Cycling of elements

Reservoirs and pathways of elements

Environmental scientists are interested in both the storage and movement of elements within and among the four spheres we described near the beginning of this chapter. Since atoms can be tracked through living and non-living reservoirs, we use the term **biogeochemical cycles** to indicate the scope of our work.

***Biogeo-chemical**: "bio" refers to living organisms, and "geo" refers to non-living rocks, air, and water. This is the study of chemical elements as they move through all Earth's systems.*

Recycling and reuse

Since Earth is closed with respect to materials, it does not gain or lose appreciable amounts of any atoms—what is present now has been present throughout history. Consider what this means: the atoms in its system must be used and reused over and over again. For example, an individual **carbon** atom has been a part of many, many different objects and systems since Earth was formed. That same C could have been a component of a gas in the atmosphere, then moved into the biosphere to form part of the tissue of an organism, then on to the lithosphere where it might have been incorporated into a rock and stored for millions of years. Eventually, the decomposition of the rock could have released that atom into water, where it combined with other atoms to build the shell of an oyster. There is a staggering number of possibilities. Keep in mind that, although atoms are reused, they often undergo changes in chemical form as they travel from one system to another. It will not likely be a surprise to you that a C atom in atmospheric carbon dioxide is connected to different atoms, via different kinds of bonds, than that same C atom in, say, a plant leaf. In fact, the C atoms in those two examples are different in one fundamental way: as part of the leaf it is said to be **organic**, whereas in the CO_2 gas it is in an **inorganic** form (see Box 4.1 for more).

Box 4.1. What is the difference between organic and inorganic?

The word "organic" appears in many places these days. For example, it can be used in reference to certain kinds of food and clothing to suggest something about the methods used to produce them. Often it is viewed as synonymous with "natural", "pure", and "good", particularly in marketing to health-conscious consumers. To a chemist, though, the term refers to a simple set of measurable criteria employed to evaluate substances. A chemical compound is classified as **organic** if 1) at least one carbon atom is present in it *and* 2) there is at least one C-H bond present, that is, at least one carbon is attached to at least one hydrogen atom. The term **inorganic** refers to a compound that does *not* meet both criteria required to be called organic. So, a compound that lacks any carbon atoms or has one C atom but no C-H bonds is inorganic. Let's compare the following three examples: $C_6H_{12}O_6$ (a molecule with 6 carbons, 12 hydrogens, and 6 oxygens), CO_2, and H_2O. The first is organic because it has at least one carbon and one carbon-hydrogen bond, but the other two are inorganic as CO_2 does not have any C-H bonds and H_2O lacks C. Note that characteristics such as source, appearance, and nutritional value are irrelevant here. Both classes of compounds can be produced by natural forces and both can be synthesized under laboratory conditions. We will use these terms to convey information about chemistry, *not* to express value judgments.

Importance of chemical form

As will become clear, both the location and chemical form of an element are equally important to us. It is vital to understand the distinction between the mere presence of an element in a sphere and the relative usability of that element to organisms (often referred to as **bioavailability**). In other words, an element that is abundant in an area will not necessarily be utilized by living things. Since only certain chemical forms of elements can be assimilated by members of the biosphere, processes that affect bioavailability of individual atoms, either by increasing or decreasing it, are of great concern.

The details vary from element to element, but cycles have some basic processes in common.

Fixation. Broadly applied, this term refers to chemical transformations that increase the bioavailability of an atom. Not all organisms can carry out fixation, so those that can are critical to the success of ecological systems (Chapter 5). Fixation tends to involve the production of relatively large molecules from the transformations and re-combinations of smaller ones.

Assimilation. Here, biologically available forms of elements are absorbed by members of the biosphere and used to construct various cells and tissues. An atom in one organism cannot be used by other organisms until it is excreted or moved via **consumption** (like it sounds, eaten by another; more below).

Decomposition. This general term can be applied whenever a large, complex object is reduced to smaller, simpler products. For example, the frustration brought on by a misbehaving computer could be alleviated, albeit temporarily, if one uses a hammer to smash the offending machine into tiny pieces—*physical* decomposition or disintegration would be the result (along with the need to spend money on a new CPU). A related, but fundamentally different kind of breakdown occurs when large complex molecules are *chemically* decomposed. In this case, bonds are broken by biological or non-biological means and individual atoms or smaller molecules are released. As we will see, fungi and some bacteria are responsible for a great deal of chemical decomposition. Although reality is a little more complicated, it is useful to imagine this process as essentially the opposite of fixation. Keep in mind our cycling context: decomposition frees atoms, allowing them to undergo any number of transformations, including fixation. Without this process, waste and remains of dead organisms (grouped together as **dead organic matter**) would pile up on Earth.

Respiration. This process enables organisms to gain energy from the food they eat. Note that decomposition and respiration are closely related because, when chemical compounds in food are broken down, some of the released energy can be captured by the organism and converted into a form that is biologically available. Often, that energy is stored inside the cells of organisms until needed. Realize that not all organisms use the same strategy to obtain energy. For example, **aerobic** organisms like humans need O_2 gas to keep respiration going. **Anaerobic** organisms (certain bacteria and fungi) use molecules other than O_2. We will see more about the ways organisms get nutrients in Chapter 5.

Excretion. Nutrients that have been assimilated will ultimately be broken down and released in waste such as urine and feces. Since excrement generally contains forms of elements that are poisonous to the organism producing it, the ability to effectively release it is crucial for survival. Interestingly, though, toxicity and usefulness are rather relative: waste to one organism can be food for another. To decomposers, your unwanted products make up a nutrient-rich meal. Insects feasting on a pile of animal dung similarly thrive on what you (likely) consider unpalatable and unhealthy. In addition to feeding themselves, these feces-eating organisms also make an essential contribution to the recycling of Earth's materials because they free up elements that would otherwise remain unavailable.

Weathering and erosion. These two processes are often grouped together, although they are technically distinct. Weathering is a general term that refers to the breakdown and change to Earth's surface features brought about by forces such as rain, wind, and ice. Erosion is responsible for the transport of weathering products—bits and pieces of rocks, if you will—to other locations (Chapter 3). Think back to the hypothetical C atom we imagined earlier for a minute: weathering and erosion would have been responsible for its release from the rock and movement to the ocean.

Diffusion. This physical process moves materials from locations in which they are relatively abundant, referred to areas of **high concentration**, into areas in which they are present in **low concentration**. Imagine what would happen if you added a few drops of grape juice to a clear glass of water. At first you would see obvious purple clouds floating near the top of the container. In a short period, though, differences in color would disappear as the juice spread out evenly throughout the system. In all likelihood, the entire volume of water would eventually become the same light purple color due to diffusion of the grape juice from its point of highest concentration—where it entered the water—to those with lower concentrations. Similar processes are active on much larger scales and profoundly affect the distribution of elements among systems. The flow of CO_2 between the atmosphere and the hydrosphere is just one important instance in which diffusion has important consequences for organisms living in water as well as global climate (more about these later).

Combustion. Objects can ignite and burn as a result of both natural and human activity. In this chapter we will focus more on the latter than the former because of its potential to quickly change the distribution of large amounts of material in a relatively short period of time. The combustion of fuels such as coal and oil along with the use of fire to clear large forested regions are two notable and important examples of human activities responsible for the conversions of solids to gases and the movement of large amounts of materials into the atmosphere. We will return to these topics in detail in Chapters 10, 12, and 14.

Limits of tolerance

How much of a substance is present—often referred to as its dose or concentration—will determine if that substance is beneficial or detrimental. Put another way, there is an optimal level for all materials, including chemical elements. Too little *and* too much of an essential nutrient can harm a system. If insufficient amounts of a vital material are available to organisms, they will not be able to properly grow and reproduce. On the other hand, an excess of that same material can bring about drastic and unwelcome changes. We will see what can happen if levels of tolerance are exceeded shortly.

Rate of movement

The time required to move and recycle different chemical elements varies. Because gaseous forms of elements can travel quickly among reservoirs, elements that do not spend much time in the atmosphere tend to move more slowly than those that do. Also, biological processes generally are faster than non-biological ones. The elements we will discuss here have dramatically different residence times in different reservoirs, ranging from days (some atoms in the atmosphere) to hundreds of years (some atoms in the hydrosphere) to hundreds of millions of years (some atoms in the lithosphere).

Three important examples: carbon, nitrogen, and phosphorus

Of the many known elements, we will focus on the movements of just carbon, nitrogen, and phosphorus because of the important roles they play and the useful examples they provide.

The carbon (C) cycle.

We need carbon

Why is carbon on our short list? The first part of the answer is its value to the members of the biosphere. Organisms must have access to sufficient biologically available C atoms or they will not be able to construct the organic compounds that make up their cells and tissues. Carbon is simply crucial for the survival of all living things. Carbon-containing compounds are also important sources of energy for the vast majority of organisms. Secondly, it is central to many of the environmental concerns we will explore in upcoming chapters, including ecosystems, agriculture, energy resources, waste management, and air pollution (Chapters 5, 9, 10, 13, and 14, respectively).

Carbon is versatile

Its chemical properties allow C atoms to bind in many different ways to a wide range of atoms. As a consequence, we see tremendous diversity within the biosphere, with organisms taking on a variety of physical forms. But carbon does not stop there—it is central to the structure of many non-living things as well, from natural objects such as diamonds, coal, and oil to synthetic polymers, building materials, drugs, industrial chemicals, and pesticides. We could devote several textbooks to the wonders of carbon. Instead, though, we will focus on three important compounds in the carbon cycle.

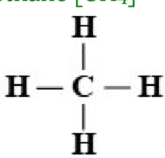
Carbon dioxide [CO_2]

C = O = C

- Chemistry. It is inorganic.
- Phase. It is a gas under most conditions on Earth (but solid dry ice when frozen).
- Sources. It is released as a product of aerobic respiration (e.g., each time animals like us exhale), from volcanoes, and as a product of the combustion of organic carbon compounds (e.g., wood, oil, coal). It can also be the result of chemical transformation of methane (see the next compound on this list).
- Biological availability. This form cannot be utilized directly to build cells and tissues. Carbon-fixing organisms like plants

and algae, called **producers**, can convert it into glucose, a form that is usable to all organisms (see the third compound on this list). The actions of producers are essential to Earth's **consumers**, including animals and many microorganisms, which cannot fix carbon. The relationship between producers and consumers is explored in more detail in Chapter 5.
- Reservoirs *and fates.* It is stored primarily in the atmosphere and hydrosphere, but some is also present in the lithosphere. As we have just seen, it can also be fixed by producers.
- Important pathways. Decomposition, respiration, and fixation are carried out by organisms and can cause rapid conversions among this and other forms. Combustion of organic carbon compounds found in materials such as wood, coal, oil, trash, and animal waste releases CO_2 into the atmosphere, and diffusion can move it from the atmosphere into the hydrosphere (and back again).

Methane [CH_4]

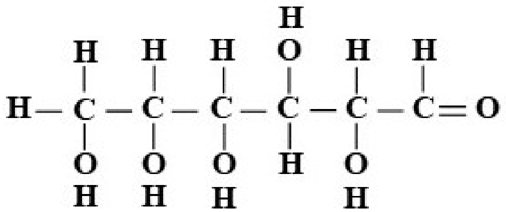


- Chemistry. It is organic.
- Phase. It is a gas under most conditions on Earth.
- Sources. It is released from anaerobic respiration and decomposition (e.g., certain bacteria), from some animals (e.g., cows) and as a product of combustion (as seen with carbon dioxide, above).
- Biological availability. Other than a small number of very specialized bacteria, organisms cannot use this form of C directly.
- Reservoirs and fates. It is stored primarily in the atmosphere and hydrosphere, but some is also present in the lithosphere.

As noted, it can be used as a nutrient source by certain bacteria and some is converted to carbon dioxide.

- Important pathways. Anaerobic respiration (see above) and combustion are the two major pathways affecting the distribution of methane.

Glucose [$C_6H_{12}O_6$]



- Chemistry. It is organic.
- Phase. It is typically solid.
- Source. Carbon-fixing organisms (see above) can convert carbon dioxide into glucose.
- Biological availability. This form of C can be used by all members of the biosphere as a basic starting point for the construction of many cell components such as fats, proteins, and carbohydrates.
- Reservoirs and fates. Carbohydrates like glucose are present in living organisms and the lithosphere. Since it is the preferred nutrient source for many organisms, glucose is converted to carbon dioxide by aerobic organisms or a gas such as methane by anaerobes. In other words, consumption and respiration lead to the release of unavailable, gaseous forms of carbon. Consult Box 4.2 for the story of what happens to carbon compounds that enter organisms.
- Important pathways. Fixation produces glucose, and decomposition and respiration break it down to smaller and less useable compounds like carbon dioxide. Consumption in the biosphere, as well as erosion and combustion, also affect

the distribution of glucose and its precursors and products.

Box 4.2. Just what are we doing with that food we eat?

As we have learned, organisms, including humans, need to find biologically available forms of carbon to survive. To do that, of course, we eat food. But just what are we doing with glucose after we obtain it? Three fates are possible. First, when the bonds holding the molecule's atoms together are broken, stored energy is released. Some of that energy is absorbed and used immediately, or later, to do work. The carbons that were in the glucose are released to the atmosphere as a product of our respiration. Second, some of the C atoms stay inside our bodies and are used as building blocks to construct new cells and tissues. Finally, we release some of the compounds in our waste. Those unwanted products still have stored energy and usable carbon in them, even if we can no longer use them. Keep in mind that *some*body can use them, though: certain organisms live off excrement, extracting needed energy and other nutrients from it.

Carbon fixation is a crucial step in the cycle

There are a few mechanisms by which carbon dioxide is fixed into glucose, although **photosynthesis** is the most common (and likely the most familiar).

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$$

Plants, algae, and other organisms can use the sun's energy to produce one glucose from six carbon dioxide and six water molecules. This fixed carbon is then used by both producers and consumers in the biosphere. You should also notice the six oxygen molecules, O_2, released during carbon fixation. That second compound is of no small importance to all of us aerobes on Earth: without photosynthesis, the atmosphere would not contain breathable oxygen (more in Chapter 1).

Carbon moves among reservoirs via pathways

As with all the materials we will consider, C atoms are distributed among Earth's spheres as they travel along the pathways described above. Just as important as their location is their chemical form: remember that carbon dioxide may be abundant, but success of organisms depends on how much of it is converted to glucose. Figure 4.5 connects the reservoirs and pathways of the carbon cycle.

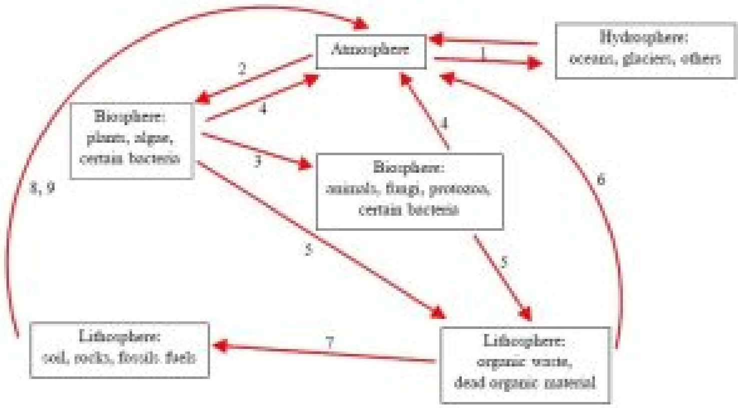


***Figure* 4.5.** *The carbon cycle. The boxes are reservoirs, and the numbered arrows are pathways.*
1. Diffusion moves CO_2 *between the atmosphere and hydrosphere. Note there is no change in chemistry of the C, just a change in its location.*
2. Carbon fixation by producers converts CO_2 *in the atmosphere (or, in the case of aquatic systems, hydrosphere) to glucose. This moves C into the biosphere.*
3. Fixed carbon moves within the biosphere from producers to consumers.
4. Respiration by all organisms moves CO_2 *(aerobes) and* CH_4 *(anaerobes) back to the atmosphere.*
5. Organic carbon in the waste and remains of living things can accumulate in the lithosphere.
6. Decomposition and respiration by fungi and certain bacteria convert some of the organic material in the lithosphere into CO_2 *(aerobes) and* CH_4 *(anaerobes). The gases can move back to the atmosphere.*
7. Some fraction of the organic material in the lithosphere escapes decomposition and is stored for an extended period in soils, rocks, and fossil fuels (more on fossil fuels can be found in Chapter 10).
8. Combustion of fossil fuels converts organic C in the lithosphere into CO_2 *and other gases.*
9. Slow decomposition of some of the organic matter in the lithosphere converts organic C into CO_2 *(aerobes) and* CH_4 *(anaerobes). Kelsey, CC BY-NC-SA.*

The nitrogen (N) cycle.

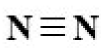
We need nitrogen

Like carbon, nitrogen is essential to members of the biosphere: without biologically available forms of N, organisms cannot make proteins. Although less nitrogen is required than carbon, it is still vital to our survival. We will also see that nitrogen-containing compounds play important roles in our study of agriculture (Chapter 9), water pollution (Chapter 13), and air pollution (Chapter 14).

There are many forms of nitrogen

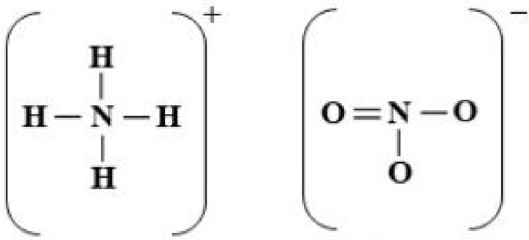
Nitrogen is a versatile element and can be found in different types of materials. For the sake of efficiency and simplicity, we will consider just four compounds here.

Dinitrogen [N_2]



- Chemistry. It is inorganic (no C atoms are present).
- Phase. It is a gas under most conditions on Earth.
- Sources. Recall from the beginning of this chapter that N_2 makes up about 78% of the gases in the atmosphere. It became so abundant during Earth's long history due to several geologic processes. On today's Earth, decomposition and combustion of organic material release a variety of N gases that can be converted to N_2 in the atmosphere.
- Biological availability. This form of nitrogen cannot be used to build any biological structures. A specialized group of bacteria, collectively known as **nitrogen fixers**, can absorb dinitrogen and convert it into forms that are directly usable by microorganisms and plants.
- Reservoirs and fates. This form is found primarily in Earth's atmosphere, although a little is present in water and soil. As noted above, it can be fixed under certain conditions. Nitrogen fixation is described in more detail below.
- Important pathways. Fixation, decomposition, respiration, and combustion all participate in the movement of N_2.

Ammonium [NH_4^+] and Nitrate [NO_3^-]

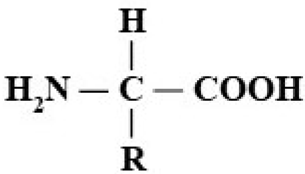


(ammonium) (nitrate)

- Chemistry. Both are inorganic.
- Phase. These ions usually are dissolved in water or bonded to soil solids. Note that the + indicates a lost electron and positive charge and the – means a gained electron and negative charge (ions are described above).
- Sources. Ammonium is produced from N_2 by **nitrogen-fixing** bacteria, and nitrate is produced from ammonium by other bacteria. The two compounds can also be synthesized artificially and applied to soil in fertilizers. Decomposition of waste products and amino acids (the next N compound on our list) in the remains of organisms will yield them as well.
- Biological availability. Both can be used as a source of N by microorganisms and plants, but not by animals.
- Reservoirs and fates. These compounds are primarily found in soil and water. Microorganisms and plants can absorb and convert them into amino acids, the next N compound on our list. Under different conditions, **denitrifying bacteria** can convert ammonium and nitrate back into nitrogen gases such as N_2.
- Important pathways. Fixation, decomposition, assimilation, excretion, and denitrification all play important roles in the cycling of ammonium and nitrate. Downward movement through soil in flowing water, known as **leaching**, and horizontal movement or **runoff** across the Earth's surface will also carry these and other compounds (more about water

movement is presented a bit later).

Amino acids [$RCH(NH_2)COOH$]



- Chemistry. These are organic compounds (recall: at least one C bonded to at least one H). Note that the structure shown above is a generalized model of the twenty different amino acids commonly found in the biosphere. The R does not represent an atom of a particular element, rather it is used to indicate a variety of different possible atoms or groups of atoms that can be located in that place—each of the twenty has a different group in the R position.
- Phase. They are generally found in the solid phase.
- Sources. As noted above, the ammonium and nitrate produced by microorganisms are converted into these compounds. Animals can then obtain the amino acids they need through consumption of plants, microorganisms, or other animals
- Biological availability. They are readily usable by all organisms.
- Reservoirs and fates. These are bonded together into long chains which are then folded into proteins. The specific form and function of the resulting proteins depend on both the identity and sequence in which they are linked. Note that proteins can be decomposed into inorganic nitrogen-containing compounds within organisms. These products are released into soils and waters through processes such as urination. Another group of microorganisms may convert them to amino acids (assimilation) or begin the process which ends with the re-formation of N_2 gas (denitrification).
- Important pathways. Assimilation, decomposition, and excretion influence their distribution.

Nitrogen fixation requires specific conditions

The most common form of nitrogen, N_2 gas, is certainly abundant on Earth—we are surrounded by it—but its two atoms are not useable if they remain bonded together. Complicating matters, it is rather difficult to separate them, and only a limited number of organisms can break that triple bond (see the structure, above) and incorporate the resulting N atoms into ammonium and nitrate. The chemical transformations that increase the biological availability of nitrogen are therefore both relatively rare and essential to the biosphere. Here, we will take just a brief look at the process.

Organisms. About 85% of the nitrogen fixation on Earth is carried out by organisms. Lightening and other phenomena can also lead to the conversion of some nitrogen gas to ammonium, but these forces are relatively insignificant. Certain bacteria have the capacity to carry out the first of many steps in nitrogen fixation. Omitting many of the details, different nitrogen-fixing bacteria carry out this or a similar process under varying conditions. Some do it inside of specialized plants, living cooperatively with the host organism. In these cases, the plant receives useable nitrogen compounds, and the bacteria receive a suitable habitat and fixed carbon compounds (see Chapter 5). Other bacteria fix N while living independently in soil or water.

Environment. In addition to the energy requirement suggested in the reaction above, biological nitrogen fixation requires an environment that lacks O_2 gas because the enzyme that catalyzes the reaction is poisoned by oxygen.

Movement between bioavailable and non-bioavailable forms is key

As with carbon, both chemical form and location of atoms are relevant to the nitrogen cycle. It is fair to say that environmental scientists pay a great day of attention to those forces affecting biological availability because of how limited fixed nitrogen can be in certain systems. Speaking broadly, then, we can visualize two possible directions or pathways within the nitrogen cycle: one associated with increasing bioavailability and one associated with declining availability (Figure 4.6 provides an overview of the processes and Figure 4.7 provides more details).

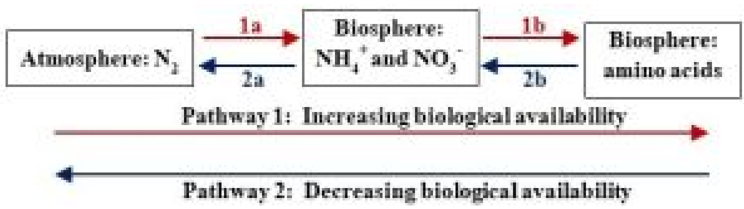


Figure 4.6. *Nitrogen cycle: overview. Pathway 1: a general increase in bioavailability of nitrogen. 1a, fixation by bacteria, 1b, assimilation by plants and microorganisms. Pathway 2: a general decline in bioavailability. 2b, decomposition and excretion by animals, 2a, denitrification by bacteria. Kelsey*, CC BY-NC-SA.

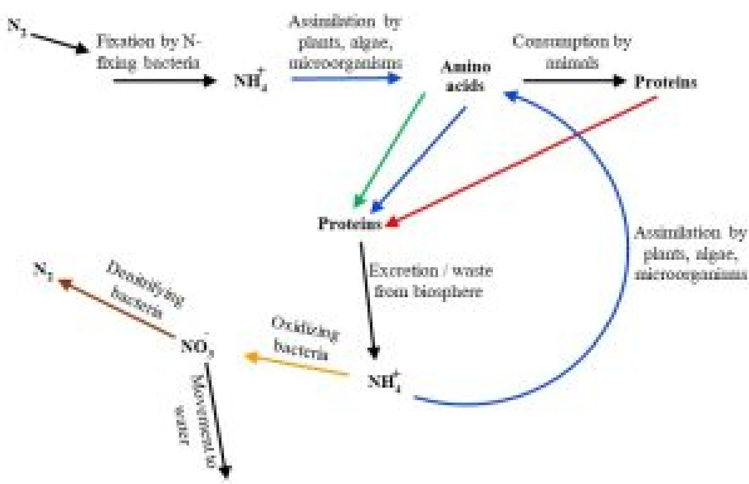


Figure 4.7. *Pathways and reservoirs from Figure 4.6.a: details. Fixation, assimilation, and construction of proteins from amino acids are part of Pathway 1 of 4.6.a (i.e., increasing availability); organic nitrogen excreted in waste can be processed by microorganisms into forms that can move to water and ultimately undergo transformation to biologically unavailable dinitrogen (i.e., Pathway 2 of 4.6, decreasing availability).* Kelsey, CC BY-NC-SA.

The phosphorus (P) cycle.

Phosphorus is essential for organisms

The availability of phosphorus has a profound influence on organisms. If it is lacking, they cannot effectively store energy, synthesize proteins, encode genetic information, and, in animals, construct bones and teeth. It is also important to us as environmental scientists because of its roles in the stories of ecosystems (Chapter 5), agriculture (Chapter 9), and water pollution (Chapter 11).

Phosphorus is part of many chemical compounds

Both organic and inorganic forms of phosphorus are possible, and many of Earth's systems depend on P for proper functioning. For the purposes of this discussion, though, we will simplify matters and only consider whether P is present in a particular system. The effect of chemical form, so critical to the carbon and nitrogen cycles, will not be of concern here.

The distribution of phosphorus depends on slow processes

Much of Earth's phosphorus is stored in the lithosphere and the oceans, and only a miniscule amount is in the atmosphere. Since phosphorus has no relevant gas phase, it moves far more slowly than do carbon and nitrogen. This low rate of recycling is important: because phosphorus is often in short supply relative to demand for it, both natural and human systems are limited by its availability.

Uptake by terrestrial organisms. The word uptake is often used by environmental scientists to refer to the movement of materials into organisms (we will see it in future chapters as well). So, varying mechanisms like ingestion through the mouth of an animal, absorption across the cell membrane of a microorganism (Chapter 3), and even movement into plant roots are typically grouped together.

Excretion and death. As we have seen before, waste products and remains contain important elements. Decomposition releases some of the stored phosphorus, making it available for uptake by a new group of soil dwellers. Some of the phosphorus binds to soil, though, decreasing its availability and increasing the likelihood that it will be lost from the system (see the next item on this list).

Movement with water. Some of the solid materials to which phosphorus has bound can be dislodged by flowing water and carried away. As a result, there is net movement of this important element from land through surface runoff and rivers to the ocean (we will study water movement later in this chapter).

Uptake by aquatic organisms. Certain aquatic microorganisms will absorb phosphorus from water, and consumption will move it into larger members of the biosphere.

Deposition onto land. A portion of the phosphorus in the ocean moves into seabirds that eat fish and other marine organisms. Since these birds tend to nest in the same coastal regions year after year, a large amount of their excrement, known as **guano**, accumulates in selected areas. Guano is rich in phosphorus and other nutrients, so it is part of an important, albeit quirky, pathway for this element. Countries in the western part of South America, for example Peru, once generated a lot of income by mining guano from shockingly large piles and then exporting it (Figure 4.8). Currently, this valuable resource tends to be used within the countries that have it to support their local agriculture.

Figure 4.8. *Guano mining in Peru in 1800s. For scale, the red circle at the lower right of the photo highlights an image of a person. American Museum of Natural History, Public Domain (modified by Kelsey).*

Accumulation on the ocean floor. Phosphorous in the waste and remains of aquatic organisms is generally bound to solids, rather than dissolving in water, and sinks to the bottom of the ocean. Once incorporated into sediments, the nutrient can stay in place for thousands or even millions of years.

Uplift. Although it is nearly impossible to see in the short run, the surface of the Earth is constantly in motion: rocks, sediments, and other materials on the bottom of the ocean can rise up to become dry coastlines and even tall mountains millions of years later (Chapter 3). Given enough time, then, phosphorus on the seafloor will move back to terrestrial areas and be taken up by soil-based organisms again. The processes responsible for uplift are extremely slow, with rates of input of phosphorus into land-based systems lagging far behind rates of output of the element from them.

Upwelling. Some phosphorus in ocean-floor sediments is moved into surface waters due to interactions among wind, Earth's rotation, and ocean currents that lift deep water (Figure 4.9). Due to the abundance of nutrients in them, coastal areas subject to upwelling are typically teeming with marine life.

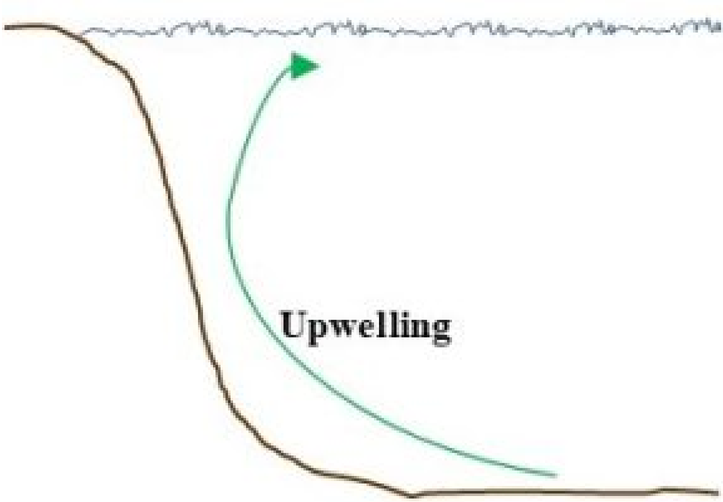


Figure 4.9. *Coastal waters, with upwelling moving nutrients from the ocean floor to the ocean's surface. Not to scale, in cross section. Kelsey,*

Phosphorus moves along pathways among reservoirs

The phosphorus cycle lacks a gas phase, so, as described above, recycling of this vital nutrient can be very slow. Figure 4.10 is a simplified model of the phosphorus cycle. See its legend for more details.

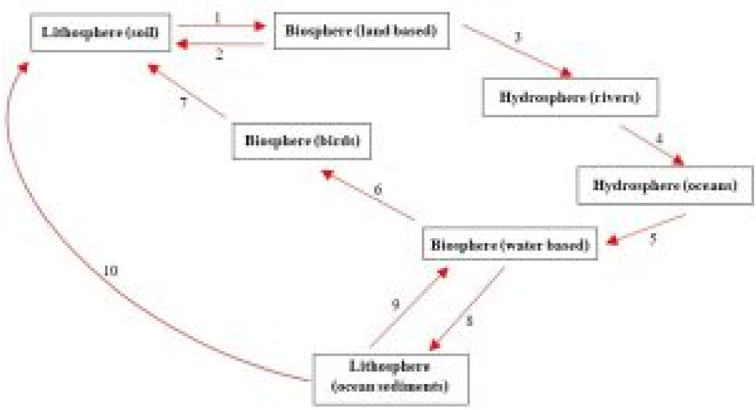


Figure 4.10. *A simplified model of the phosphorus cycle. Boxes are reservoirs, and the numbered arrows are pathways. Kelsey, CC BY-NC-SA.*
1. Phosphorus in soil is taken up by plants and microorganisms and then moves to animals through consumption.
2. Waste products and remains of organisms return some phosphorus to soil.
3. Some phosphorus is carried away from soil in flowing water. Much of the water ends up in rivers.
4. Rivers carry phosphorus to oceans.
5. Marine microorganisms take up phosphorus from water and pass it to the animals that consume them.
6. Consumption moves some phosphorus from fish and other marine organisms to seabirds.
7. Phosphorus is returned to soil in coastal regions via bird excrement.
8. Much of the phosphorus is not passed to birds but settles to ocean sediments with waste and remains.
9. Upwelling can bring phosphorus into surface waters, making it available for marine organisms.
10. Uplift can slowly push phosphorus in rocks and sediments on the ocean floor up onto land.

Elements must be present in the right amount

Recall the principle described near the beginning of the section on biogeochemical cycling: both too much and too little of a necessary nutrient can be detrimental to systems. Organisms will neither grow nor reproduce effectively if the supply of an element such as phosphorus cannot meet demand. On the other hand, excess phosphorus (in a form called phosphate) is problematic because it can lead to **eutrophication**, a dramatic and typically undesirable phenomenon affecting aquatic systems. Note that nitrogen and phosphorus can both bring about eutrophication, so we will consider them together here.

Sources of excess nitrogen and phosphorus. Many processes add P and N to aquatic systems. We learned earlier that these two elements can enter water via natural pathways involving organisms, leaching, and runoff. Human activity can release them as well. For example, farmers often apply phosphorus and nitrogen to soil in fertilizers when natural quantities of nutrients are too low to support crop growth (Chapter 9). These and other synthetic chemical compounds not taken up by plants, bound to soil, or otherwise removed, are transported in flowing water to rivers, lakes, ponds, and oceans. Detergents and some industrial products released into waterways can also be an important source of what is sometimes termed **nutrient pollution** or **cultural pollution** (note that detergents, a big problem historically, have become a less important P source in recent years due to changes in laws regarding their formulation).

Input-output analysis. Recall what happens when a substance is added to a system faster than it is removed from it: the substance accumulates (Chapter 2). Eutrophication occurs when the rate of input of phosphorus and nitrogen to a small aquatic system (e.g., a pond) exceeds the rate of their output from it. It can also be a problem in larger systems such as rivers and even coastal oceans if very large amounts of nutrients are involved.

Steps in eutrophication. When nutrients accumulate, an important series of events is set in motion that can transform a clear pond into a murky swamp or even dry land. In systems terms, an imbalance in inputs and outputs leads to positive feedback and an ever-increasing rate of change. See Figure 4.11, below, for more information.

1. Nitrogen and phosphorus from natural and human sources enter a pond faster than they are removed.

2. The excess nutrients stimulate growth of algae and bacteria.
3. Algae grow so rapidly and that they form a floating mat on the surface of the pond (Figure 4.12, below). Since little sunlight can penetrate the water, aquatic plants and other carbon-fixing organisms can no longer carry out photosynthesis (C cycle, above). Thus, very little new O_2 is added to the pond from within. The mat also restricts movement of O_2 from the atmosphere into the water. Oxygen content in the water drops precipitously as aerobic organisms use it up faster than it is replaced.
4. When oxygen levels drop low enough, aerobic organisms that dominated the pond prior to the addition of excess nutrients begin to die (such as fish—review Figure 1.4).
5. The remains of the dead fish from step 4 are decomposed by certain microorganisms. At first, this process is carried out aerobically. Eventually, though, anaerobic organisms take over when O_2 in the water disappears. Some of the products of decomposition, including organic C, N, and P compounds, start to accumulate. The availability of more and more nutrients further stimulates decomposers.
6. High rates of die-off of aerobic organisms lead to accumulation of their remains because decomposition cannot keep pace. The water gets increasingly murky, cloudy, impenetrable to sunlight, and inhospitable to all but anaerobic microorganisms.
7. The pond gets filled with organic remains and other debris, turning it into a swamp.
8. If allowed to continue, the process will eventually dry out the pond, and a terrestrial system will be the result. Land-based organisms like trees, grasses, and animals will dominate the area.

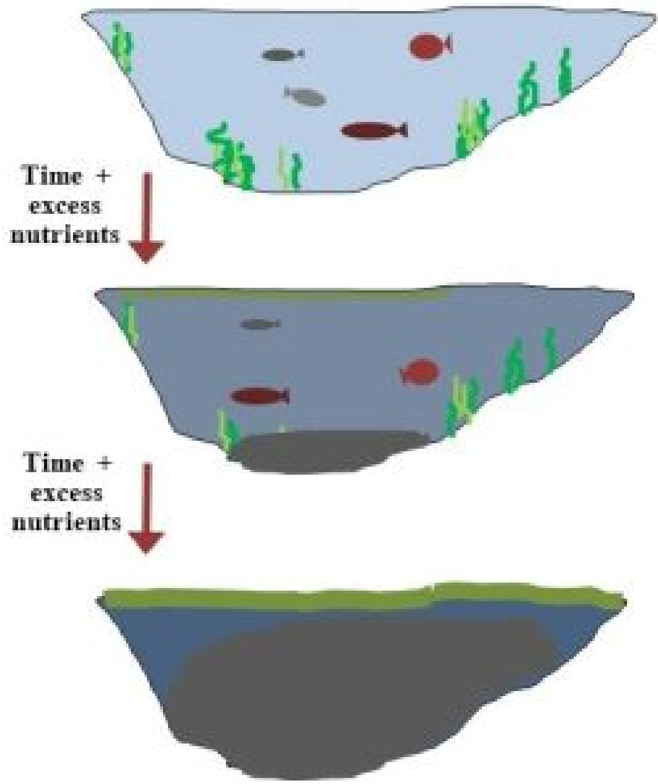


Figure 4.11. *Eutrophication: a clear pond (top) can become somewhat murky (middle) and then a swamp (bottom). The process could continue until the area is transformed into a terrestrial ecosystem such as a forest.* Not to scale, in cross section. See main text for details. Kelsey, CC BY-NC-SA.

Figure 4.12. *Eutrophication of a pond. Note the green layer of algae covering much of the surface, blocking sunlight from entering the water. Trlabarge,* CC BY-SA.

More about feedback and eutrophication. In eutrophication, accumulation of nutrients stimulates the growth of microorganisms and the changes they cause. As less and less oxygen is present in the water, more and more aerobic organisms die; the nutrients in the remains drive more growth of microorganisms and even greater demand for oxygen. Thus, *eutrophication is an example of positive feedback*: the transition from pond to land accelerates with time (Chapter 2).

4.2.3. Cycling of water

As with other materials, the amount of water on Earth is finite.

Since all living and many non-living systems require readily available water to function properly, it must be cycled and recycled rapidly enough to meet demand. The many reservoirs and pathways responsible for the storage and movement of water are linked to make up the **hydrologic cycle**, one of the most dynamic and important systems we will see (a summary is presented in Figure 4.13, with details to follow).

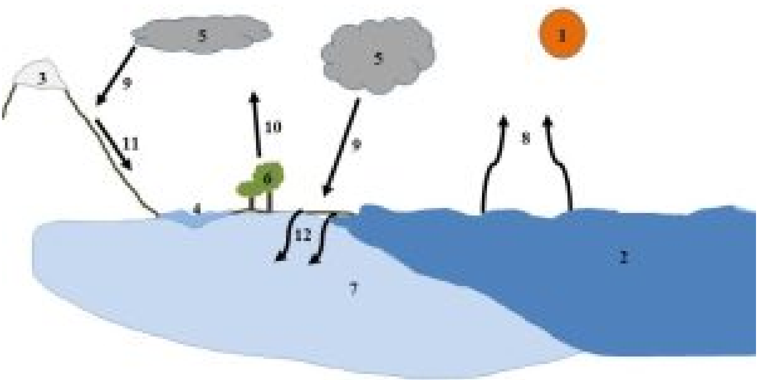


Figure 4.13. *An idealized overview of the hydrologic cycle. In cross section, not to scale. As always, arrows represent pathways that move water among the reservoirs shown. See the main text for details.*
1. The sun, the source of energy for the hydrologic cycle.
2. Oceans, the largest reservoir of water on Earth.
3. Ice, the second largest water reservoir (affected by freezing and thawing, not pictured).
4. River, one example of a freshwater reservoir on the surface.
5. Atmosphere, a reservoir in which water can be found as a gas, liquid, or solid (shown as clouds in this diagram).
6. Trees, one example of a living reservoir of water.
7. Groundwater, a reservoir beneath the surface.
8. Evaporation is a pathway that moves water from the surface to the atmosphere.
9. Condensation followed by precipitation moves water from the atmosphere to the surface.
10. Transpiration moves water from soil to plants to the atmosphere.
11. Runoff, water flowing along the contours of the surface, is one of the fates of precipitation and melt water.
12. Infiltration, water percolating into the ground, is another fate of precipitation and melt water. Kelsey,

Water is stored in natural reservoirs

There is a lot of water on this planet, around 326 million trillion gallons of it. However, water is by no means evenly distributed, and importantly to we humans, almost none of it is directly available for drinking.

Several important reservoirs store water. Here we look at the important characteristics of each, keeping in mind that accessibility and usability are crucial concerns. See Figure 4.14 to get a sense of the relative amounts of water in Earth's major reservoirs (details about the reservoirs are presented below the figure)[2].

2. Volumes of water according to The US Dept. of The Interior (DOI.gov) and U.S. Geological Survey (usgs.gov)

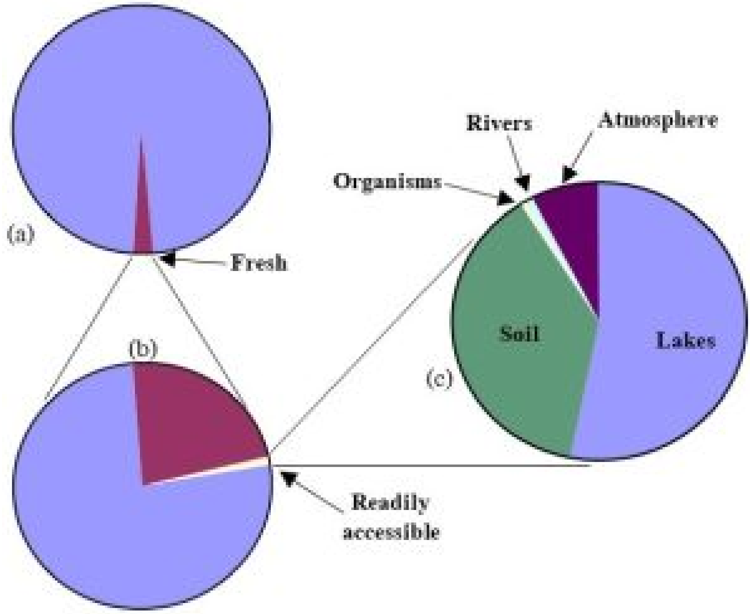


Figure 4.14. *Relative amounts of water on Earth: a, the total amount of fresh vs. salt, b, how that small amount of fresh water from a is distributed, c, how that small amount of readily available fresh water from b is distributed.* *Kelsey*, CC BY-NC-SA.

Oceans and seas: 97% of Earth's water (this and all numbers in this section are rounded; thus, they do not add up to 100%). Nearly all of Earth's water is stored in salt-water reservoirs. Although the oceans are home to a large and diverse number of organisms, and ocean resources are valued and indispensable, the water itself is not in a form that is consumable by land-based organisms. In other words, almost none of the water on Earth is available for drinking. Sometimes **desalination** is employed to remove salt and produce fresh water, but that process requires the proper technology, as well as a sizeable financial investment, to carry out on a large scale (more in Chapter 11). Note there is also a small amount of saltwater stored in certain lakes as well as underground (see more about these two types of reservoirs, below), but this combined volume adds a trivial amount to that in the oceans.

Ice: 2%. The water frozen in glaciers and other places on Earth makes up the second largest reservoir. Like that found in oceans, this water is not readily accessible or available.

Groundwater: 0.6% (but 98% of Earth's fresh, unfrozen water). Since it is not generally as well understood relative to surface waters, and because of its importance as a resource, we will give groundwater some additional attention here.

Rarely seen

This is a reservoir found beneath the surface of the Earth. It can be located as shallow as a meter or less, or as deep as a few hundred meters or more, but in all cases, access to this water is limited by our ability to locate it and pump it to the surface (see Chapter 11 for limitations on groundwater supply).

Rarely in underground pools

Contrary to what is believed by many, the bulk of groundwater is *not* held in underground lakes and rivers. Yes, some water beneath the surface is stored in open, pool-like structures (notably visible at commercial attractions that allow visitors to explore caverns and the rivers that pass through them), but this is not the typical environment in which groundwater is found. Instead, groundwater tends to occupy the small spaces within rocks, sands, and other porous materials. In a situation that more closely resembles what occurs in a drip coffee maker than that in a channel of a rushing river, water moves downhill, in response to gravity, very slowly through these connected openings (See Figure 4.15).

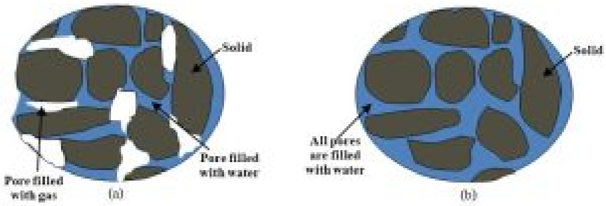


Figure 4.15. *Idealized microscopic view of water in the unsaturated zone, a, and water in the saturated zone, b.* Kelsey, CC BY-NC-SA.

When a layer of comparatively non-porous rocks or clays is present beneath the porous material, water is stopped from further vertical movement. The pores above such an **impermeable boundary** will then fill from the bottom up. If this accumulation occurs to an extent sufficient to concentrate useable quantities of water, the structure is referred to as an **aquifer** (more in Box 4.3).

Box 4.3. Next laboratory field trip: the beach!

If you have ever put beach sand into a plastic bucket you have played with a pretty good model of an aquifer. Water you add to the top of the permeable sand will soak down until it encounters the impermeable bottom of the bucket. The top of the sand may appear to be relatively dry, although you know a fair amount of water must be present below the surface because you put it in yourself.

The type of aquifer present in an area depends on several variables, including the arrangement of permeable and impermeable layers. We should make two additional notes about groundwater here.

First, not all water under the surface of the Earth is considered part of this reservoir. There is an upper region, known as the **unsaturated zone**, in which the pores described above contain both water and air. Water here, termed **soil water**, behaves differently than groundwater and cannot be extracted through wells (soil water *is* used by soil organisms living in the unsaturated zone). At sufficient depth, in the **saturated zone**, the air is gone, and pores are filled exclusively with water. Water in this lower zone is classified as groundwater. The boundary between these two zones, which can also be thought of as the top of the saturated zone, is referred to as the **water table** (Figure 4.16).

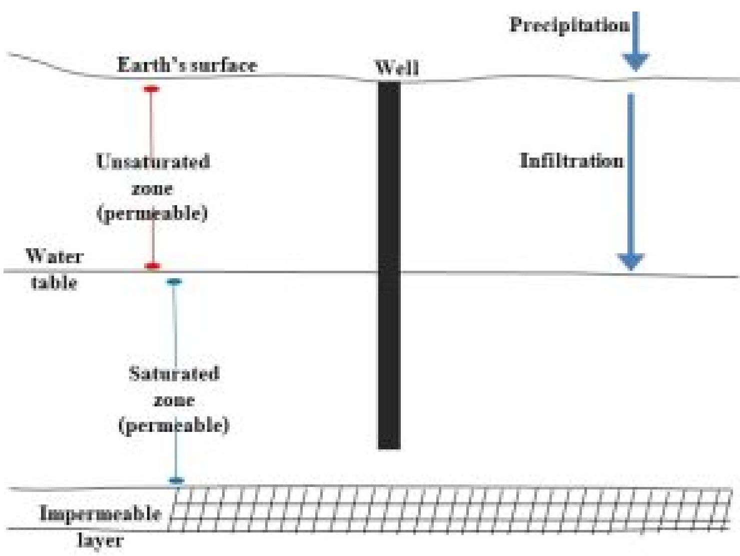


Figure 4.16. *Idealized diagram of the groundwater environment, with a human-constructed well. In cross section. Kelsey,* CC BY-NC-SA.

Water flows vertically downward through the unsaturated zone to the saturated zone via infiltration (more below). This movement is critical because it replenishes water that has been removed from the groundwater reservoir (a process known as **groundwater recharge;** we will see more about the importance of recharge in Chapter 12). Second, in nearly all cases, water must be pumped upward—which can be a great distance—through a well. Very rarely does water rise to the surface naturally. In other words, humans need to invest money and energy if they are to use groundwater. Aquifers are very important to environmental scientists because they often are sources of a substantial amount of water. If they are not managed carefully, demand for water may go unmet. See Chapter 11 for a description of the way human activities can reduce the availability of groundwater.

Fresh lakes, rivers, swamps: 0.01%. These reservoirs are familiar to most people, yet collectively they hold only about 1/100th the amount of freshwater stored below ground. Despite their relatively insignificant volume compared to groundwater, they account for about three times the amount of freshwater withdrawals by humans each day (according to the United States Geological Survey, 230 billion gallons and 76 billion gallons, from lakes and groundwater, respectively[3]). As we will see in Chapter 11, surface water is more susceptible to pollution than is groundwater, so it has its limitations as a resource.

Atmosphere: 0.001%. The bulk of the water in this reservoir is in a gaseous (vapor) form, although there is some liquid here as well. Water in the atmosphere plays many important roles on Earth. We will return to it in the section on precipitation, below, and again in Chapter 14, during our discussion of air pollution.

Soil moisture: 0.001%. A small fraction of underground water is not part of the groundwater reservoir but instead is held much closer to the surface. As we noted above, this reservoir supplies

3. The US Dept. of The Interior (DOI.gov) and U.S. Geological Survey (usgs.gov)

water to plants and other organisms living in the soil (more about soil can be found in Chapter 9).

Organisms: 0.0001%. Individual plants, animals, and microscopic organisms are mostly water, and they cannot survive without a constant supply of it. However, as a percentage of the total amount on Earth, very little water is held in living things.

Water moves along natural pathways

The reservoirs of the hydrologic cycle are neither isolated nor static. Keeping in mind that the net amount of water in each does not change much with time, multiple pathways are responsible for the near-constant movement of water among Earth's systems[4]. Review Figure 4.13 to see how pathways connect reservoirs.

Evaporation. When heated, water is converted from a liquid to a gas. The sun heats the oceans and other surface waters, so it powers evaporation and ultimately the whole hydrologic cycle. Evaporation accounts for nearly all the movement of water from Earth's surface to its atmosphere, approximately 90% of it, and all but about 10% of *that* water returns directly to the ocean through the process known as precipitation (below). The rest ends up falling onto land. The rates of water input and output to and from the surface are roughly equal on a global basis, although there are small-scale systems that experience unequal rates. For example, evaporation generally outpaces precipitation over dry land but occurs at relatively lower rates over bodies of water. In addition to the way it affects the global distribution of water, evaporation can also have a deleterious effect on a small scale (see Box 4.4).

4. Relative importance of pathways according to USGS (usgs.gov).

Box 4.4. Dams do not always give you what you expect

Humans often construct reservoirs by diverting or damming rivers. At first glance, this might seem to be a reasonable strategy because water can be artificially stored in areas of high demand rather than being allowed to move downstream. Unfortunately, the transformation of a stream channel into a stagnant, relatively flat lake can dramatically increase the rate of evaporation from an area and provide far less water than hoped (Figure 4.17).

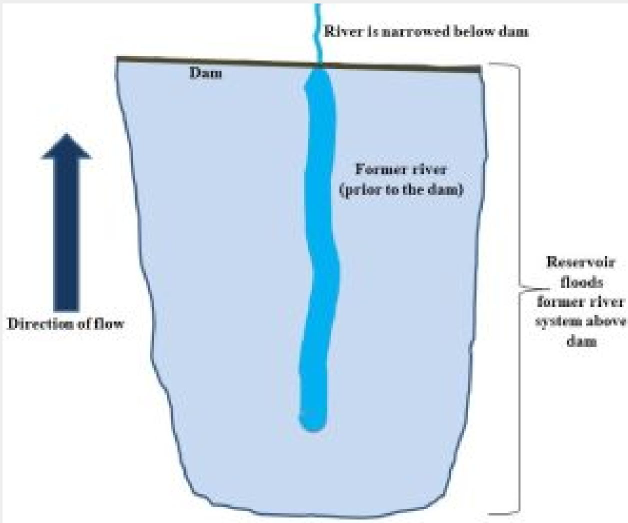


Figure 4.17. *The surface area covered by water increases after the river is dammed. Rate of evaporation from the reservoir is much higher than that from the river; see main text for more. Map view. Kelsey,* CC BY-NC-SA.

Evaporative loss is the term used to describe what can happen to a substantial fraction of the water trapped behind a dam. Lake Mead, the reservoir created by the Hoover Dam on the Colorado River is just one example of an area affected by this phenomenon. Estimates suggest that approximately 0.75 cubic kilometer (nearly 2 meters of vertical height) of water is lost each year from the lake[5]. Other potential consequences of dam building include decreased flow of water and nutrients to areas downstream, a change in the character of the system that was flooded to create the reservoir, increased risk of flooding if a dam breaks, increased rates of erosion of the stream bed downstream, accumulation of sediments upstream, and interruption of movement of migratory aquatic species such as salmon.

Transpiration. This process is responsible for about 10% of the movement of water from the surface to the atmosphere. In short, water and nutrients enter plant roots, and then nearly all that liquid water moves up and exits as a gas from microscopic openings in leaves. Since it is difficult to separately measure biological and non-biological pathways of evaporation, we often use the term **evapotranspiration** to refer to all the processes involved in the movement of water from the land to the atmosphere.

Condensation. Condensation is the reverse of evaporation: when water vapor is cooled, it can be converted to a liquid. This phenomenon is responsible for the dew that coats plants and outdoor structures on a cool morning. In the atmosphere, water will condense and combine with very small circulating dust particles to form clouds. At first these wet particles remain aloft, but eventually enough water can be added to them that they become too heavy to stay in the atmosphere. The result is precipitation, the next pathway on our list.

5. Lake Mead's Water Budget, National Parks Service. 2023. Public domain.

Precipitation. Water that moves toward Earth's surface due to the force of gravity is known as precipitation. It includes solid forms like ice and snow as well as liquid rain. Precipitation is critical to continuing the hydrologic cycle because it is the principle mechanism by which water returns to the ground from the atmosphere. It supplies organisms with the moisture they need to survive, replenishes the reservoirs described above, and is an important agent in the cycling and recycling of Earth's crust and rocks (Chapter 3).

Runoff and Infiltration. Rain that hits the ground experiences two important fates: it can flow across the surface, following contours of mountains, hills, and valleys, or it can move downward, penetrating the ground and percolating slowly through soils, rocks, and other materials making up the Earth's crust. The first of these fates is termed **runoff**. Water that encounters a material that is not immediately permeable will move from high to low elevation along the surface. For example, water will flow over rocks, hard soil, pavement, buildings and other human structures. This pathway is very important because it replenishes surface reservoirs such as streams and lakes. It also can lead to flooding during rainfall events. Note that for the purposes of this discussion, runoff is classified as a pathway because it is temporary phenomenon that comes after a precipitation event; rivers and streams are grouped with reservoirs because although they clearly move water, they also hold water on a long-term basis. **Infiltration** is the second fate of water falling to Earth's surface. Materials such as soil and gravel are permeable, so water will move vertically downward through them in response to gravity. Infiltration is an important pathway because it replenishes water that is removed from underground reservoirs by both human and non-human activities. The balance between infiltration and runoff is of concern to environmental scientists because it has a substantial effect on the likelihood that a rain event will cause flooding along nearby water ways. In areas having relatively impermeable surface materials, runoff will outpace infiltration. Therefore, water will move more efficiently downhill to streams than in areas in which water is diverted as it percolates into and through soils and rocks. We will see more about flooding in Chapter 7.

Freezing and Thawing. Water can be locked up in glaciers or other frozen reservoirs when it freezes. Thawing will release water, subjecting it to evaporation, infiltration, and runoff. Some high-latitude regions on Earth depend on spring thawing to replenish depleted reservoirs.

Consumption and Excretion. As we know, organisms are by far the smallest reservoir of Earth's water. These pathways are similarly trivial. However, water does enter and exit organisms, and biological usage of water can have an impact on a local scale. Environmental scientists are particularly interested in the ways human activity can pollute and limit the supply of freshwater, a subject we will explore in Chapter 11.

4.2.4. Rock and tectonic cycles

The movement and recycling of the materials making up the lithosphere are also important to a study of environmental science. Please refer to Chapter 3 for a review of those concepts.

4.3. MOVEMENT OF ENERGY

There is a good reason the bulk of this chapter has been devoted to the cycling of materials: their movements, transformations, and recycling are necessary to maintain Earth's closed systems. We would be remiss, however, if we did not consider the energy that powers the many pathways and processes involved in the constant renewal of nutrients and other resources. Therefore, here we briefly explore the sources and behavior of energy, including how the rules governing it are quite different that those that apply to materials.

4.3.1. What is energy?

Most people have at least a basic understanding that energy is related to one's ability to accomplish tasks. Casual comments such as "I could not possibly run anymore—I'm out of energy" speak to the importance we assign to it: energy is clearly required for life. Environmental scientists use a similar, if more formal, definition: **energy is the capacity to do work.** What is work? It can be thought of as the product of an effort to change the status or position of an object. So, when you carry a heavy box upstairs you have used energy to do work—you have moved the box. Two important types of energy are briefly described below.

Potential

In short, this is stored energy; it *could* provide the capacity to do work, but currently is not used to do anything. If you imagine what it would be like to hold a large boulder in place at the top of a steep slope, you have some idea about the nature of potential energy. As long as you keep your hands on it, the boulder will not move—it is ready to roll, though. Another familiar example of potential energy can be seen when you lift a delicate glass up over your head: the higher you hold it, the more energy is stored (by virtue of its ever-higher position). Chemical compounds can also be viewed as stored energy. In certain circumstance, reactions involving them release energy and propel some important changes, such as the production of a protein from the bonding together of individual amino acids (as in the nitrogen cycle, above).

Kinetic

This is defined as energy in motion. In other words, energy that has been stored for some period of time is now mobilized to do work. The boulder imagined above becomes a manifestation of kinetic energy because, once you let it go, its potential energy drives downhill movement.

4.3.2. Energy inputs to Earth's systems

A lot of work is carried out on Earth. From the movement of organisms to the evaporation of water to the fixation of nitrogen to the movement of tectonic plates (Chapter 3), demand for energy is high. Moreover, since energy is neither recycled nor reused to do the same work again (*unlike* materials), new energy must be supplied constantly. Essentially all of Earth's energy demand is met by three sources.

Solar energy

This is by the most important source of energy to Earth, providing over 99.9% of the energy we receive. It should not be surprising, then, that most systems are powered by the sun, including the hydrologic cycle, all but a small fraction of the biosphere (more in Box 4.5), weather events, wind, and other phenomena in the troposphere. We will see how humans make direct use of the sun for energy in Chapter 10.

Box 4.5. Exceptional bacteria

The biosphere is mostly powered by the sun—nearly all producers use a form of photosynthesis to fix carbon, and then consumers gain bioavailable forms of carbon they need by living off those producers. However, a few specialized bacteria can produce glucose from carbon dioxide without light. They use energy released in chemical reactions to power their work. These organisms are crucial to supporting larger organisms in dark places like the deep ocean (Chapter 5).

Geothermal energy

It accounts for less than 0.1% of the total, but this still is an important source of energy. Recall from Chapter 3 that Earth began as a completely molten body which has been cooling for about 4.6 billion years. Although the outer crust is relatively cool, much evidence suggests Earth's interior, namely the core and parts of the mantle, are still very hot. This interior heat slowly moves toward the surface and drives the movement of **tectonic plates**. Dramatic phenomena such as volcanic eruptions and earthquakes (Chapter 7) are therefore powered by geothermal energy. Humans can tap into this source as well (Chapter 10).

Gravitational energy

This energy is associated with the pull that all objects in space have on each other. The two most important to the Earth are the moon, due to its proximity, and the sun, due to its proximity *and* size. In simple terms, the changing position of the sun and moon leads to the familiar phenomenon of high and low tides (many other factors beyond the scope of this textbook contribute to the cycling of the tides). Although the amount of energy here is a very small percent of the total—even less than that contributed by geothermal—it still drives a process that affects human and non-human systems along coastlines in profound ways.

4.3.3. Energy outputs from Earth

Energy that enters Earth's system experiences two fates: it is used to do work and irreversibly changed in the process, or it is reflected back to space, unchanged, without being used to do any work.

Conversions

Energy is converted from forms that have the highest capacity to do work into forms that have the least capacity to do work. Importantly, energy does **not** disappear from the universe—it is simply transformed. These are such fundamental ideas that scientific laws (Chapter 2) have been developed to describe them. The two most relevant to our discussion are included here.

The first law of thermodynamics. This law states that there is a fixed amount of energy in the universe: energy can be neither created nor destroyed. The energy we saw stored in chemical compounds and then released, for example, is *not* produced from scratch, it is simply freed or made available to do work. Another way to express the first law is to say that the amount of energy in the universe is **conserved**.

It is unfortunate that people often say "energy is created" when they really mean "energy is released". The first law tells us why the distinction between those two statements is important.

The second law of thermodynamics. Energy cannot be destroyed, but it is certainly subject to *change* when it is used to do work. That is, energy undergoes conversions among many different possible types. The second law of thermodynamics describes the nature of these energy conversions. Its essence consists of the following ideas.

Energy is degraded

To understand this concept, we can imagine high-grade energy as concentrated and highly ordered. It becomes less concentrated and less ordered any time it is converted to another form (e.g., after being used to do work). Eventually, the energy is completely dissipated. Keep in mind that all reactions and conversions lead to an increase in the total amount of disorder, known as **entropy,** in the universe.

Energy loses its ability to do work

Energy conversions occur whenever energy is used to do work. Loss in order, therefore, is linked to lowered ability to do work. Once it becomes dissipated and disordered, energy cannot be used to do the same work again (see Box 4.6 for a familiar example of energy degradation).

Box 4.6. Your car must obey the laws.....of thermodynamics

The next time you are filling up your car's gas tank, you ought to contemplate the second law of thermodynamics. Consider it takes energy to perform the work of moving your vehicle down the street. If it is like nearly every other car on the road, it has an engine that can convert the potential energy in petroleum into the kinetic energy of moving pistons, spinning wheels, and forward motion. Now, you might ask just how is gasoline used to make the car move? In simple terms, when it is burned, chemical reactions lead to the release of energy. Some of that energy is converted into useful work, like turning of the wheels, but far more than half of it is lost from the system as heat. What happens to the heat? The answer can be found if you touch the outside of the hood of a car that has been running for a few minutes: it is warm (the tailpipe and other components are also warm, but you ought to keep your hands off them). This heat is radiated from the car to space and cannot be used to perform work. As is likely familiar to you, after the car has been driven for several hundred kilometers, it will require more gasoline. At this point your choices are to find a station to refill the tank, that is, provide new high-grade energy and enable additional work, or allow your car to run out of gas. If you want to continue your journey with the second option, you will need to do the work yourself by walking, biking, etc. Be sure you have enough fuel stored in your body before you embark. Once your tank is full and you have paid, you should think about one final, yet crucial, idea. What is the origin of the potential energy stored in the fuel you just purchased? Present-day reservoirs of materials like oil and natural gas are the result of millions of years of processing of the remains of ancient organisms (more in Chapter 10). Recall from our discussion of the carbon cycle earlier in this chapter that some fraction of dead organic matter escapes decomposition and can be transformed into fossil fuels. So, most of our current technology is powered by extremely old organic chemicals never fully broken down by members of the biosphere. It is fair to say your car runs on *indirect* solar energy!

Earth constantly needs new energy

Once high-grade energy in Earth's system is degraded, it moves out to space in the form of **heat** (low-grade energy) and cannot be used again. Only new high-grade energy can replace lost energy if additional work is to be done. In other words, Earth must remain open with respect to energy. In this context, you should review the straight-arrow model used in Figure 4.2 and note how the tail is *no*t connected to the head: again, unlike materials, energy is not recycled.

No energy conversion is 100% efficient

Imagine a hypothetical example involving two types of energy. The first is classed as high grade relative to the second, but both still have the capacity to do some work. Some fraction of the useable energy present in the first form can be converted into the second form. However, much of the energy in the first form never makes it to the second form. Instead, it is completely degraded to heat during the conversion.

Energyon Earth may be dissipated in steps

There are several possible fates of energy. In all cases, high-grade is converted to low-grade energy. The specific number of reservoirs and pathways involved can vary, however. Some of the conversions involve individual small steps from one relatively high-grade form to a lower-grade form, whereas others change the energy from the highest to lowest in one step. The path that energy can follow as it moves from the sun through the biosphere and out to space is modelled in Figure 4.18.

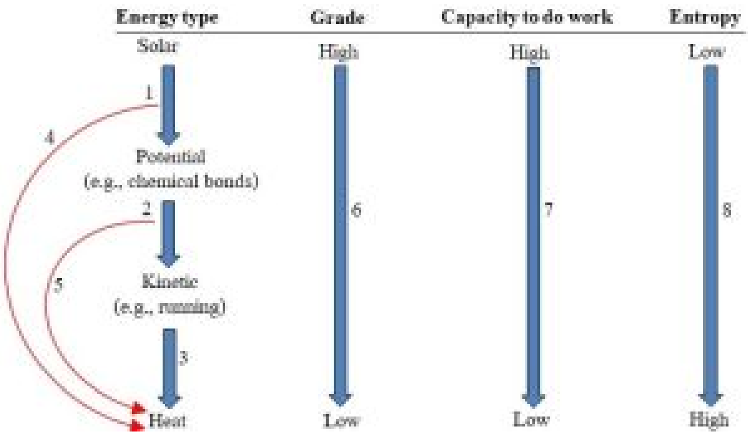


Figure 4.18. *Hypothetical path of energy entering Earth's biosphere*
1. Some of the high-grade solar energy is absorbed by photosynthesizing organisms, say, plants. They convert it to chemical bond energy and store it in their cells. Some of that energy is passed to plant-eating animals and stored in their cells.
2. Some energy stored in animals is converted to work: for example, envision a plant-eater chased by a carnivore.
3. After it is used to do the work of running, the energy is converted to heat and radiated off both animals.
4. No energy conversion is 100%; some of the solar energy is lost to heat during the conversion to chemical energy in step 1.
5. The conversion from potential to kinetic energy shown in step 2 is inefficient. Some of the stored energy is converted to heat and not used for running. This inefficiency would mean the bodies of the running animals would feel warm.
6. There is a progressive and irreversible change from high-to low grade energy with each conversion.
7. The capacity for a particular form of energy to do work decreases with each conversion.
8. Entropy increases as energy is used to do work. Kelsey, CC BY-NC-SA.

Reflection

A portion of the sun's energy striking any planetary body is not absorbed. Just how much is lost this way can be expressed as **albedo**, the fraction of incoming radiation reflected off a surface. Earth has an albedo of about 0.30, meaning that 30% of light hitting it bounces back to space, whereas Mercury and Venus reflect around 7% and 70%, respectively[6]. Albedo matters because it affects Earth's temperature; as more energy is absorbed, the warmer the planet becomes. Several factors can change reflection, either increasing or decreasing it. Perhaps most familiar is the tendency of darker colors to absorb more energy than lighter ones. Ice, snow, and other materials will absorb less than will soil, asphalt roads, and most buildings. Green vegetation and water are intermediate in their reflection. Since Earth's surfaces can change color, the amount of heat absorbed can also change. Consider, for example, what could happen if large-scale transformations of forested land (mostly green leaves) into farms (mostly brown soil) occurred: less energy would be reflected and the temperature of the Earth could increase (Chapter 9). Urbanization would have a similar impact. The melting of ice and snow can also enhance absorption and bring on higher temperatures as more and more dark soil is revealed. We will return to the effect of albedo on temperature in Chapter 14 when we consider human activities influencing Earth's climate. Figure 4.19 is a simplified representation of the sources, pathways, and fates of energy that enters Earth's system from the sun.

6. NASA. Mercury and Venus fact sheets. 2021. Public domain.

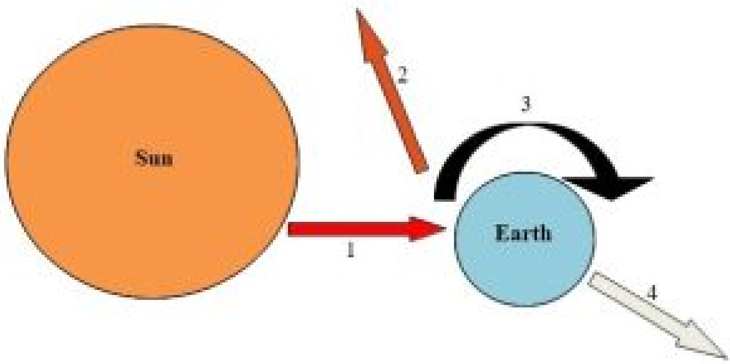


***Figure 4.19.** A simplified diagram of energy flow into and out of the Earth. The numbered arrows are pathways. Not to scale.*
1. High-grade solar radiation strikes the Earth.
2. About 30% of incoming radiation is reflected to space.
3. Energy not reflected is absorbed and used to do work. It undergoes conversions from high- to low-grade forms (see the main text).
4. After it has been degraded, energy is released to space from Earth as heat.
Kelsey, CC BY-NC-SA.

THE CHAPTER ESSENCE IN BRIEF [7]

Earth is closed with respect to materials, thus environmental scientists carefully track the movements, transformations, and storage of materials within and among the atmosphere, hydrosphere, biosphere, and lithosphere. Energy behaves

7. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 4 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

differently than materials, so Earth must remain open with respect to it.

Think about it some more...[8]

Are the four spheres of the Earth separate entities or do they overlap?

Is more always better when it comes to essential nutrients? Think about the nitrogen and phosphorus cycles as you pond your answer.

How are biogeochemical cycles good examples of the fundamental principle that Earth is closed with respect to materials (Chapter 1)?

Why might we care about the rate of evaporation from a constructed water reservoir?

How is most groundwater stored (under natural conditions)?

8. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

What does energy enable?

What is wrong with comment such as “we have run out of energy”? Is there a more appropriate way to put it?

5. Fundamentals of Ecology

JASON KELSEY

As we saw in Chapter 1, the relationship between organisms and their surroundings is of great interest to environmental scientists. Here we will take a closer look at the ways the *living* things in an environment affect each other—that is, we will explore the science of **ecology**. Before we study dynamics involving groups, though, we will briefly consider the needs of individuals, because it is the drive to meet nutritional requirements that ultimately pushes a single organism into interactions with others.

Key concepts

After reading Chapter 5, you should understand the following:

- How different organisms acquire the materials and energy they need to survive
- How organisms can obtain and use energy
- The hierarchy of organization of natural systems, ranging from individuals to ecosystems
- The roles that different organisms play in ecosystems
- The distinction between ecosystem function and structure

- How biomes are defined and why they matter
- Various ways organisms interact with each other to fulfill their needs
- The rise and relevance of antibiotic-resistant bacteria
- Trophic levels and the movement of materials and energy through ecosystems
- The process of ecological succession and factors affecting it

Chapter contents

5.1. WHAT IT TAKES TO SURVIVE

Despite obvious and dramatic differences in size, shape, structure, and lifestyle, every organism on Earth faces the same challenges: each must secure sufficient **energy** and **materials** if it is to succeed. We will soon see, however, that the **biosphere** features many different strategies to obtain necessary nutrients.

5.1.1. Energy

Demand for energy is high

Generally, organisms only carry out work that contributes to their basic survival. They struggle in the face of scarcity just to obtain the nutrients required to function each day, protect themselves, move (as appropriate), and stay alive long enough to produce offspring. The last item on that task list is extremely energy intensive and one of the most critical functions of living things. Put simply, current generations are driven to replace themselves with subsequent generations to ensure the continuation of their species (of course humans are not bound by all the same rules as other organisms, as noted in Box 5.1).

Box 5.1. Humans are not obliged to follow ecologists' rules about energy and reproduction!

1. Energy. Humans must meet basic needs but are the only species that seems to find surplus energy for non-essential activities like video games, jogging, dancing, tennis matches, sex solely for the sake of pleasure, and Sunday drives.

2. Reproduction. Clearly, one can contribute to the continuation of a species without having offspring. But that discussion is beyond the scope of this book.

It is worth noting here that the reproductive strategies used to ensure the success of offspring and the continuation of a species vary. In some cases, for example, numerous fish and plant species, energy is devoted to producing as many individual progeny as possible rather than providing much (or any) parental care once the next generation emerges. On the other hand, organisms like humans concentrate their energy on raising one or a few offspring as a way to increase the likelihood that one's genetic information survives into the future. Consult Box 5.2 for more details.

Box 5.2. How to best use your reproductive energy: overwhelming numbers or good parenting?

Most species allocate their reproductive energy in one of two ways. In the first, parents concentrate their energy on the release of large numbers of offspring—think clown fish, bacteria, frogs, and even dandelions (although the word is a bit clunky in some of our cases, we will assume "parent" is descriptive enough), and then do little or nothing to care for their young. This approach works because, although nearly all the new individuals are likely to die, enough survive to keep the species going. The second strategy involves the production of a few (or one) offspring that are protected and cared for to maximize their chances of survival. For example, relatively long-lived gorillas, elephants, orcas, cheetahs, and humans depend on intensive parenting (and all the energy it requires) to prepare their young before turning them out on their own. Ecologists employ a shorthand when referring to these two options: organisms using the first are known as **r strategists**, and those in the second group are called **K strategists**. What's with the anti-intuitive letters? Suffice it to say that an equation employing several variables, including r and K, can be solved and yield either r- or K- dominant (more commonly called **r- or K-selected**) results. Finally, r and K are endpoints, and most organisms trend toward one of them. Note, though, that the reproductive strategy employed by some organisms does not fit neatly into the r / K classification scheme.

Phototrophs and chemotrophs obtain energy from different sources

Phototrophs. Light is the energy source for these organisms. Specialized structures in their cells allow them to transform solar radiation into chemical energy (note the suffix -troph is from the Greek for feeding—so, literally translated these are “light eaters”). Often, photosynthesis is the process by which these conversions are carried out (Chapter 4): using energy from the sun, atoms from inorganic molecules are linked together into the important products glucose and O_2. The resulting molecules are stored in their cells until energy is needed to do work. Plants, algae, and some bacteria are in this group.

Chemotrophs. These organisms cannot harness solar energy but gain what they need through reactions involving pre-existing chemical compounds. Nearly all chemotrophs get their energy from organic molecules formed through photosynthesis or similar reactions, that is, directly or indirectly from phototrophs. Put into simpler terms, this group includes organisms that are *not* phototrophs: all animals, fungi, protozoa, and many bacteria.

5.1.2. Materials

Living systems are built of atoms

Recall from Chapter 4 that atoms can be bonded together in various ways to produce a dizzying number of different structures, including cells and tissues. Although many elements are required for an individual to be successful, organisms are classified into one of two groups according to their source of carbon (consult Box 5.3 for more about the elements that are most important to members of the biosphere).

Box 5.3. The ingredients for life

It would be fair to say that all living things are largely constructed of the same basic building blocks. As you might imagine, though, the amount of **carbon**, nitrogen, **hydrogen**, **oxygen**, phosphorus, and other elements present in cells varies among organisms. For example, **bacteria**, marine fish, land mammals, grasses, and trees need different relative amounts of carbon and nitrogen, but they all must obtain and concentrate large amounts of these vital elements if they are to be successful. Table 5.1 provides a very brief list of the nutritional requirements of two types of organisms, revealing both similarities and differences among members of the biosphere.

Table 5.1. *Relative amounts of essential elements in bacterial and human cells. The numbers are the*

approximate percentage, by mass, of the total atoms present.

Element	Bacteria[1]	Humans[2]
Oxygen	20	65
Carbon	50	18
Hydrogen	8	9
Nitrogen	14	3
Phosphorus	3	0.5
All others	~5	~4.5

Autotrophs and heterotrophs use different sources of carbon

Autotrophs. When the prefix "auto" is combined with "troph" the resulting word can be taken to mean "self feeder." Obviously these organisms are not able to conjure up carbon out of nowhere, so this literal translation overstates their powers. They do have an ability that most organisms do not possess: autotrophs produce useable carbon molecules from biologically unavailable carbon precursors.

1. Todar, K. 2016. Online Textbook of Bacteriology, textbookofbacteriology.net.
2. Shah, R. 2015. Elements that keep us alive also give color to fireworks. National Institutes of Health. biobeat.nigms.nih.gov

In other words, using **carbon fixation**, they can generate glucose from carbon dioxide they obtain from the air or water in which they live. Plants, algae, and some bacteria are autotrophs.

Heterotrophs. These "other feeders" are not able to fix their own carbon. They can only use organic, biologically available compounds. Heterotrophs therefore consume the glucose produced by autotrophs. Most species are in this group, including animals (like humans, fish, and insects), fungi, protozoa, and most bacteria.

5.1.3. Defining properties: energy and carbon sources

Many other characteristics affect survival and success, but we can say a lot about an organism if we know how it obtains energy and carbon. Using that information, we divide the biosphere into four basic groups.

Photoautotrophs

As with all the terms on this list, this one combines the prefixes related to energy and carbon source we saw in the past few paragraphs to create one descriptive word. Organisms in this category use the sun as their energy source (i.e., are phototrophs) and can fix CO_2 into glucose (i.e., are autotrophs). Land-based plants, algae, and certain bacteria are photoautotrophs. They can succeed in many environments, both terrestrial and aquatic, as long as light is available at least some of the time (carbon dioxide, their source of carbon, is found in both air and water).

Chemoheterotrophs

These organisms use existing chemical compounds for energy and biologically available compounds such as glucose as their carbon source. Most of the biosphere falls into this group. Animals (e.g., insects, humans, rodents, fish), fungi, protozoa, and most bacteria are chemoheterotrophs. Since they do not use the sun directly, they can live in light or dark places provided useable carbon is accessible to them.

Photoheterotrophs

A relatively small number of specialized bacteria use light energy but *cannot* fix their own carbon. Again, we can conclude that they must live in areas that receive some sunlight. Appropriate pre-formed carbon compounds—typically products and remains released by other organisms—must be present as well, though. These unusual organisms are usually found near the surfaces of bodies of water.

Chemoautotrophs

This last category, like the previous one, is made up of a comparatively small number of members. It consists of bacteria that need pre-existing chemical energy but fix inorganic carbon into glucose. Note that most chemoautotrophs do not use organic carbon compounds such as carbohydrates for energy, instead, inorganic compounds serve as their energy source. The sulfur-dioxide-eating bacteria mentioned near the beginning of this chapter are, arguably, the most famous (among those who keep track of such things) members of this group. In fact, the notion that chemoautotrophy could sustain life was only an unsupported hypothesis until the 1970s, when these organisms were discovered thriving in the complete absence of light near seafloor volcanoes.

5.2. ORGANISMS INTERACT TO MEET THEIR NEEDS

Now that we have been introduced to the basic needs of individuals, we are ready to study the many types of interactions seen within and among groups. Again, as we proceed through this discussion, you should not lose sight of the reason an organism operates as it does: it must fulfill the requirements presented in the beginning of this chapter if it is to survive. Put another way, the behaviors we will explore below only occur out of necessity. Non-human organisms do not act out of malice, greed, compassion, or other such impulses.

5.2.1. The hierarchy of ecology

One way to model ecological systems is with an inverted triangle. As is illustrated in Figure 5.1, complexity increases from bottom to top, with several different types of interactions possible at the top three levels.

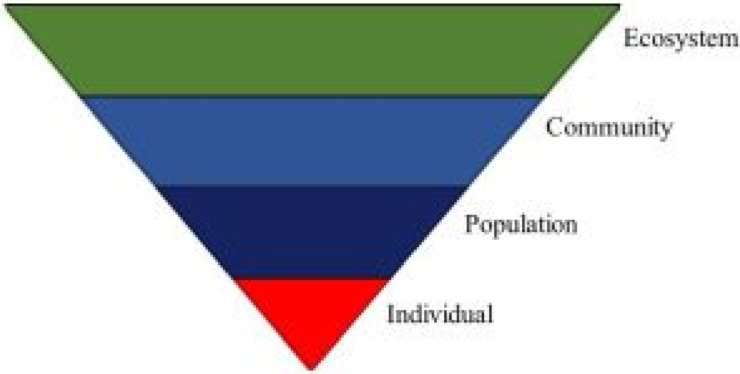


Figure 5.1. *An inverted triangle models the hierarchy of organization in ecology. The size of each color band indicates the relative amount of complexity within each level. The main text below describes the terms shown.* *Kelsey*, CC BY-NC-SA.

Individual

We know from Chapter 4 and elsewhere that an entire living being is hardly the least complex entity in the universe. Nonetheless, an individual is the simplest unit we will recognize as we begin to consider interactions. The list of possible examples would of course be extremely long, consisting of many millions of names of organisms. For our purposes, we will use one red-tailed hawk to represent this first level in our hierarchy.

Population

The next level in our scheme is the first one in which we start to see interactions. A population is defined as a group of individuals that are all members of the same **species**, that is, they can breed with each other to yield fertile offspring (see Box 5.4 for more about the species concept). In addition to reproducing, members of a population also can cooperate or compete. Note that this term can be applied on different scales: in principle, it includes every individual of a species on Earth, but it is not uncommon to identify a small subgroup in an area as a population of interest. For example, all the red-tailed hawks living and directly interacting near a mountain in, say, the state of New York, U.S.A., might make up a population studied by ecologists.

Intraspecies *interactions are those occurring* ***within*** *a population, and* ***interspecies*** *interactions are those occurring* ***among*** *multiple populations.*

Box 5.4. Horses and donkeys and bacteria (oh my?)

The definition of **species** is one of the more straightforward ones in ecology. Two individuals that can mate and produce fertile offspring are, by definition, members of the same species. It is therefore simple to objectively group organisms in this way. How similar they appear, how much they have in common genetically, and other potentially subjective criteria about whether they resemble each other sufficiently do not come into play. So, although a dog and a cat may share many characteristics (four legs, fur, tail, claws, and so forth) their inability to breed with each other puts them in separate groups. What about horses and donkeys? As you may know, they are closely related—so much so, that a union between a male donkey and female horse produces a mule. However, since these offspring are sterile (i.e., cannot reproduce), they are not members of the same species. It is all very logical and pleasing…until we consider bacteria. It turns out that the species concept only neatly applies to organisms that reproduce sexually, that is, where two parents combine their genetic material to produce one new individual. As we will see in Chapter 6, bacteria and some other microorganisms are asexual: a single cell divides into two progeny that are identical to the original cell. Further complicating things, bacteria are not particularly different in their shapes and appearances (i.e., there are many more species than there are different physical forms; Chapter 3). So they sort of look alike, and the offspring test is not applicable. How then can we make sense of this diverse group? Scientists have been wrestling with the problem for quite some time. When bacteria were first identified, microbiologists had to fall back on a comparison of physical and chemical characteristics, an approach fraught with subjectivity and uncertainty. During the past few decades, methods that allow for an objective analysis of DNA have evolved, making genetic relatedness increasingly important as a way to categorize organisms. Research in this area is ongoing.

Community

A group of populations that interact directly within a given space make up a community. Returning to our example, all the hawks, other birds of prey, songbirds, rodents, grasses, trees, and other organisms found on that New York mountain would be a community.

Ecosystem

This term is used by many ecologists to describe a level that encompasses the others in our hierarchy. It consists of all biological, physical, and chemical components active within a defined space. So, we would add the soil, air, and water—the relevant physical components and their properties—to the list of organisms imagined in our community above to define an ecosystem.

Why ecosystems matter. This, the most *complex* entity on our list, is the minimum unit necessary to sustain an individual, the *simplest* entity on our list. As we have seen previously, organisms are inextricably linked to their environments; they neither exist nor succeed in isolation, and they interact with their surroundings to fulfill their requirements. If an ecosystem fails to provide adequately, the individual organisms depending on it will fail as well. Furthermore, humans derive benefits from what are termed **ecosystem services**, including habitat for valued organisms, primary production, food, and important contributions to the **hydrologic cycle**.

Characteristics of ecosystems. In principle, Earth is one large ecosystem, although it usually is conceived and studied as a collection of many distinct ecosystems. For each, two important characteristics can be identified.

Structure

Although it has a different meaning outside of ecology, we will use this term to refer to the **identity of the organisms present** in a given ecosystem. Structure varies among ecosystems. For example, it is likely unsurprising that the community in a desert in northern Africa is not the same as that in a tropical rain forest in South America, the ocean near northern Australia, or the tundra in Russia. Since environmental conditions are so different in these places, so too will be the survival strategies used by organisms living in them. However, even in locations with similar conditions, such as the deserts in North America, structure will likely differ as well.

As we know, environmental conditions affect organisms—it does not take much variation in temperature, moisture, or other properties to alter the organisms dominating an area.

Function

This term refers to the processes active within an ecosystem. Unlike structure, this characteristic does not vary much from place to place. Regardless of the specific organisms present, the following processes or functions are active in all ecosystems.

Primary production. As we saw near the beginning of this chapter (as well as in Chapter 4), certain organisms are able to convert unusable carbon compounds into forms that are accessible to all members of the biosphere. Put another way, they produce fixed carbon for themselves and others. Remember that these autotrophs are almost always phototrophs as well: so, they also convert biologically *unavailable* energy from the sun into biologically *available* chemical energy (see Box 5.5 for a note about dark environments).

> **Box 5.5.** What about places where the sun doesn't shine?
>
> Note that in those systems lacking sunlight, primary production is carried out by chemotrophic bacteria (see the early part of this chapter and Box 4.5). They are indispensable because of their ability to take a non-biological source of energy—in their case, that in inorganic chemical compounds like hydrogen sulfide instead of solar energy—and turn it into a biological form.

In terrestrial systems such as forests, deserts, and prairies, photoautotrophic plants such as trees, cacti, and grasses carry out primary production, respectively. This critical function is typically the work of photosynthetic algae and **phytoplankton** in oceans, lakes, and other aquatic ecosystems (**plankton** are generally small, free-floating organisms). Figure 5.2 shows examples of primary producers.

***Figure* 5.2.** *Different primary producers: left (a), green plant; middle (b), microscopic green algae (individual cells are approximately 10 microns in length); right (c), visible colonies of algae color surface water in a pond.* Vinayaraj, CC BY (a); Aaron Carlson, CC BY (b); Micropix, CC BY (c).

Despite the dramatic diversity in form, they share some common traits. Notably, their green color is associated with the machinery necessary to carry out the same function: **photosynthesis**. In other words, although adapted to very different environments, they all serve as primary producers for the ecosystems of which they are members.

Consumption. Recall from the discussion of the carbon cycle (Chapter 4) that **carbohydrates** and other organic forms of C move within the biosphere as organisms eat—consume—each other. This function is necessary in all healthy ecosystems as it provides chemoheterotrophs with the nutrients they cannot produce themselves and ultimately leads to the recycling of elements after the consumers die (see the next function on this list). The number of consumers on Earth is enormous, so coming up with examples is not difficult. On land, herbivores (plant eaters) such as zebra and deer and **carnivores** (meat eaters) like lions and wolves are among the organisms that carry out this function (Figure 5.3.) (more about herbivores and carnivores is presented below). Many **microorganisms**, including **protozoa**, **fungi**, and certain **bacteria**, are also consumers.

***Figure* 5.3.** *Four terrestrial consumers, above left (a), white-tailed deer; above right (b), zebra; below left (c), wolf; below right (d), lion. Averette Vinayaraj, CC BY (a); Danesman1, CC BY (b); Malene Thyssen, CC BY (c); Winfried Bruenken Vinayaraj, CC BY (d).*

In water-based ecosystems, consumers range from microscopic **zooplankton** (they eat phytoplankton) to small fish, seals, tuna, sharks, and dolphins (Figure 5.4).

Figure 5.4. *Four aquatic consumers: above left (a), zooplankton (~ 4 mm long); above right (b), seabass; below left (c), tuna; below right (d), dolphin. Uwe kils,* CC BY *(a); w:en:*Aquaimages, CC BY *(b);* US NOAA, *Public Domain (c); Faraj,* CC BY *(d).*

Decomposition. This third function is carried out by specialized consumers that eat the remains and waste products of organisms. Through it, nutrients are released and recycled. Decomposers also ensure that **dead organic matter** (Chapter 4) does not accumulate to any significant extent. Fungi and certain bacteria (distinct from the producers and consumers noted above) are responsible for much of the decomposition that occurs both on land and in water (Figure 5.5).

Figure 5.5. *Several images of white-rot fungus, an important terrestrial decomposer. Among other things, it can break down the complex chemical building blocks of wood. Henk Monster, CC BY.*

Of course, as we know by now, the specific decomposers found in different ecosystems must different—that is, the structure of decomposers varies even as the function they perform is consistent. For example, places that lack oxygen feature **anaerobic** decomposers whereas **aerobic** decomposers are found in oxygenated environments. Related to decomposers are organisms known as **detritivores** (detritovores are distinct from decomposers because the former are large enough to ingest distinct clumps of decaying material whereas the latter are so small that they act on and break down materials at a microscopic level). They feed on small organic fragments—known as **detritus**—of dead tissues and excrement. Terrestrial earthworms and some insects as well as ocean crabs and shrimp are good examples of these types of organisms (Figure 5.6).

Figure 5.6. *Two detritivores: left (a), terrestrial earthworm; right (b), aquatic ghost crab. Rob Hille, CC BY (a); Laslovarga, CC BY (b).*

Habitats and niches. These two terms denote distinct ecosystem properties. **Habitat** refers to the location in which an organism lives, whereas **niche** refers to the function an individual carries out—the role it plays, how it survives—in its ecosystem. Consider the three simple examples below, and note the difference in the way the words are used.

Great white shark

This fish lives in the coastal waters off the state of California, U.S.A. (its habitat) where it consumes large mammals such as seals (part of its niche).

Oak tree

This large plant grows in a forest in New Jersey, U.S.A. (its habitat) where it provides nesting opportunities for birds and squirrels, acorns for squirrels to eat, and is a primary producer (its niche).

Mosquitoes

These insects live near a lake in the African country of Kenya (their habitat) where they consume the blood of mammals and serve as a food source for bats (their niche).

Biomes. Simply put, a **biome** can be thought of as a type of ecosystem, an area characterized by a particular set of living and non-living features. Ecologists have identified several different biomes on Earth, defining and distinguishing them on the basis of climatic factors such as prevailing temperatures and abundance of precipitation as well as the types of dominant organisms (generally, the primary producers) present. The details of this topic are unnecessary for us, but future discussions (below and in Chapter 6) will benefit from an understanding of some essential concepts related to biomes.

How many biomes are there?

Unfortunately, the answer depends on who you ask. Some ecologists prefer to define a small number using very general criteria whereas others have established a larger number of narrowly defined biomes. In that context, it would be pretty safe to say the answer is in the range of five to fifteen.

What are Earth's biomes and what organisms live in them?

Clearly the response to the first biome question suggests a complicated and variable answer is likely to be found here as well. We will keep things appropriately simple for our purposes. The two broadest types are aquatic (water) and terrestrial (land). Beyond that distinction, Earth's major biomes are presented in Table 5.2.

Table 5.2. A *sampling of biomes and their defining features*[3].

3. Karla Moeller. 2013. Boundless biomes. Arizona State University School of Life Sciences Ask A Biologist. https://askabiologist.asu.edu/explore/biomes

BIOME TYPE#, NAME[4]	DEFINING CHARACTERISTICS	DEFINING ORGANISMS	SAMPLE LOCATIONS
Aq: Marine	High salt content, wide temperature range	Algae, fish	Atlantic and Indian Oceans
Aq: Freshwater	Low salt content, wide temperature range	Algae, fish	Hudson River (U.S.A.), Nile River (Sudan)
Ter: Tropical rainforest	Rainiest biome, fewer than four seasons, warm all year	Many trees (not fir), ferns, flowers, monkeys, birds, insects	Brazil, India, Madagascar
Ter: Temperate rainforest	Second rainiest biome, experiences all four seasons	Many trees (including firs), rodents, birds, insects	Pacific northwest of North America, Coastal Chile, Britain
Ter: Temperate forest	Third rainiest biome, experiences all four seasons	Many trees (mostly leaf, not needle), rodents, bears, birds, insects	Northeastern U.S.A., Southern Canada, China
Ter: Grasslands	Generally less rainy than temperate forest, hot, dry summer, cold wet winter	Grasses and animals that eat grasses	Midwestern U.S.A., Argentina, Russia
Ter: Savannah	Intermediate between forest and desert	Grasses, trees, lions and their prey	About half of Africa
Ter: Desert	Second driest biome, can be very hot in the summer, cool in the winter in equatorial regions (very cold winters in polar regions)	Cacti, sagebrush, small mammals, reptiles, birds (polar: algae, grasses, penguins)	Southwestern U.S.A., Ethiopia, Antarctica

Ter: Tundra	Coldest and driest biome	Algae, shrubs, caribou, polar bears	Alaska (U.S.A.), Siberia (Russia)

Some words of caution should be offered before we move on. First, note that the table is not an exhaustive list, but a sampling for the purposes of illustration. A number of other biomes and characteristics could be included if we had additional space. Second, you should realize that, although some of the same names are used to indicate organisms present in different areas, species present vary considerably among biomes. For example, the term "birds" is used several times. As knowledgeable ecologists we know that only birds adapted to the narrowly defined environmental conditions in an ecosystem will survive there. So, the birds in a tropical rainforest will surely differ from those in a temperature forest.

Can the same biomebe found in more than one place on Earth?

If you think of a biome as a type of ecosystem you will conclude that the answer to this question must be "yes". Additionally, a further examination of the final column in Table 5.2 indicates that these types of ecosystems can be found in many places around the world.

5.2.2. Organisms can help or harm each other

Whether participants receive benefits or are inhibited is a fundamental and defining feature of the many possible interactions among individuals and groups. Three commonly observed interactions are briefly described here, although you should realize

4. Aq, aquatic or water based; Ter, terrestrial or land based

that these groups represent what might be viewed as end points on a continuum. In other words, the extent to which organisms are helped or harmed is not always easy to describe or categorize, and the boundaries among them can get blurred.

Competition

Everybody suffers. In this case, two or more individuals or populations are rivals for space, light, water, food or another resource that is in limited supply. Because none of them is able to access all of what it could utilize in the absence of competitors, **all participants in this interaction are harmed**. Natural systems are often characterized by scarcity, so organisms must develop effective strategies to obtain energy and materials in any case (what we have called *adaptations* previously). When demand for a critical nutrient exceeds supply, pressures on individuals and populations increase. If you recall how important the process of reproduction is to organisms, the reason competition brings on adverse consequences should become clear. Instead of devoting all its available energy to ensuring its genetic information is successfully passed to subsequent generations, an individual must expend precious nutrients in its struggle with others seeking to utilize the same resources. Note that competition does not necessitate or even imply direct confrontation among the participants, just that multiple parties inhibit each other through their efforts to meet their needs.

One ultimately wins. Ecologists have repeatedly observed that rivals for the same resources do not persist indefinitely in the same ecosystem. **The competitive exclusion principle** is an important concept—rule, really—that has developed as a result. In brief, it states: individuals of different species that have the same nutrient requirements cannot coexist. Put another way, they cannot occupy the same niche in the same space and time. What happens when multiple species try to access the same resources in the same place? The weakest competitor, the one least able to obtain what it needs, will disappear from that ecosystem. When this occurs, the winner, if you will, is unburdened by competition and will likely enjoy greater success as it can devote more of its resources to reproduction. On the other hand, the so-called loser could experience one of a few fates: become extinct (complete disappearance from Earth; see Chapter 6) if it is so specialized that it cannot exist elsewhere, adapt to occupy a different niche and exploit different resources in that ecosystem, or migrate to a location characterized by less intense competition.

Competition is common. It would be fair to say that competition is one of the most important forces shaping the living members of Earth's numerous ecosystems, and this type of interaction is widespread. Given our space limitations, we will only consider a few good examples here.

Red Squirrels and Grey Squirrels

Aside from their obvious color differences, these two distinct species are strikingly similar (Figure 5.7).

Figure 5.7. *Competitors: left (a), red squirrel; right (b), grey squirrel. awel Ryszawa, CC BY (a); Natalie-S, CC BY (b).*

They can live in the same habitats and consume the same food sources, but they cannot coexist without diminishing the success of one other. Observations made of populations sharing habitat in

England have revealed the nature of their relationship and taught us a great deal about the importance of this type of interaction. The story gets somewhat convoluted, but in short: red squirrels are native, or endemic, to England, whereas grey squirrels are native to North America. Although these two species are equally successful when they are the only squirrels present in an area, red squirrels are currently being outcompeted by grey squirrels where they share habitat. Why do they overlap at all if the two species come from different continents? Well, back in the late 1800s humans intentionally imported them to Scotland, possibly for purely aesthetic and whimsical reasons. The two organisms never fought directly, but they became instant competitors. Unfortunately for the native reds, due to their higher rates of reproduction, more efficient use of available local food sources (abundant Oak Trees), and other advantages, the grey population grew quickly and spread aggressively. Now, many people fear that the red squirrels will disappear from the UK completely unless the grey squirrels can be controlled—something that has so far eluded us. Keep in mind that the competitive exclusion principle does not favor natives, just better competitors!

> ***Endemic species*** *developed in their current habitat whereas* ***exotic species*** *developed elsewhere and somehow have moved into a new habitat.* ***Invasive species*** *are exotics that outcompete endemics.*

Black Walnut and other trees

The next time you take a walk in the woods, look for walnut trees growing in fields surrounded only by grasses (Figure 5.8). How is this possible? Black Walnut trees inhibit other trees that could encroach on their space and interfere with their success by releasing toxic chemical compounds into the soil. The poison affects other trees, but not grasses in an example of **selective toxicity** (Box 5.6, below, plus Chapters 9 and 15).

Figure 5.8. *Black walnut trees release a chemical into soil that inhibits other trees. Thus, they often stand alone in fields of grasses. Jim Linwood, CC BY.*

Two similar species of protozoa

Our final example comes from the unseen yet vibrant microbial world we explored in Chapter 3. Even though extremely small by our standards, organisms such as protozoa and bacteria interact in ways like organisms we can see with the naked eye. Similar to the squirrels above, *Paramecium caudatum* and *Paramecium auerlia* (two species of protozoa found in water—Chapter 6 presents more about the classification and naming of these protozoa and all members of the biosphere) are equally successful when grown in isolation and fed the same bacterial food source. However, when they occupy the same space, like a glass flask in a laboratory, both are slowed to some extent. *Paramecium auerlia* grows more quickly though, leading to the eventual disappearance of *Paramecium caudatum*; this is then followed by an increase in the success of the remaining, better-adapted, organism.

Mutualism

Everybody wins. Unlike competition, this cooperative interaction **provides an advantage to all involved**. Through different possible mechanisms, mutualism increases the success of participating individuals and populations. Under certain circumstances, these associations can even enable organisms to survive in environments they would otherwise find too hostile.

Mutualism is common. We are surrounded by cooperative relationships and even participate directly—generally unconsciously—in some ourselves. In fact, mutualism is widespread, affecting organisms throughout the biosphere. The specific ways organisms are connected are so diverse, however, that it would be difficult to generalize about them beyond the fact that they are mutually beneficial. A small number of examples will again have to suffice to illustrate this important and complex phenomenon.

Microorganisms and plants

Some bacteria can live within specialized structures of certain plants (mostly legumes such as alfalfa, clover, and beans) and fix nitrogen gas (N_2) to a usable form of nitrogen such as ammonium (Chapter 4). The plant can then use the ammonium to produce its own proteins. In return, the plants provide fixed carbon for the resident N-fixing bacteria. The importance of this mutualistic relationship is hard to overstate as farmers take advantage of it to grow crops without the need to add nitrogen fertilizer to soil. We will return to the topic of fertilizers and soil in Chapter 9.

Microorganisms and cows

You likely have noticed dairy cows grazing in fields at some point in your life, but you may not have considered the complex mutualistic relationships they have developed with some tiny organisms (Chapter 3). Although they can ingest enormous amounts of grass, cows are utterly[5] dependent on the microbial community living within them to obtain nutrients from their food. In short, one of the chambers along the digestive tract of a cow, known as the **rumen**, is a very hospitable environment for certain bacteria and protozoa. Chemical bonds within the swallowed grass are broken by the microorganisms in the rumen, and the released nutrients are used by the cooperating symbionts. This association is critical because cows would otherwise starve to death—with full stomachs.

Microorganisms and humans

People are similar to ruminants in their reliance on bacteria and protozoa to help digest some of the food they ingest. Like cows, humans obtain nutrients from plant materials through the actions of specialized microorganisms living inside the intestines. Symbiotic organisms perform other services for us as well, including providing certain important vitamins and discouraging infection and disease by pathogens. For our part, we contribute nutrients and an appropriate environment for these microscopic partners. Review Chapter 1 for more on this relationship.

Deep-sea fish and bioluminescent bacteria

Many organisms living in deep, dark waters have developed mutualistic relationships with **bioluminescent bacteria** (i.e., they can convert some of their energy into light). In the case of the anglerfish, the bacteria live inside a long, thin structure, an esca, that is attached to the animal's head and dangles near its mouth (Figure 5.9).

5. Pun intended??

Figure 5.9. *An angler fish is well adapted to live in dark water due to the mutualistic bioluminescent bacteria living in its esca. Emőke Dénes, CC BY (modified by Kelsey).*

They receive nutrients and a suitable habitat, and in return light up and cause the end of that protrusion on the fish to glow. This arrangement provides a useful advantage for anglerfish: the esca serves as a lure for unsuspecting prey attracted to one of the only light sources in the dark water (hence the name "anglerfish").

Fungi and algae grow as lichens

Our final example is one of the most famous and well-studied cooperative associations in ecology. In this case, the relationship is so highly developed and specific that two separate organisms, fungi and algae, appear to be a third entity, a lichen (Figure 5.10).

Figure 5.10. *Certain algae and fungi can enter into a tight mutualistic association and take on a third form: left (a), common green alga growing on a rock; middle (b), common fungus (with a leaf balanced on top of it); right (c), common lichen.* Ryan Hodnett, CC BY-SA (a); Jenstasi, CC BY-SA (b); Kelsey, CC BY-SA-NA (c).

Because of the contributions made by each of the partners, lichens are successful in locations too hostile for fungi or algae to grow on their own. Algae are primary producers, providing fixed carbon and energy, and fungi have structures that enable them to obtain and hold onto water as well as other nutrients that may be in short supply. Lichens can grow on barren rocks, dead trees, and even the sides of buildings, and play an important role in the process of ecological succession (described in an upcoming section of this chapter).

> *Note that **symbiosis** is often erroneously equated with mutualism, but the former general term only refers to associations in which organisms live together (for mutual benefit or otherwise).*

Parasitism

Winners and losers, part 1. This is one of two types of interactions through which **organisms derive benefits at the expense of others**. Here, a parasite causes harm to its **host** by living on or inside that larger organism (the other case, predation, is described below). Although we will only consider a small number of examples of parasitism, keep in mind these types of relationships are varied and numerous.

Pathogenic bacteria and humans

We have seen previously that members of the group known as bacteria carry out an enormous number of vital functions and are indispensable to the biosphere. Those that are parasitic, though, can bring on diseases that run the gamut from mild to deadly. For example, strep throat, tuberculosis, syphilis, cholera, and pink eye are among the many conditions caused by bacterial infections. A few cells of the pathogen can enter a person, evade defense mechanisms, and rapidly increase in number inside the body system that provides the best environment for them (throat, eye, etc.). The details vary widely among different diseases, but in general terms the bacteria are parasites: they divert nutrients, inhibit normal functions, cause tissue damage, or otherwise harm their host. Now, humans employ some natural mechanisms to fight infection, ranging from non-specific responses like fever to cells that target specific invaders, but they are not always successful. What if body defenses fail to eliminate the parasite? Permanent damage or death may occur without medical intervention such as the administration of chemical antibiotics (see Box 5.6 for more). Viruses and other microbes can behave like parasites as well, with outcomes like those brought about by bacteria. Treatment of diseases they cause, however, is not the same.

Box 5.6. Antibiotics: winning a battle, losing the war, against pathogenic bacteria

We have been fighting disease-causing bacteria for as long as humans have been on Earth, but up until relatively recently, there was little we could do against many of them. People infected with pathogenic bacteria had to, for the most part, depend on their immune responses that are only partially effective against such organisms or on medical interventions that might treat or alleviate the severity of symptoms yet do little or nothing to actually eliminate the invaders and the diseases they cause. Then, in the 1920s, an important discovery changed the balance of power: penicillin, a chemical substance produced naturally by a certain type of mold, was found to destroy many harmful bacteria. Starting in about the 1940s, antibiotics gained wide availability and were able to cure diseases such as tuberculosis, cholera, syphilis, strep throat, and others that had killed so many people throughout history. The success of penicillin and the other antibiotics used since is attributed to a phenomenon known as **selective toxicity.** Put simply, different organisms can react very differently when exposed to the same substances. Although they have much in common as living entities, humans and bacteria are distinct in several important ways. For example, most pathogenic bacteria build cell walls outside of their cell membranes (Chapter 3) whereas animals lack such structures. Consequently, humans can ingest penicillin, which targets wall synthesis, to kill off bacteria that are making us sick without causing undue harm to ourselves (aside from people who are allergic to the drug). As brilliant as these medicines may seem, some serious adverse effects come along with their use. For one, in addition to killing pathogens, antibiotics can also kill mutualistic bacteria inside places like our intestines, making it temporarily difficult to properly digest food. However, a second phenomenon is far more problematic: the development of **antibiotic-resistant bacteria**. During the decades since we discovered penicillin, many populations of bacteria that were once vulnerable to antibiotics are no longer killed by such chemicals. In practical terms, this means you might effectively treat strep throat with penicillin a few times but then discover the drug has no effect on a subsequent bout with the disease. Because bacteria can quickly evolve a way to overcome the adverse effects of the drug, they continue to live and cause harm to their host. How? Recall that organisms are affected by the stressors they encounter in their habitats. To be blunt, a species will either fail (die or at least be forced out) or adapt a survival strategy and persist. An antibiotic introduced into any system will kill only susceptible bacteria, leaving the few unaffected cells to reproduce with reduced competition. Once their numbers increase, these resistant individuals can continue to cause disease unhindered. We have a lot of evidence of this phenomenon in public-health records, but it can also be witnessed in real time. Laboratory experiments have repeatedly shown how a population of bacteria exposed to an antibiotic can change from sensitive to resistant in a matter of days (more in Chapter 6).

Should we be concerned about antibiotic-resistant bacteria? The short answer is: yes. As a population of pathogens becomes resistant to an individual antibiotic, a new chemical that targets different sensitive structures in the bacteria may be employed to fight a disease (e.g., instead of penicillin to block cell wall construction, you use tetracycline because it disrupts other prokaryotic processes). Unfortunately, the response to that second antibiotic will be as it was for the first, yielding a population that is now unaffected by *two different* drugs. A third and a fourth, and many other antibiotics can be tried, but the outcome is always the same. Certain pathogenic bacteria have even become resistant to every known antibiotic. One instance of multi-drug resistance recently identified by public health officials is a strain of bacteria that causes the common sexually transmitted disease gonorrhea. This once readily treatable condition now could potentially persist in infected individuals, causing many long-term health problems. A return to the situation that prevailed prior to the discovery of penicillin, that is, we possess no effective defense against bacteria, may well occur in the not-so-distant future.

Can we do anything about antibiotic-resistant bacteria? The short answer is: not much. Arguably we have vastly increased the pace of an inevitable process by the way we have used antibiotics, but it may be impossible to prevent in any case. We could probably *slow* the spread of antibiotic resistance among all of Earth's bacterial populations by taking the following steps.

1. Use far fewer antibiotics. Keep in mind antibiotics are not effective against all microorganisms. For instance, viruses and protozoa are insensitive to penicillin and other drugs that target bacteria. Public health officials note that many physicians, particularly in the U.S., overprescribe antibiotics, using them to

combat conditions that will generally run their course without chemical therapy. So what's the harm in being extra cautious? Well, a large amount of those antibiotics we ingest exit our bodies in urine and enter natural water systems. Thinking like an environmental scientist, you might picture what happens when populations in soil and water are exposed to antibiotics: they adapt. Put somewhat differently, as we use more and more of these chemical substances, the very organisms we are trying to kill gain more and more of an advantage.

2. Develop new antibiotics. This step may seem rather straight forward and obvious. In principle, we might be able to keep better pace with bacterial evolution by quickly and regularly designing new drugs, but there is currently little effort being devoted to this goal. The reasons for this apparent lack of interest are numerous. Suffice to say, the outcome of it could include an enormous public health crisis in the coming years.

Some have called it an arms race, and humans have gained only a temporary advantage in it.

Tape worm inside a dog

Anyone who has ever lived with dogs knows that these animals will eat nearly anything that they perceive might be food—roadkill, excrement (their own or others'), debris they remove from their own bodies, many inanimate objects—with gusto and seemingly no concern about the potential consequences. Unfortunately, among other problems arising from their indiscriminating palate is the tendency to ingest fleas and the tapeworm larvae they can carry. These parasites grow inside the digestive system and cause harm by eating some of the nutrients consumed by the dog. This association between the two organisms can last for a relatively long time, as it is generally not all that dangerous to the animal (medical intervention will kill the worm). Incidentally, humans are also susceptible to tapeworm infestation. In some cases, people have been known to intentionally ingest the parasite in the interest of weight loss[6].

Wasp larvae and tomato hornworms

One of the scourges on tomato growers is a green caterpillar known as the tomato horn worm. These insects can eat the leaves of the plants and destroy crops. Interestingly, this pest is itself subject to a parasite. In short, a certain type of wasp will lay its eggs on the skin of the hornworm. The **larvae** that hatch from the eggs slowly eat the caterpillar until they emerge as adults. While they are feeding and developing, the larvae can be seen protruding from the back of the still-living hornworm (Figure 5.11). This relationship ultimately ends in the death of the host, unlike some of the parasitic relationships described above.

6. For the record: this is a bad idea.

***Figure* 5.11.** *A tomato hornworm with parasites. Max Wahrhaftig, CC BY.*

Predation

Winners and losers, part 2. Predators are like parasites in that **they obtain what they need while bringing about some level of harm to another organism**. However, this relationship differs from the former in three critical ways. First, parasites tend to be small relative to their hosts whereas predators are larger than, or at least capable of overpowering, their prey. Second, the length of time parasites can interact with their hosts is almost always longer, ranging from days to years, than that of predators and their prey, which is in the range of minutes. Nutrients stored in **prey** organisms move to **predators**, as the former are killed and consumed by the latter (more about this flow of nutrients within ecosystems is described in the next section). Third, although predation is almost always fatal to the prey, parasitism need not kill the host (think dogs and fleas). One final point bears consideration:

generally speaking, the term "predator" is reserved for organisms that eat other heterotrophs like animals or certain microorganisms (animals that eat the flesh of other animals are also known as **carnivores)**. Plant-eating consumers tend to be called **herbivores** to reflect the fact that they eat plants or other plant-like cells. A small number of organisms, humans, as well as certain rodents and insects for instance, are appropriately classified as **omnivorous** because they can eat both animals and plants.

Predators eat to survive. Consumers get energy and materials by eating other organisms. Recall from earlier in this chapter that, although they could appear to be violent or cruel when viewed through the lens of human emotions, predators simply act in response to their fundamental needs. Predation is active throughout the biosphere, but as before, we will only briefly consider a few examples of it.

Spiders consume insects

A number of spider species build webs to catch unsuspecting insects. Once ensnared, the prey is wrapped up by the predator to be eaten later (Figure 5.12).

Figure 5.12. *A wasp caught in a web is wrapped up by a spider. Erik Baas, CC BY.*

Orcas consume seals

Killer whales, also known as orcas, are large predatory animals which eat other animals such as salmon, seals, and tuna (Figure 5.13). Those prey species are themselves predators on smaller organisms. We will examine feeding relationships in ecosystems, known as **trophic interactions**, in more detail shortly. The role that predation has on the movement of pollutants in ecosystems is another important topic, one we will cover in Chapter 15.

Figure 5.13. *An orca with its prey. Robert Pitman, Public Domain.*

Protozoa consume bacteria

Predation is a way of life at all scales, from the familiar visible animals noted here to the microscopic world we explored in Chapter 3. Like orcas, spiders, lions, wolves, and so forth, protozoa are consumers that obtain nutrients through ingestion of other, smaller, organisms such as bacteria. In fact, a single protozoan can eat thousands of individual bacterial cells during its life cycle. This predation, or grazing, plays an important role in many ecosystems because it can control the numbers of bacteria present and can serve as the basis for many ecosystems.

5.3. MATERIALS AND ENERGY MOVE THROUGH ECOSYSTEMS

As we learned in earlier parts of this chapter, individuals must fulfill specific needs to survive. Although a variety of strategies can be used, most often materials and energy are obtained from other organisms. Put simply, producers obtain nutrients and convert them into biologically available forms, and then consumers live off of those producers directly or indirectly.

5.3.1. Hierarchy of feeding relationships

Trophic levels

Organisms are grouped according to the number of steps they are from the original source of energy (typically the sun, as we saw above) for a particular ecosystem. Each of these groups is called a **trophic level** (Figure 5.14).

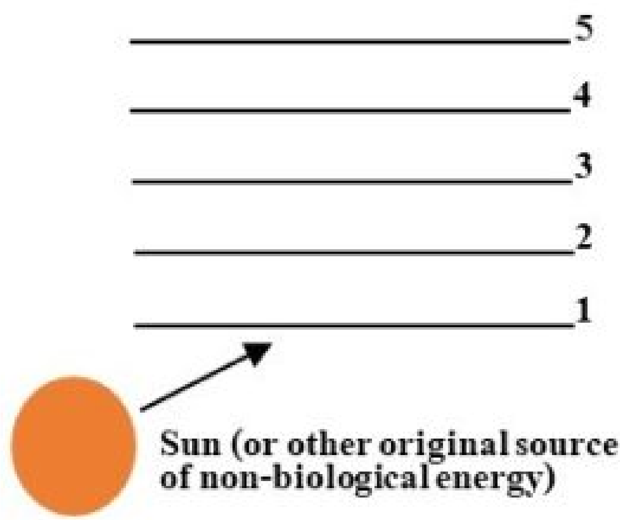


Figure 5.14. *A simplified model of trophic levels in any ecosystem. Each line represents all the organisms feeding the same number of steps away from the original source of energy, and the numbers indicate trophic levels. Kelsey,* CC BY-NC-SA.

For example, since they obtain their energy directly from the sun, photoautotrophs such as plants and algae, as well as all other primary producers, are in the same category: the first trophic level. Organisms that consume those at the first are in the second trophic level and are themselves consumed by predators at higher levels. A simple, idealized representation of the arrangement and characteristics of trophic levels is presented in Table 5.3.

***Table* 5.3.** *Characteristics of different trophic levels. In some cases* "**RANKING / NICHE**" *provides more than one possible term for organisms in a given level to suggest different ways to view their roles.*

TROPHIC LEVEL	RANKING / NICHE	DIRECT ENERGY SOURCE	EXAMPLES
Fifth	Quaternary consumers, Tertiary carnivores	Organisms at the fourth and likely third and second trophic levels	Eagles, sharks
Fourth	Tertiary consumers, secondary carnivores	Organisms at the third and likely second trophic levels	Foxes, seals
Third	Secondary consumers, primary carnivores	Organisms at the second trophic level	Songbirds, small fish
Second	Primary consumers, herbivores	Primary consumers	Caterpillars, zooplankton
First	Primary producers	Non-biological energy source like the sun	Plants, phytoplankton

Note that organisms are ranked on the basis of the highest level they could *possibly* occupy. So, in the table, the eagle is placed in the fifth trophic level because it can feed on organisms in the fourth trophic level (e.g., the fox). However, it may skip down to a songbird in the third trophic level if presented the opportunity to do so. Since few organisms can eat both plants and animals (i.e., are omnivores, as define above), the diets of high-level consumers tend to be restricted to prey in the second level and above.

Food webs

One useful way to visualize trophic relationships is through a diagram known as a food web. This kind of graphical representation shows the possible interactions among trophic levels, or put somewhat whimsically, can be used to answer the question "who eats whom?". Importantly, our model is *not* a chain, but a web. Although people commonly use the term "food chain" in reference to trophic interactions (and even outside the realm of ecology as a metaphor in any hierarchical system, e.g., "since she is higher on the food chain, she makes more money"), a linear structure vastly oversimplifies the situation. Figure 5.15 enables a comparison between the use of chains and webs in this context.

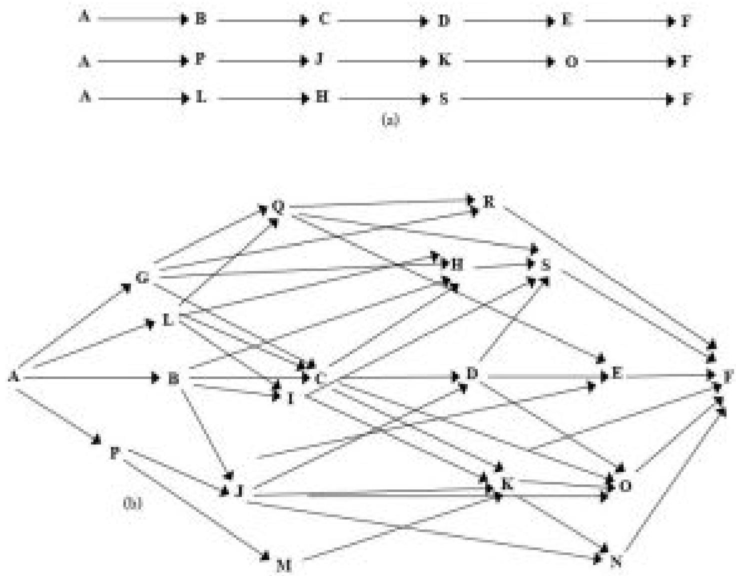


Figure 5.15. *Three separate food chains (a) vs. one food web (b) to model trophic interactions. Each capital letter represents a different organism, and arrows indicate direction of energy flow (e.g.,* A *is eaten by* B, B *is eaten by* C, *etc.). Kelsey,* CC BY-NC-SA.

Note how chains necessarily and unrealistically suggest that each organism has only one source of food and one predator, whereas a web more accurately presents the complex and varied nature of the movement of materials and energy from lower to higher trophic levels. Put another way, one organism can eat and be eaten by a number of different other organisms. The food web modelled by Figure 5.15 implies six possible trophic levels (**A** is at the first, **F** is at the sixth), although a careful examination of it reveals that energy does not always move through six levels from organisms **A** to **F**. In fact, multiple possible pathways exist (examine the figure to see for yourself, reading left to right). In any case, the organism that occupies the highest trophic level, the one that is not ordinarily a source of food for any other, is referred to as an **apex predator** (more below). A diagram of an example food web studied in the Chesapeake Bay is shown in Figure 5.16.

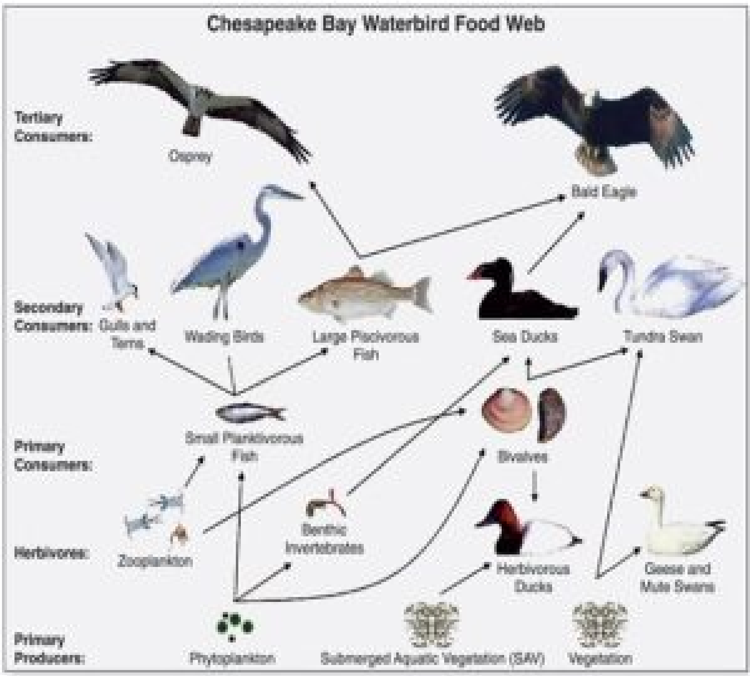


Figure 5.16. *An example food web from the Chesapeake Bay, U.S.A. Note: "benthic" refers to those living in the sediments of the bay. USGS, Public Domain.*

The number of trophic levels and the complexity of the webs can vary considerably among ecosystems. Broadly speaking, extreme, barren, or ever-changing environments are characterized by fewer organisms and simpler models. In a very dry desert, for example, it is more difficult to develop proper adaptations and survive, so there are fewer different species present. Put another way, there is low **biodiversity** in such an environment Trophic interactions are therefore relatively simple in such places. On the other hand, a more hospitable, stable, or nutrient-rich habitat such as a coastal ocean or tropical rainforest tends to support a much wider variety of species and has a more complex food web as a result.

> *In simple terms, **biodiversity** refers to the total number of different species present in an area. See Chapter 6 for more.*

5.3.2. How many trophic levels are possible?

On first pass, this question might appear to be inconsequential, one that is a simple matter of book keeping suited for academic discussions only. However, the answer is rooted in some fundamental principles we studied in Chapter 4 and is critical to a discussion of ecosystems.

Available energy decreases with increasing trophic level

We learned previously in this chapter that energy is converted into a biologically available form by members of the first trophic level. In serving as food sources, these primary producers then supply energy to higher-level consumers. The amount of usable energy that can pass from one level to the next is quite low, though, because most of it is used to do work or passes to decomposers before consumers can access it. You might be surprised to learn that somewhere between 90 and 99% of the total energy entering a trophic level is lost—that is, nearly all of it becomes unavailable by the time organisms at the next level can access it. This phenomenon can be visualized with a graphical representation known as an **energy pyramid** (Figure 5.17).

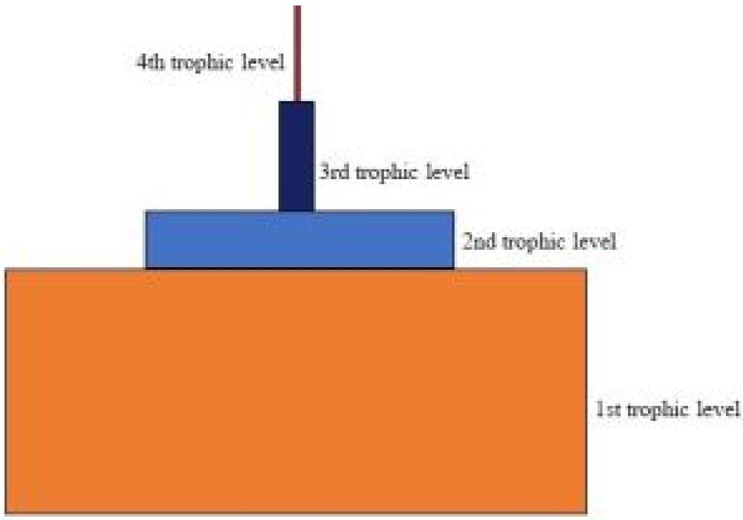


***Figure* 5.17.** *An energy pyramid. Each box represents the relative amounts of useable energy entering trophic levels. See main text for more.* Kelsey, CC BY-NC-SA.

Just where does this lost energy go? Two important processes divert energy away from consumers, bringing about crucial consequences to all ecosystems.

1. **High-grade energy is degraded to heat.** The laws of thermodynamics we first encountered in Chapter 4 are ever present, governing ecological and all systems. Organisms at each trophic level use the relatively high-grade chemical energy to which they have access for work such as reproduction and movement. In the process, that energy is degraded to heat, a form that is unusable by higher-level consumers. Picture, for example, the feeding relationship between a gazelle and a cheetah. The prey species is an herbivore, obtaining necessary nutrients by eating grasses. The predatory big cat gets what it needs by eating the gazelle, but it must expend a substantial amount of energy hunting and killing its prey. For its part, the gazelle utilizes stored energy to power all its functions, including running for its life when pursued by a cheetah.

A protracted chase will result in that much less available energy for the successful cat, and its net gain will be a function of how much both it and the gazelle expended. The prey may or may not elude its predator, but there is no escaping the second law!

2. **High-grade energy passes to decomposers.** In addition to the work and associated conversions described in the previous section, various other types of losses also remove usable energy from a particular organism and trophic level. As we saw in our discussions of ecosystem functions in the early parts of this chapter and the carbon cycle in Chapter 4, waste products released by organisms contain a lot of high-grade chemical energy. It is not just excrement, though: plants also lose branches, leaves, bark, and the like, and animals shed hair, skin, and other body parts during their normal cycles. Decomposers are adapted to extracting energy stored in these waste products and using it to enable the work they must do. Ultimately, low-grade heat energy is released as those fungi and bacteria grow and reproduce.

Declining available energy limits growth at high trophic levels

The dramatic drop in available energy just described greatly restricts the number of trophic levels present in any ecosystem. In fact, no more than approximately six consumer levels are possible in most food webs, with fewer seen in many situations. The amount of energy required to support yet another consumer level above the current apex predator simply exceeds that which is available to do so. Furthermore, the total amount of living material, termed **biomass,** in each level declines quickly and dramatically with distance from the original energy source (Figure 5.18).

As counter intuitive as it might seem, in terms of mass, there are far more primary producers than there are herbivores, more herbivores than primary carnivores, and so forth up to the apex, because of the massive energy inefficiencies described above. Yes, higher-level animals do have a higher *concentration* of stored energy in their tissues than do plants (that is, amount of energy per gram of flesh), but the *total* amounts of both mass and energy in a given trophic level are far lower than those in the levels beneath it.

Figure 5.18. *A biomass pyramid. Sizes of layers represent the relative amounts of living matter at each level. See the main text for details. Kelsey,* CC BY-NC-SA.

5.4. ECOSYSTEMS DEVELOP AND CHANGE IN RESPONSE TO STRESS

The final section of this chapter will address two fundamental questions about ecology we have yet to ask: first, how do ecological communities come into being and, second, what happens to an ecosystem that is damaged or otherwise subjected to stress?

5.4.1. Ecological succession changes ecosystem structure with time

Contrary to what might be our instincts or expectations, a system such as a mature forest did not simply emerge fully formed at the beginning of time and remain stagnant since. Instead, it and all ecosystems developed through **ecological succession**, a dynamic process by which the organisms dominating an area are replaced sequentially. Using a term we encountered earlier in this chapter, we could also say succession changes the *structure* of an ecosystem with time. These changes continue until a final (really long-lived) stage is reached (see Figure 5.19 for an overview).

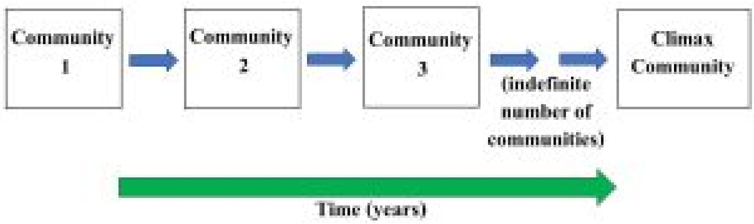


***Figure* 5.19.** *Schematic representation of ecological succession. As time progresses, the structure of the dominant community in an area changes.* *Kelsey*, CC BY-NC-SA.

Importantly, unlike the earlier communities, the climax is not easily replaced but persists until conditions change sufficiently to disrupt it (more below). We will see shortly that the concept of ecological succession is not nearly as straight forward as the idealized model described here. We devote space to it, though, because it is a fundamental idea in ecology that can be used to study and predict how ecological systems develop, change, and respond to stress.

Ecologists recognize two types of succession

The two types are distinguished on the basis of how much destruction precedes them. **Primary succession** follows a severe disturbance, one that obliterates a community. In other words, a new ecosystem is formed on a stark, barren landscape. Alternatively, **secondary succession** restores an existing community subjected to relatively minor damage. Here, recognizable remnants of the previous ecosystem are left behind, and succession can be viewed as a mechanism of recovery.

Primary succession. Many forces can produce barren environments, either by generating completely new habitat or by wiping out an existing community. A volcanic eruption, for example, is a powerful and transformative event; among other things, it can eject large amounts of lava onto Earth's surface (more in Chapter 7). This extremely hot molten rock tends to engulf anything in its path—forests, grasslands, houses and so forth—burning and burying it. Once the lava has cooled, it becomes solid rock, devoid of any life. Other natural causes of this level of destruction include melting of large glaciers that have covered an area for thousands of years, and collisions with asteroids. Humans can also clear land through agriculture (Chapter 9), construction, pollution (Chapter 15), and even warfare. We will devote some more attention to the ways human activity affects succession near the end of this chapter. Whatever the force, primary succession can follow it, changing an apparent wasteland into an area characterized by a complex, productive, and thriving community of organisms. You should keep in mind that this process is active in any location subject to severe changes, be they terrestrial or aquatic, hot or cold, wet or dry, etc. The upcoming section (along with Chapter 6) presents more information about the different types of ecosystems found in different places.

Secondary succession. Unlike primary, secondary succession begins on a landscape that still includes organisms and visible sources of nutrients and other vital means of support. Events such as forest fires, floods, earthquakes, landslides, and hurricanes can damage but not obliterate existing communities and are among the precursors to secondary succession. Here too, anthropogenic sources of stress, activities like tree cutting, various types of harvesting, and others, can set the stage for this type of ecosystem response.

More on the relationship between the succession types. Note that the two types of succession we just defined are distinct processes and should *not* be understood to be sequential. That is, despite what you might assume from their names, primary does not precede secondary. Instead, they each represent different paths by which an ecosystem can progress toward a climax community. Yes, the mechanism of replacement is integral to both, and it can be reasonable to even think of them as starting at different points on the same continuum. But given the many possible variables affecting the trajectory of succession, including the nature of the destructive event that came before it, it is appropriate to keep them separate from each other.

Some clarifications and context

Several important assumptions require our attention before we proceed. First, when studying succession, ecologists generally focus on the primary producers present in an area. In other words, first-trophic-level organisms, and how they change with time, are used to define different stages of succession. We expect that the identity of consumers at the second level and above will be a function of the producers present and will also change accordingly. So, in a land-based habitat, we would only keep track of plants and plant-like organisms. Second, primary producers attempt to procreate by releasing seeds (or analogous structures). Through various mechanisms, for instance on the wind, seeds are carried from their points of origin (he distances travelled can be enormous, hundreds or even thousands of kilometers). Usually, the seeds land in places inhospitable for them to grow. On occasion, though, they fall into exactly the right place at the right time, finding an environment that is well suited for them. Keep this second point in mind as we explore succession because it is important to remember that communities will most likely not be restricted by availability of organisms, rather by lack of necessary resources to support growth. Third, an area that exhibits little or no obvious growth and contains no visible sources of nutrients (e.g., no soil) is termed barren for the purposes of our discussion. Finally, succession should really be thought of a continuum, but for ease of study, we separate it into discrete stages as described here.

Succession involves many different species and steps

As noted, we will view ecological succession as a series of replacements. Earlier-stage species, are replaced by later ones that are themselves subsequently replaced. Unsurprisingly, a community that takes over an area does so because it is well suited to live there. What we might not expect, though, is that the actions of those successful organisms can change the environment so much that it becomes a better place for other organisms—members of a later-stage community—to live than it once was. In other words, earlier-stage organisms set the stage for their own departure because they often improve conditions for later-stage organisms that otherwise would not have been able to survive there. Changes continue until a stable end point is reached, what we termed the climax community. An important feature of this phenomenon is that the organisms dominating the different phases of succession possess different adaptations.

Pioneer species. These organisms begin the process of succession, so, in the idealized diagram in Figure 5.19 we saw above, they would make up Community 1. They colonize areas such as barren rocks, sand, ocean floor, or others that lack an obvious existing community. Although most species will fail to grow under such harsh conditions, pioneers possess traits that enable them to reproduce and take over seemingly unwelcoming landscapes. Notably, they devote an unusually large fraction of their energy to reproduction, dispersing large numbers of seeds (or the equivalent) far and wide. In a suitable environment, the seeds will quickly grow into mature forms, release their own reproductive structures to start the next generation, and die after a relatively short life. Their fast and intense existence requires an enormous amount of available energy, generally in the form of abundant, direct sunlight. The adjective “opportunistic” is often applied to them for they are well adapted to take advantage of habitats few others can tolerate.

Climax community. Although this is not the second but the final stage, we consider it immediately after pioneers because it represents the other endpoint in the process of succession and provides us with an opportunity to compare the traits common in these species to those found in pioneers. It is not quite this simple, but it is broadly fair to say that organisms in the climax community are everything pioneer species are not. Relative to **colonizers**, these species devote less energy to reproduction, grow slowly, live longer, and require less direct exposure to sunlight. Table 5.4 summaries the major differences between these two types of organisms.

Table 5.4. *A comparison of important traits of pioneer and climax species.*

TRAIT	PIONEERS	CLIMAX
Amount of energy devoted to reproduction	High	Low
Growth rate	High	Low
Length of life	Short	Long
Shade tolerance	Low	High
Likely to be replaced?	Yes	No

Early-, mid-, and late-stage communities. Now we consider the species that dominate between the two end points described above. Even as pioneers colonize an environment, many kinds of seeds continue to blow into the area. This time, though, since the conditions have been changed somewhat, a different group of organisms begins to appear. It has much in common with pioneers—need for direct sun, rapid growth, short lives—but it is not adapted to succeed on a completely barren landscape. In addition, these **early-stage successional species** will likely be larger than pioneers. As time progresses, the pioneers are outcompeted, and the character of the community changes. What happens to those newcomers? Their presence will likely facilitate the growth of additional new species, termed **mid-stage successional species**, that eventually take over the area. Regarding the important traits we have identified, they are intermediate between pioneers and climax organisms. More time brings further alterations away from the earliest conditions, allowing for the growth and dominance of increasingly long-lived, shade tolerant, and large primary producers. **Late-stage successional species** have more in common with those in the climax community than they do with pioneers but are themselves slowly pushed out by the species that are best adapted to this final, hypothetical environment. You should bear in mind that, beyond the changes noted in the survival strategies used by organisms, the community itself can take on different characteristics as well. Notably, although this trend is not universally seen, some researchers have observed how the total number of different species present is generally low in the earliest communities, increases through the middle stages, and then may decrease again at the climax (more in Chapter 6). If we think about the traits that are most advantageous at the various stages, we might predict this trend. The ability to colonize a hostile area requires specializations that are not common in the biosphere. We learned earlier, for example, that lichens can grow in very nutrient-poor settings. Consequently, they are among the few species able to act as pioneers. As challenges presented by an environment become less severe, there is a potential for a greater biodiversity because the necessary traits are less unusual. A decline in the total number of species in late stages could be explained by the fact that such a community is crowded, nutrients are scarce, and sunlight is not as available as it once was. Again, the conditions require a somewhat narrow and specialized set of adaptations. Despite the apparent logic in this reasoning, though, many ecosystems defy expectations. Clearly the many variables affecting community composition warrant additional study.

Succession affects all environments

Ecological succession surrounds us, controlling the way natural systems develop. The small sampling of examples included here demonstrates that, although the process is fundamentally the same wherever it is active, the details vary considerably as a function of prevailing environmental conditions. Furthermore, succession within a biome (described in an early part of this chapter) often follows a predictable pattern.

Temperate forest.

Succession can transform a landscape left barren by a volcanic eruption

Note that the following steps represent an idealized view of what is *possible*, not what is inevitable in this biome. That is, some variations from location to location are likely.

Bare rocks support little or no life (Figure 5.20). Whatever is living on this cooling field of volcanic rock is not immediately apparent.

Figure 5.20. *A lava flow incinerated the existing community, leaving behind a smoldering field of barren rock. Brocken Inaglory,* CC BY.

Lichens and mosses colonize the rocks (Figure 5.21). Seeds of many grasses, shrubs, and trees will land here, but due to the lack of soil and readily available nutrients, they will not grow. After a short time, the presence of pioneers starts to produce small amounts of soil. How does soil develop? Recall from our discussions of the carbon cycle and decomposition (Chapter 4 and early in this chapter) that some of the decomposed waste products and remains of organisms can accumulate in the lithosphere. In this case, they could fill small cracks in the rocks, fundamentally changing the character of this environment. We will see much more about soil formation in Chapter 9.

Figure 5.21. *Pioneers colonize barren rock. Frank Schulenburg,* CC BY-SA.

Grasses start to grow in the soil created from the remains of pioneers (Figure 5.22). The presence of even a small amount of soil, combined with the direct sunlight that would be available in the absence of any large organisms, allows these small organisms to quickly spread across the area.

Figure 5.22. *Grasses replace pioneer lichens and mosses. Kelsey*, CC BY-NC-SA.

Small shrubs, plants like milkweed, start to grow (Figure 5.23). The presence of the fast-growing grasses leads to the production of even more soil and a more welcoming habitat for slightly larger organisms.

Figure 5.23. *Small shrubs begin to appear. Kelsey,* CC BY-NC-SA.

Small trees like black walnut and birch can now grow (Figure 5.24). These early-stage organisms quickly mature and partially block some of the incoming sunlight. Soil and nutrients continue to accumulate.

Figure 5.24. *Small trees begin to appear. Kelsey, CC BY-NC-SA.*

Saplings of slow-growing trees such as maple and oak start to emerge (Figure 5.25). About this time, the early-stage trees from the previous step approach their maximum size and age and drop seeds to start the next generation. However, since the presence of shrubs and trees greatly reduces the availability of direct sunlight, seeds of the early-stage trees will struggle to grow.

Figure 5.25. *Trees get larger and shade increases. Kelsey,* CC BY-NC-SA.

The maple and oak trees get larger, reducing sunlight even further (Figure 5.26). The earliest-stage organisms are likely gone by now, as are many of the early and mid-stage trees. Seeds of the late-stage trees can grow in partial shade, and those species begin to take over the area. Often, such organisms are very straight and tall, growing densely enough that their crowns (i.e., the top sections containing the most leaves and branches) are in contact with each other. The resulting phenomenon is called a **canopy**, a largely continuous structure at the top of the forest. Like canopies built by humans, it blocks much of the incoming sun, creating a shady forest floor.

Figure 5.26. *Later-stage trees take over and shade increases even more. Kelsey*, CC BY-NC-SA.

The climax community made up of tall maple, oak, and similar trees persists (Figure 5.27). Under idealized circumstances, these species will dominate and not be subject to replacement. In reality, a climax forest may not last particularly long—a number of factors and forces can disrupt it (as stated above; we will see more below as well).

***Figure* 5.27.** *The climax forest. Kelsey*, CC BY-NC-SA.

How long does it take?

The time required for a forest to arise from a barren rock field varies, as you might expect. However, to provide a sense of scale, it would not be surprising if the entire process was completed within a century. In fact, the development of a field filled with grasses and small shrubs (review Figure 5.23) could occur after only a few years, depending on the specific conditions affecting the relevant environment.

Before we leave forest succession, we should recall one of the clarifications made earlier: succession is generally described according to the primary producers present in an area. We would expect the identity of animals to change from small insects, rodents, and birds in the early stages to larger mammals, larger birds, and small reptiles in the later stages. The climax forest could be habitat for deer, bears, lizards, owls, and so forth, animals that would be largely absent from open, grassy fields.

Sand dune.

Barren sand can become a forest

Although not formally identified as a biome like the temperate forest we just studied, a coastal sand dune is a distinctive area characterized by unique conditions and challenges for members of the biosphere attempting to live there. Through the following process it can undergo succession from a seemingly lifeless mound to a teeming, if small, forest.

A barren sand dune is an unwelcoming habitat for most (Figure 5.28). Organisms struggle to live in such an environment for two important reasons. First, sand is quite nutrient poor so supplies little of what organisms need to survive. Second, sand is not stable, but moves due to wind and precipitation. Most plants cannot survive in such a harsh habitat.

Figure 5.28. *A barren sand dune supports little obvious life. Famartin, CC BY-SA.*

Specially adapted plants withstand moving sands (Figure 5.29). Because they have extremely long roots—several meters is possible—certain grasses can colonize a dune. The roots anchor the plants and allow them to grow as the sand shifts. Even more importantly, though, the presence of the root network stabilizes the dune somewhat. Organic material also begins to accumulate as the remains of the grasses are decomposed.

Figure 5.29. *Beach grasses begin to colonize the dune. Kelsey*,

Other grasses and small shrubs begin to appear. Thanks to the actions of the pioneers, the dune is a more hospitable habitat than it was: it is now both comparatively stable and nutrient rich. The addition of organic material also increases the amount of water available for plants (we will see much more about the factors affecting soil fertility and water retention when we discuss agriculture in Chapter 9).

The climax community is dominated by small trees (Figure 5.30). Since sand dunes can be located on beaches and coastlines all over the world, the specific identity of climax trees found on them varies enormously. In any case, fast-growing trees such as beeches, pines, and birches are likely to make up the stable community in these environments.

Figure 5.30. *A good example of a climax community seen on a sand dune. It is quite different than that seen in the climax forest we saw in Figure 5.27, above. Kelsey,* CC BY-NC-SA.

Not everything can grow on a dune. As we know from several previous discussions (e.g., Chapters 1 and earlier in this chapter) organisms that are best adapted will dominate an area. So, although both are terrestrial systems, the temperate forest supports a different set of organisms than does the sand dune. The climax community near the coast must withstand salt, wind, intense storms, and other stresses that an inland community would not face. We will return to this point shortly.

Freshwater pond.

An aquatic system becomes terrestrial

Succession within this biome can lead to dramatic changes, shifting from a community dominated by algae, fish, and other water-based organisms to one dominated by trees, rodents, and other land dwellers.

Nutrient accumulation can cause eutrophication

We have seen this phenomenon before, albeit in a different context, in Chapter 4. If you recall how excess nitrogen and phosphorus entering a body of water can cause a pond to fill with sediment, you already know the mechanism by which freshwater succession occurs. Since this process was presented in detail earlier, we will not repeat it here. You should review it as necessary and add it to our list of examples (Figures 4.11 and 4.12)

Climax communities vary in structure

Previously we noted that the organisms present in the climax community on a sand dune are different from those in a forest. It is worth our attention to briefly revisit and expand on that idea here. Keep in mind that time is just one of the many variables affecting ecosystem structure. Moisture, temperature, chemistry, sunlight, and other properties also are very influential. Consequently, even if given thousands of years to develop, a sand dune will never support the organisms that flourish in a temperate forest because its conditions are simply inappropriate for such a community. Similarly, neither tundra nor desert environments will be able to support oak and maple trees. The same is true throughout Earth's ecosystems: dominant species vary enormously among biomes. Review Table 5.2, above, in this new context to see more examples (the heading on the column labelled "DEFINING ORGANISMS" could also be loosely taken to mean "organisms in a climax community").

Reality often defies predictions and expectations

The processes described here are idealized representations of what is likely to occur under natural conditions. In other words, early-stage organisms are replaced by later stage-organisms and so forth, but, for numerous reasons, the transitions and progression may not be as predictable or smooth as suggested.

Some organisms resist replacement. Organisms devote a tremendous amount of energy to reproduction and their own continuation. It stands to reason, therefore, that it is not in their best interest to facilitate competitors. Succession occurs because organisms possess different adaptations and abilities to survive under different conditions, it does not result from pioneers willingly or even knowingly serving their fellow members of the biosphere by improving conditions for them. Indeed, some species have developed ways to resist and interfere with later-stage species. A detailed description of the many possible mechanisms by which this can occur is unnecessary for our purposes, but generally speaking, some plants just make it difficult for others to establish themselves. One interesting example of this phenomenon is seen in the way certain grasses form a very thick mat, one that is quite challenging for a seed of a would-be replacer to penetrate. In short, the grasses shield all that newly formed and inviting soil from larger organisms that need it to grow.

Environmental conditions are not static. An ecosystem that is destroyed by a volcano or other event may not be exactly replicated again in the future, even if allowed sufficient time to do so. Environmental conditions like water availability and prevailing temperature can change, favoring different organisms. In addition, as suggested earlier, organisms not native to an area can outcompete those that once dominated. Such **invasive species** are so well adapted that they can quickly take over, profoundly changing ecosystem structure (more in Chapter 6). Other forces such as disease and nutrient limitations can also influence succession.

Humans affect succession in many ways. Many anthropogenic activities can alter, slow, or even stop ecological succession. A small sampling is provided here, and we will encounter these and other concerns in several places in this textbook, particularly in the chapters found in Part Three.

Ecosystems manipulated for homes, recreation, and sports

Back yards, public parks, golf courses, baseball fields, and similar areas tend to be carefully managed so they are appropriate for their designated uses. Often, they are dominated by grasses that are *never* replaced by later-stage species. How can this be? The answer is probably familiar to most people: extensive use of lawn mowing, fertilizer, watering, and pesticides help us control the structure of ecosystems (Figure 5.31).

Figure 5.31. *Lawn mowing maintains early-stage conditions indefinitely.* *Tobias Barkskog*, CC BY.

Since it is generally not in the interest of, say, a baseball team, to allow shrubs and ultimately tall trees (assuming the biome in question supports such organisms) to grow in their outfield, conditions are maintained so a grassy field persists indefinitely. In fact, each time you mow your lawn you are preventing succession from occurring—larger plants are simply not given a chance to establish themselves.

> *Try this: stop cutting the grass in your yard for a month or so and watch what appears. Oh, be sure to check the bylaws of your neighborhood association first, though.*

Agriculture

If you have ever noticed fields in which crops like corn or wheat are grown, you have seen one way humans stop succession from proceeding (Figure 5.32).

A climax forest might even be cleared from a plot of land, leaving only barren soil behind. Plowing and pesticides help inhibit anything other than the desired crop from growing. We will see much more about this important topic in Chapter 9.

Figure 5.32. *Climax forests are cleared and maintained in early stages of succession for farming. Pam Brophy, CC BY.*

Release of pollutants

Certain biological, chemical, and physical substances are products of industrial, agricultural, domestic (i.e., occur in homes), and military activities. Through numerous mechanisms, these toxins can bring about adverse effects on individuals, populations, communities, and ecosystems. In some cases, succession is stalled because early-stage species such as grasses are better adapted to survive the stress brought about by pollutants than are later-stage ones. One related phenomenon is known as an **ecological gradient**. In short, dominant species change with distance from a pollution source, providing an important indication that toxic substances are present in an area. Consult Chapter 15 for more on the ways pollutants can affect human and natural systems.

THE CHAPTER ESSENCE IN BRIEF [7]

Ecosystems are made up of the physical and biological components that interact with each other. A variety of interactions are possible and are necessary for organisms to obtain the materials and energy needed to survive and reproduce. Furthermore, ecosystems are not static, but they change with time in response to various stressors.

---

7. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 5 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[8]

How do organisms use energy? What happens to the energy after they have made use of it?

What would likely happen to heterotrophs (like humans) if autotrophs (like plants and algae) disappeared from Earth?

Is predation always bad for a prey population?

How does the use of antibiotics provide a selective advantage (we also might call it selective pressure) for pathogenic bacteria?

Why would an organism provide a benefit to another? Is that first organism simply generous?

Is it appropriate to refer to your body as an ecosystem?

Why do pioneer species improve conditions for later-stage organisms? Awfully nice of those pioneers, don't you think?

Given sufficient time, can ecological succession ever

8. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

convert a desert into a forest ecosystem dominated by oak and maple trees? Why or why not?

6. Biodiversity, Evolution, and Extinction

JASON KELSEY

Biodiversity is a term that gets heavy usage among people in and out of the sciences. Simply put, it is related to the number of different species on Earth. You can get an anecdotal, visual idea of biodiversity the next time you go into a forest or other natural ecosystem: take note of the myriad of sizes, forms, adaptations, and survival strategies around you. From different types of trees, to grasses, shrubs, birds, insects, rodents, and all the rest, there is a great deal of variety readily visible in the biosphere. In this chapter, we will see that biodiversity varies with space and time due to many factors, including the processes of evolution and extinction.

Key concepts

After reading Chapter 6, you should understand the following:

- How biodiversity can be measured and defined
- How the theory of biological evolution is linked to biodiversity
- The ways organisms are related to each other and classified

- How the endosymbiotic theory can be used to explain the development of complex life forms
- The evidence for and examples of biological evolution
- The many factors that can increase and decrease biodiversity
- The phenomena that contribute to extinction of species
- The role of human activity in causing and mitigating extinctions
- The relevance of biodiversity

Chapter contents

6.1. THE DEFINITIONS OF BIODIVERSITY

Several varied, although linked, definitions of biodiversity have been developed and are used in different situations. Here we begin with a brief introduction to three widely used terms.

6.1.1. Species richness

This property refers to the total number of different species alive in a defined area. The biodiversity of Earth as a whole, as well as that of its individual ecosystems, can be assessed using this concept.

How many species are on Earth?

There is no universally accepted answer to this challenging question and no easy way to go about counting all the species on this planet. Estimates provided by ecologists during the past few decades have typically ranged from 3 million to 100 million, although a recent study that tried to include vast numbers of yet-to-be-discovered **microorganisms** put it close to 1 trillion[1]. Clearly, given the size of the error term in this measurement (Chapter 2), more research will be needed to come up with a reasonable answer. When only **eukaryotes** are counted, the number declines substantially to a range of approximately 8 to 10 million[2]. In any case, most readily agree that the number of species yet to be tallied exceeds the number we have cataloged already.

1. Locey, KJ and Lennon, JT. 2016. Scaling laws predict global microbial diversity. Proceedings of the national academy of sciences. 113:5970–5975
2. Tittensor, MC, Adl S, Simpson AGB, Worm B. 2011. How Many Species Are There on Earth and in the Ocean? PLoS Biol 9(8): e1001127. doi:10.1371/journal.pbio.1001127. CC BY

How many species are in a given ecosystem?

Species richness can be used to assess and compare different ecosystems. The total number of species in a tropical rainforest, for instance, is higher than that in a desert. It is also higher near an ocean coral reef than it is on the portion of a sandy beach that is only covered during high tide (Figure 6.1).

Figure 6.1. *Biodiversity is both visibly and measurably higher in a tropical rainforest, left (a) than it is in a desert, right(b). Hjvannes, CC BY (a); US FWS, Public Domain (b).*

How broad or narrow are the niches in an ecosystem?

The relative size of the **niches** (Chapter 5) is a good proxy for species richness in any ecosystem.

Narrow. In this case, each species only carries out a small number of functions, lives in a very specific type of habitat, and is highly selective in choosing its nutrient sources. The presence of a large number of these **specialists** is a hallmark of high biodiversity.

Broad. Low-diversity environments are characterized by only a few organisms living in varied habitats and utilizing an assortment of different nutrient sources. In contrast to those in the preceding paragraph, low-diversity ecosystems are populated by **generalists**, species that are capable of a comparatively wide range of functions.

6.1.2. Species evenness

Species evenness is a measure of the relative number of different species present in an area, information that cannot be gleaned from an assessment of species richness. Put another way, evenness describes the balance among the species in an ecosystem. Keep in mind that richness and evenness do not depend on each other, that is, even if one is high, the other can be low. For example, a managed grassy field will likely have very low evenness because there are high numbers of only one or a few dominant plants, whereas an adjacent forested area with the same richness as the field could have high evenness, consisting of many individuals of multiple tree and shrub species.

6.1.3. Genetic diversity

A few words about genetics

Before we can consider how it can be used to quantify biodiversity, we need to take a brief look at the science of genetics.

Organisms are built from codes. Each living thing possesses a unique code in its cells. This code is stored in complex molecules known as deoxyribonucleic acid, or **DNA**. DNA molecules are long, yet microscopic, strands made up of shorter sections that are each called **genes**, and each gene controls the physical properties of a particular system. For example, different stretches of DNA contain the separate genes for human eye color, hair color, and skin color, as well as thousands of important traits related to growth, reproduction, and survival. The properties of all organisms are controlled by their DNA codes (also called their **genomes**), influencing which individuals are most likely to survive and thrive in their environments and which are not. Keep in mind that each time it needs to construct or repair a molecule or tissue an organism taps into the appropriate stretch of DNA. The term **genetics** broadly refers to all the information contained in DNA.

Appearances do not tell the entire story. Although all physical traits are a function of genetic sequences, not all data stored in DNA ends up translated into physical traits. Certainly the vast biodiversity on Earth results from differences in DNA, both small and large, but many organisms carry unused sequences, stretches of DNA that apparently are not used to construct any biological structure. Further complicating matters, some organisms having similar physical traits may have different genetic codes. In other words, it would be fair to say that genetic diversity exceeds physical diversity in the biosphere (particularly among bacteria, Chapter 3). Even members of a single species can differ substantially in their genetic sequences, despite appearances that might suggest otherwise. We will return to the distinction between physical form and genetic characteristics when we study biological evolution as well as how organisms are classified. This topic will also come up again in our discussion of **genetically modified organisms** (Chapter 9).

Genetic codes are passed from one generation to the next. Recall from Chapter 5 that reproduction is central to the continuation of living things. Although the details of how it is done vary from species to species, the general process is the same throughout the biosphere: mature individuals (we will informally call them parents) pass their DNA sequences to their offspring. For **prokaryotes** and other organisms that are **asexual**, a single cell will make a copy of its entire genome (termed DNA replication) before diving into two new cells. Under normal circumstances, each progeny cell is identical to its parent. In the case of sexual organisms such as humans, specialized reproductive cells are produced (called **gametes**, e.g., sperm and egg) which allow each separate parent to contribute half of the genetic code to an offspring. As a result, the new individual will have a complete genome that codes for a novel combination of traits; accordingly, it will resemble, not be identical to, its parents.

The meaning of genetic diversity

Now that we have gained the appropriate background, we are ready to consider how genetic sequences are used to characterize biodiversity of a population in an ecosystem. Genetic codes of individuals can be studied (it is no easy task to do this, but it is possible if the proper tools are available), and the degree of variation *within a species* determined (**not** the total diversity across multiple species). High genetic diversity would be indicated if a relatively wide variety of genetic sequences is detected. Many forces likely affect genetic diversity of a **population**. It is sometimes reduced in environments characterized by long-term stability because there is no pressure to adapt to new conditions on a regular basis—only a limited number of strategies are needed. Put another way, systems that have been subjected to a substantial amount of change and stress (from human and natural processes—Chapter 7) tend to display higher diversity within populations than those that are stable. As we will see below, ecologists tend to view high genetic diversity as a sign that a population is resilient and likely to successfully adapt to changing environmental conditions. These rules do not apply universally, and additional research is needed to better understand the factors influencing genetic diversity.

6.2. THE ORIGINS OF BIODIVERSITY

No matter how it may be measured and expressed, we can observe that Earth is populated by a wide range of organisms. What is less immediately obvious, though, is just *how* all that biodiversity came to be. This is not a new conundrum, for people have been struggling with it for millennia, although ideas and answers offered have changed with time and study.

6.2.1. The theory of biological evolution

This theory is the prevailing scientific explanation for Earth's current biodiversity, offering an evidence-based mechanism by which the biosphere has developed. Its modern form started to take shape in the 1750s, but it has changed a great deal since that time. Like other **theories** (Chapter 2), it has been influenced by generations of researchers making objective observations, proposing **hypotheses**, taking measurements, and drawing conclusions. It has withstood over two centuries of scrutiny and continues to serve as a foundation for our understanding of life on this planet. Here we will explore the many principles and ideas comprising the theory of evolution.

Change with time

The term **evolution** refers generally to change during some time frame and can be applied to a wide range of entities, ideas, principles, relationships, and feelings. In short, anything that is dynamic and responsive to stimuli has the potential to evolve. Now, although most people understand the word to imply an improvement from a previous state, it is important to recognize that not all change is for the better, particularly when it comes to organisms. Yes, biological evolution is often characterized by changes that enhance a species' ability to survive, but as we will see shortly, the story is more nuanced than that.

Processes affecting the evolution of organisms

The change in the inherited characteristics of a population is the essence of **biological evolution**. Put another way, modified genes and the physical traits controlled by them are passed between generations. Very importantly, evolution does *not* refer to or imply changes that occur during the lifetime of a single organism. Several processes interact and influence biological evolution.

Mutation. Recall that DNA contains the codes—the blueprint—that can be used to produce offspring. One of the assumptions underlying reproduction (both asexual and sexual) is that the complex genetic code stored in parents is faithfully copied before being passed to progeny cells. However, under many circumstances, replication is not perfect, and newly constructed DNA is different than the parental template from which it was built. In other words, the new DNA sequence contains errors or **mutations**.

Causes of mutations

DNA sequences can be changed via a number of mechanisms, the details of which are unnecessary in this discussion. Two broad types are very briefly described here.

External agents. As is known by many people, exposure to certain types of radiation, some toxic chemicals, and a small number of diseases can bring on mutations. We will see more about the effects of these types of agents on living systems in Chapter 15.

Endogenous errors. Mutations can also occur in the absence of any kind of external stressor. A phenomenon known as **normal error** can lead to changes in the DNA code during replication. The body can tap into several mechanisms to repair most errors before they cause trouble. Some mutations occur as a matter of course, though, and some of those progress into substantive outcomes for an organism.

Potential outcomes of mutations

Many outcomes are possible after the code has been altered.

Death of the offspring. If vital physical structures encoded by the mutated gene are affected, the organism will die.

Damage to the offspring. A wide variety of adverse consequences is possible. A non-lethal mutation often makes an organism less fit for survival and greatly reduces the likelihood it will reproduce.

No apparent changes. We know that for many organisms, in addition to genes employed to produce cells, tissues, and other structures, DNA contains information that is not translated into any physical traits (noted earlier). A mutation in one of those unused regions may lead to no obvious consequences.

Increased fitness. Although this outcome is rare, mutations can alter genes in ways that actually improve an organism's ability to survive in its environment. Like those that damage an offspring, many specific mechanisms by which fitness is improved are possible. For example, as we saw in Chapter 5, DNA mutations can lead to antibiotic-resistant bacteria. An individual with an altered genome has the capacity to produce a new physical structure that enables it to survive exposure to penicillin or related compounds. It therefore has a great advantage over the cells around it that lack the mutation, and its offspring will dominate future populations (creating a public-health menace). Keep in mind that any kind of mutation increases genetic diversity of a population: more mutations lead to a greater variety of genes and possible physical traits in a group. Often, the total collection of genetic information stored in DNA of all the members of a population is referred to as a **gene pool**. In principle, as its gene pool gets larger and more diverse, a population becomes more and more adaptable, that is, able to survive under a wider and wider range of conditions. It follows, then, that a gene pool may contain many unused genes at any one time. Even a potentially beneficial mutation will not lead to an advantage until and unless environmental conditions change in such a way that the new traits provide a benefit to individuals possessing them. So, a bacterial cell that can withstand the toxic effects of, say, five different antibiotics, will only have increased success in an environment that contains those antibiotics. Many other examples of this phenomenon have been observed, as we will see shortly. We will also see that mutation is a powerful, but not the only, force contributing to evolution.

Recombination. Even in the absence of mutations, new genetic sequences can arise through this process. Put broadly, existing sections of DNA can be mixed together to create offspring with novel codes and traits.

Mechanism of recombination in two types of reproduction

Asexual. Individual microorganisms divide as described previously, providing a complete copy of their DNA to progeny cells (Figure 6.2a). Recall that interactions among different cells are not needed for reproduction of these organisms. However, cells still can exchange portions of their genomes under some conditions. Even though these interactions do not support reproduction, they do increase the diversity of the gene pool. Returning to the example we used above, microbiologists have noted how genes imparting antibiotic resistance can get passed to offspring as expected, but also among cells *not* engaged in reproduction. This transfer of DNA from one bacterium to another can occur both within and across species boundaries, further magnifying the scope of the problem for humans. Although seen rarely, some plants and animals can reproduce asexually as well.

Sexual. Each parent contributes to the code that ultimately makes up the genome of their progeny. Randomness in the selection of which portions of DNA actually move into sperm and egg, along with some other processes active during the formation of reproductive cells, lead to the construction of genetic sequences in the offspring that did not exist in either parent (Figure 6.2b).

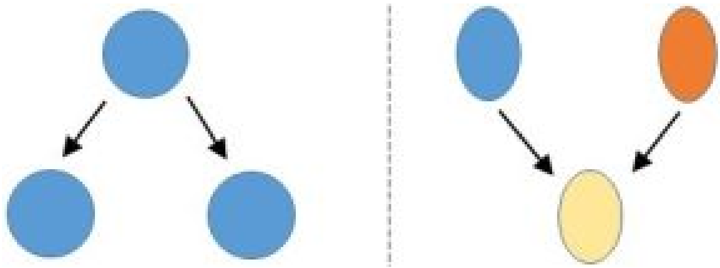


Figure 6.2. *Asexual reproduction, left (a), requires an exact copy of the entire genome of the parent cell and produces two identical progeny cells. Sexual reproduction, right (b), requires each of two parents to contribute half the total amount of genetic material to a single offspring that has a new set of traits. Kelsey, CC BY-NC-SA.*

In general terms, recombination increases the size of the gene pool of a population. So, like mutation, it leads to gene sequences that could provide an advantage or disadvantage to affected individuals.

Natural selection. People often equate biological evolution with this process, but they are not the same thing. As we learned, evolution is the change in inherited traits in a population. Natural selection, on the other hand, is one of the processes *influencing* which traits are passed on from parents to offspring.

Environments shape organisms

Natural selection is the way individuals of a species that are best adapted to their environments tend to succeed and reproduce at higher rates than less-well-adapted individuals. It can be informally viewed as a process that is able to sort out and rank individuals according to their fitness to survive. We have encountered this idea before. Recall from Chapter 1 that one of the fundamental concepts in environmental science is that organisms are tightly linked to and influenced by their environments. Pressures exerted by various conditions such as temperature, terrain, availability of food, competition and so forth shape organisms. Those possessing the most advantageous traits are most likely to pass their genes (and the resulting physical characteristics) into the next generation. Charles Darwin famously studied and proposed the basis for natural selection in the middle part of the 1800s, and his ideas, although revised and built upon by later scientists, remain central to the theory of evolution today.

Many factors influence natural selection

Since organisms are influenced by environmental stressors, it stands to reason that the traits providing advantages change as conditions change. In fact, what is beneficial to organisms at one point in time may not be sufficient to ensure survival at another. Several forces can change environmental conditions and therefore, the traits that are most helpful.

Hazards. Fires, earthquakes, volcanic eruptions, floods, and other events (Chapter 7) can quickly alter an environment, favoring individuals that are able to survive the immediate and long-term threats.

Slow changes. Environments are generally subject to gradual changes in temperature, moisture, and the like during long stretches of time: centuries, millennia, or more. Populations living under such conditions will need to adapt at a similarly slow pace.

Migration. Instead of enduring changes affecting their native habitat, a population may need to move to a completely different place. Drought, decrease in food supply, natural disaster, human activity, and competition are among the common stressors that can induce migration. Whatever the cause, pressure exerted by even subtle changes in environmental conditions in the new location will redefine which traits provide the biggest advantages to individuals that survived the journey.

Random chance. The phrase “survival of the fittest” is frequently used in reference to biological evolution and even has been known to make its way into social and business settings. It is founded on the principle that members of a population best adapted to their environments will necessarily be most successful, and weaker ones will fail. Surely natural selection does favor the best adapted individuals, but fitness alone is no guarantee of success. Under some circumstances, individuals *least* well adapted could be the ones passing their genes to the next generation. How could this anti-intuitive (and possibly upsetting) outcome occur? Random chance. Imagine, for example, a population of terrestrial animals living on an island. Because its favored food is most abundant in the surf zone, the water's edge is the best place to live. Consequently, the dominant (i.e., most fit) individuals live closet to the ocean, and the weaker ones are forced to live inland on higher ground. Everything might unfold as we expect for a thousand years, with the near-shore organisms breeding most successfully and the inland organisms struggling to survive. However, one day a giant wave appears on the shore, instantly killing the members of the population gathered in the water. That random event gives an unexpected boost to the portion of the population safely living on top of the otherwise inhospitable mountain. These weaker individuals discover the shore is now open to them, so they come down to feed and breed. Fitness still plays a role in who succeeds and who fails, but this sudden change in the content of the gene pool deflects evolution away from the path we likely anticipate. We will explore inevitability (rather, lack thereof) further in the section below.

Important considerations and cautions

The theory of evolution is complicated and often misunderstood. Here we consider and clarify some important ideas, principles, and methods before we proceed further.

Many variables affect evolution. Biological evolution results from the interplay among the several processes described above. Specifically, mutation and recombination create new genetic sequences, potentially increasing the size of the gene pool, and natural selection and random chance dictate which genes dominate within a population.

Evolution involves heritable changes. As we saw earlier, biological evolution only affects characteristics that can be passed from one generation to another, not those that are acquired during an individual's life. So, traits like eye color, hair color, height and so forth are heritable, but ability to play the piano or read a book or the loss of an ear due to an accident are not. A critical experiment conducted more than a hundred years ago helped to uncover this principle. In short, the tails were cut off rats after their births. Then, these altered rats were allowed to breed. Despite the condition of their parents, offspring were born with normal, complete tails. Even after many generations of similarly treated rats were tested (all unfortunate enough to have been relieved of their tails by researchers), newborns continued to possess tails. Scientists accept what seems like the most reasonable conclusion from these data: only the unaltered genes for tail formation were inherited, not the mutilation that occurred during the lifetime of parents.

Evolution produces new species. Interactions between modified genes and changing environmental conditions can ultimately produce a new species (a process known as **speciation**). Recall our understanding of the species concept (Chapter 5), to see an important characteristic of evolution: a group of better-adapted individuals that are *unable to breed* with members of the existing population can appear. This new species will likely become dominant because it does not mix its genes with less-fit individuals.

Biological evolution only applies to living things. Because it depends on genetics and reproduction, this form of evolution is unique to organisms.

Biological evolution is not predictable. Due to the role random chance plays in the process, as we saw above, it is impossible to predict the outcome of biological evolution. It might be explainable *after the fact* if sufficient evidence is available, but the nature of its output (what we might informally call its final product, assuming that evolution ever reached a conclusion) is by no means inevitable.

Relatedness is a function of common ancestry. You should realize that modern organisms are related to each other because of their shared ancestry, a relationship shown through a branching **phylogenetic tree**, *not* because one turned into the other (Figure 6.3.a and 6.3.b). For example, a widely held misconception is that humans evolved from monkeys. It turns out that all great apes (a group that includes us) are thought to have arisen from a single predecessor species that lived some 40 million years ago (Figures 6.3.c, 6.3.d., 6.3.e). Similarly, and perhaps surprisingly, modern hippopotamuses and whales also share a common ancestor. Again, hippos did not *become* whales, but two separate lines of descent branched from a single (now extinct) mammal about 54 million years ago. We will return to this topic later when we consider some of the evidence and data used to determine evolutionary relationships among organisms.

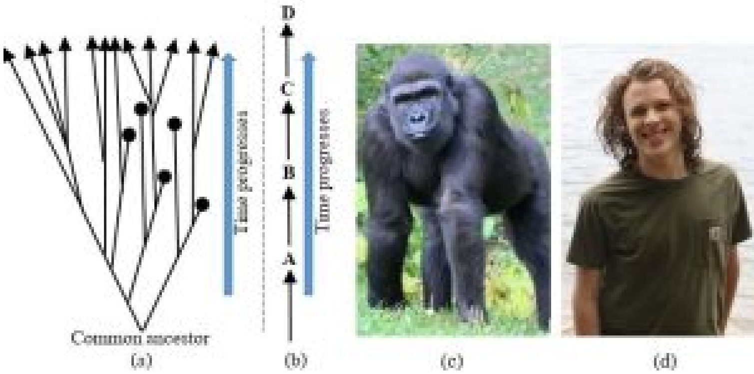


Figure 6.3 (parts a — d). *The relationships among organisms and descent from a common ancestor is visualized in modern evolutionary theory with a phylogenetic tree, above extreme left (a). Evolution proceeded from bottom to top, with each line representing a separate species. Closed dots denote a species that went extinct in the past. The transformation from one currently existing species to another, second panel from above left (b), is an inaccurate interpretation of evolutionary theory. In this incorrect second model, organism* B*—say, a monkey—is shown to become* C*—a gorilla, which then becomes* D*—a human. It is a complex idea and misunderstanding persists. So, to reiterate: evolutionary theory does* NOT *hold that gorillas (c) evolved into the humans (d). Both of these organisms are still on Earth because they proceeded along different branches from* THE SAME STARTING POINT. *Kelsey,* CC BY-NC-SA *(a and b); Rufus46,* CC BY-SA *(c); Kelsey* CC BY-NC-SA *(d). Figure 6.3 is continued below.*

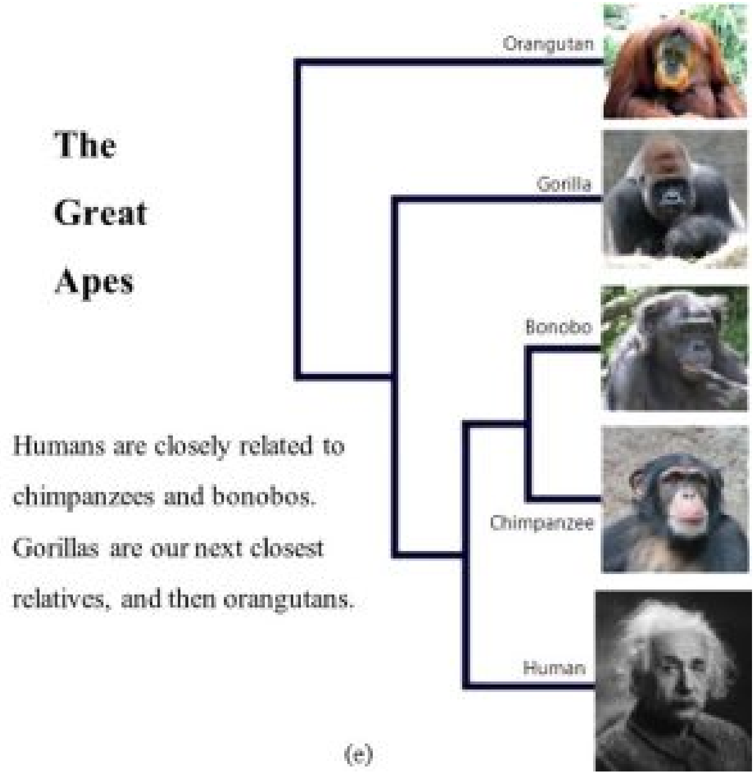


Figure 6.3 continued (part e). *All of the species shown here are still on Earth because they evolved from the same ancestor (e)! Merrilydancingape, photos in this figure were found on Wikimedia Commons and mashed together, Orangutan from Kabir Bakie, gorilla from Kilarin, Bonobo from Ltshears, Chimpanzee from Thomas Lersch, Einstein from public domain (photographer Oren Jack Turner, edited by Jaakobou),* CC BY-SA.

Scientists compare organisms to assess relatedness. Comparisons among organisms have long been used by evolutionary biologists to piece together the history of the biosphere. Similarities are assumed to be evidence of relatedness: close relatives ought to share more characteristics than distant ones (we will see it gets more complicated than that). Scientists have used two strategies to assess relatedness.

Physical traits

Historically, biologists compared the anatomy of different organisms to both determine how related they are to one another and to assign them to groups. One widely used organizational scheme divides the biosphere into five broad categories, or **kingdoms**: Animals, Plants, Fungi, Protists (which includes protozoa and some other eukaryotic microorganisms) and Monera (all bacteria). Those kingdoms are further divided into more specific groups. An underlying assumption is that organisms sharing many physical traits must have diverged from a common ancestor relatively recently. It is not always so simple, though. Firstly, similar physical traits do not guarantee close relatedness. Take, for example, two members of the animal kingdom, killer whales and great white sharks. In many ways, they look quite similar. But despite their shared characteristics, they arose from very different ancestors and should not be grouped together (we will see more about this example below, including images of the animals, in the section on convergent evolution). Secondly, it is not particularly straight forward to assess similarities between organisms lacking equivalent structures. How does one compare a barracuda to an octopus or a pine tree to a wolf, to cite just two confounding examples? What should be used as a point of reference? During the past century or so researchers have proposed and developed multiple approaches to improve comparisons, with varying degrees of success. Despite their efforts, the fact remains that an examination of physical traits only provides a limited amount of information about ancestry, relatedness, and evolution. As described in the next paragraph, modern biologists have been able to add a powerful tool to their studies.

Genetic sequences

Relatively new technology has enabled us to determine the genetic sequences of different organisms and, by so doing, make several discoveries. Notably, many of the genetic sequences within the gene pools of populations are conserved, that is, they remain largely unchanged for many generations. Even more importantly, it is now clear that a number of the same DNA sequences can be found throughout the biosphere, even in organisms that do not resemble each other physically. How is this information useful? As we learned earlier, genetic codes are passed from parents to offspring. In other words, the DNA in one organism is directly related to that in its ancestors. We also know, though, that sequences can change for various reasons during reproduction, leading to evolution as time passes. So, a scheme has been developed that uses comparisons of genetic codes to establish relatedness; new groups based on this scheme have emerged as well. The system assigns organisms to one of three broad **domains** (Figure 6.4): **Eukarya** (all eukaryotes), **Bacteria** (a group of prokaryotes that inhabit mostly familiar habitats), or **Archaea** (a group of prokaryotes that can be found in unusual, even extreme, habitats). Subdivisions within each domain have also been made, for example, within Eukarya are several kingdoms that each contain animals, plants, and many others.

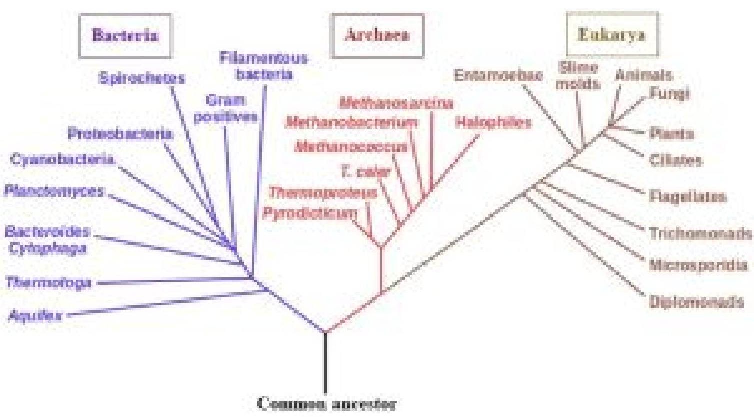


Figure 6.4. *The Three Domain System of classification based on genetic relatedness.* US NASA, *Public Domain.*

Determining relatedness through genetics provides many advantages, including the ability to study organisms with vastly different body types—the barracuda and octopus, again—and objectively compare and group them according to how well their DNA sequences match. It also bypasses potentially misleading information gleaned from studies of anatomy and focuses on a more reliable and relevant indicator of ancestry. As a result of our new understanding, older groupings based on physical traits have been revised to better reflect evolutionary relationships. Many instances of this reorganization could be cited, but we will only consider two of them here.

Monera get two domains. Because of their physiological similarities, traditional methods assigned all bacteria to a single kingdom, Monera. As we noted in Chapter 3, though, despite their somewhat limited variability in form, bacteria as a group possess an enormous amount of variety in life strategies and genetics. So, prokaryotes are now separated into two separate domains, Bacteria and Archaea.

People and fungi have a lot in common! It might be hard to imagine when you look at them side by side, but genetic analysis reveals how animals and fungi are closely related. Study Figure 6.4 again for more potential surprises.

Note that, despite lingering questions about some of its assumptions and methodologies, classification that uses gene sequencing is widely (if not universally) accepted among today's scientists. Furthermore, most still consult *both* physical and genetic data in their studies of evolutionary relationships. See Box 6.1 for a bit more about the categorization of organisms.

Box 6.1. Where do we fit in?

Like all organisms, the species formally known as *homo sapiens* can be categorized using physical and genetic characteristics. Humans, along with chimpanzees, turtles, ants, trees, and fungi are all members of the same broad domain, Eukarya (Figure 6.4, above). Of that list, all but trees and fungi are part of the less general group, kingdom Animalia. The animals are further separated into subgroups (in increasing order of specificity, their phylum, class, order, family, genus and species). Chimpanzees and humans are similar enough to share the same family, but each is assigned to a different genus. Modern humans are the only living members of the genus *Homo*—all the rest are extinct relatives of ours such as *Homo neanderthalis* and *Homo erectus*. A final note is appropriate here. By convention, organisms are referred to by their genus and species names. In Chapter 5, for instance, we encountered the competitors *Paramecium caudatum and Paramecium auerlia*, so named because they are different species in the same genus.

Evolution of the biosphere requires a great deal of time. With few exceptions, including bacteria (Chapter 3) and insects (Chapter 9), evolution is very slow. Earth's extreme age (Chapter 3), has allowed enough time for this gradual process to shape the biosphere

Biological evolution on Earth: some highlights

Earth is thought to be very old, forming about 4.6 billion years ago, and evidence of life dates back some 3.8 billion years (review Chapter 3, including Table 3.1). Clearly, it would be quite an undertaking to describe the entire history of the biosphere, so we will instead focus on just a few important highlights.

Biological molecules form (prior to 3.8 billion years ago). For living cells to develop, organic building blocks had to be available in the primordial oceans. Where did they come from? Evidence found in rocks and **fossils** suggests one commonly held answer. Briefly, Earth's early atmosphere was very different than the one we observe today. There was much more methane, ammonia, and hydrogen cyanide, and no dioxygen, in the early days. These chemical differences matter because researchers have been able to observe the spontaneous formation of amino acids and other biologically relevant compounds from inorganic precursors under conditions thought to mimic those of ancient Earth. Other answers, including an extraterrestrial origin of biomolecules, continue to be studied and debated.

The first living cells arise (approximately 3.8 billion years ago).

The building blocks noted above started to cluster and organize

Recent experiments have demonstrated that basic cell components such as membranes and genetic material (Chapter 3 and the current chapter) are self replicating (i.e., can make copies of themselves), providing evidence that primitive units possessing many properties of living things, or **proto-cells,** could have given rise to the biosphere.

The common ancestor emerges

One of the underlying principles of evolutionary theory is that all organisms share the same origin. That is, modern scientists hold that the life-like entities noted in the previous paragraph became recognizable living cells and ultimately evolved into the diverse biosphere that has existed up through today. Put simply, every living thing on Earth can trace its history back to a common ancestor (again, review Figure 6.4).

The first cells were relatively simple

Evidence strongly suggests that the earliest organisms were single-celled, **prokaryotic anaerobes** that used the sun as their energy source (i.e., they were **phototrophs**).

Photosynthesis begins (approximately 2.5 billion years ago).

Early organisms started to evolve and diversify into many species of prokaryotes. One of the most important changes occurred when certain organisms began to use the sun's energy to fix their own carbon. Three crucial consequences followed.

Changes in atmospheric chemistry

Dioxygen (O_2 gas), one of the products of photosynthesis, began to slowly accumulate in the atmosphere. By about 700 million years ago, the concentration of the gas had reached that of today, approximately 20% O_2.

Changes in dominant organisms

We know that environmental conditions shape organisms. In this case, the release of gaseous oxygen by photosynthesis made the planet toxic to the anaerobic organisms that had dominated for so long. Anaerobes had to move into new environments that lacked oxygen, and the ability to use O_2 became a substantial advantage everywhere else (review Box 1.1).

Formation of an ozone layer

Ozone is a gas made up of three oxygen atoms bonded together (instead of the two in the gas needed by aerobes). As we will see in much more detail in Chapter 14, its presence in the stratosphere (Chapter 4) protects organisms from much of the damaging solar radiation striking the planet. Without atmospheric ozone, organisms could not live in terrestrial environments. Indeed, life on Earth was restricted to aquatic environments before O_3 appeared (water shields organisms from solar radiation). What does this have to do with photosynthesis? Well, ozone is formed through reactions involving dioxygen. So in the days before free O_2 gas was present—prior to photosynthesis—there was no ozone layer and no terrestrial organisms. Again we see the mutual influences between Earth and life on Earth!

Eukaryotic organisms develop, and diversity of the biosphere increases (approximately 1.5 billion years ago). A quick look around suggests life on Earth did not stay restricted to single-celled prokaryotic forms. So, where did all these eukaryotes in your house come from? The short answer to the question is known as the **endosymbiotic theory**. There is a great deal of evidence to suggest that modern-day eukaryotes got their start when one ancient prokaryote ingested a relatively small prokaryote. Instead of being digested, this second cell was able to live **symbiotically** inside the larger one. The two organisms coexisted, and even provided benefits to each other. For example, the smaller one would have converted food ingested by the larger one into forms of energy both entities could use. They continued to reproduce independently, with one or more copies of the smaller cell being passed to subsequent generations of the larger one. Eventually, the descendants of the original cells became dependent on one another, setting up a permanent association that persists to this day. In other words, the organelles like **mitochondria** (Chapter 3) that are found inside of modern eukaryotes evolved from ancient, independent prokaryotes (Figure 6.5).

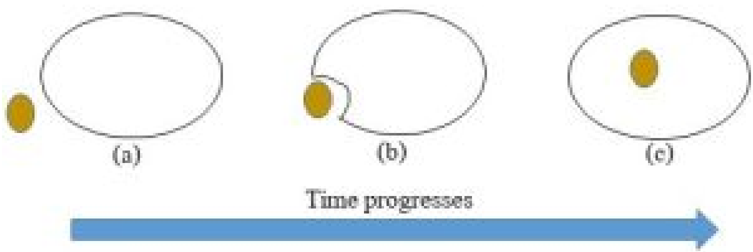


Figure 6.5. *A simple representation of the endosymbiotic theory. The sequence began when a larger cell engulfed a smaller one as shown by the circles on the left (a) and in the middle parts of the diagram (b). Instead of becoming food, the smaller cell established a mutually beneficial symbiotic relationship with the larger one, eventually becoming an organelle, right diagram (c). See the main text for more details. Kelsey, CC BY-NC-SA.*

Although some people are uncomfortable with this theory, abundant evidence in support of it, including the six points listed below, has led to its wide acceptance among scientists.

1. Mitochondria are the same size as modern bacteria

Clearly this point alone is not enough to make the case, but it allows the idea to be plausible.

2. Mitochondria have their own DNA

The genetic information used to make copies of these organelles is separate from that of the main cell.

3. Mitochondrial and bacterial DNA are similar

As we saw briefly above, it is possible to determine the specific code of any strand of DNA. The results of such work are striking: sequences found in mitochondria have more in common with those in modern bacteria than they do with those in their own main cell.

4. Mitochondria reproduce independently of the main cell

These organelles copy their DNA and divide into offspring at their own pace, not necessarily at the same time the main cell divides.

5. Mitochondria and bacteria use the same type of machinery to build proteins

All organisms can construct proteins using specialized structures known as ribosomes. Although they have similar functions, prokaryotic and eukaryotic ribosomes differ from each other. Importantly, mitochondria possess *prokaryotic* ribosomes even though they are found inside *eukaryotic* cells.

6. Mitochondria can be sensitive to antibiotics

Recall that bacterial infections may be treated through the administration of chemical compounds that are selectively toxic to prokaryotes. Antibiotics work because they generally do not harm human hosts. However, mitochondria inside human cells are sometimes affected, again suggesting they are more closely related to prokaryotes than eukaryotes.

Multicellular organisms appear (approximately 700 million years ago). Up until this time, organisms were limited to unicellular or colonial lives (Chapter 3). With the addition of oxygen to the atmosphere, though, the stage was set for larger and more complex organisms to evolve. Through their use of O_2, **aerobes** are able to extract far more energy from their food than are **anaerobes**. Since maintaining a large and complex network of cells (such as those in a plant or animal) is very energy intensive, multicellularity would have been far less probable if Earth had continued on as an exclusively anaerobic planet. The development of mitochondria and other organelles further facilitated the addition of multicellular organisms to the biosphere (for various reasons, including their ability to efficiently process energy sources). Extensive studies of fossils from different time periods have found a sudden (geologically speaking, of course) expansion of biological diversity about 570 million years ago, shortly after multicellular organisms first appeared (Table 3.1 in Chapter 3).

Organisms appear and disappear (ongoing). The history of life on Earth could be characterized as a series of trials and errors. Evolution can ultimately lead to the development of new species that persist for varying lengths of time. In some cases, a species is considered to be a dead end, a collection of traits and adaptations that did not provide meaningful long-term advantages. Some scientists use phrases like "failed experiment" to describe such organisms. However, time and again, even resilient, and seemingly permanent populations dwindle in size and become extinct; in the end, all that is left of them is fossil evidence (we will see more about the causes and consequences of extinction later in this chapter). Arguably, the story of dinosaurs is the most well-known example of this phenomenon. These animals appeared on Earth about 230 million years ago and dominated air, land, and seas for some 165 million years. Then, for reasons that are not entirely understood (likely a collision with a giant asteroid), they ceased to exist. Most believe an ability to adapt to rapid changes in environmental conditions was at least partially responsible, but the larger point is: species are temporary. As biological evolution shapes and expands biodiversity, extinction limits the number of different species present, as we will see in more detail shortly (see Box 6.2 for a bit of perspective on dinosaur history).

Box 6.2. These critters are OLD.

It's time for another fun fact about dinosaurs and the vastness of history! The oldest dinosaurs (they arrived on the scene about 215 million years ago) are separated in time from the youngest dinosaurs (they went extinct about 65 million years ago) by MORE years than the youngest dinosaurs are separated from humans (appearing about 2 million years ago). You are encouraged to review Chapter 3 for more about geologic time.

Evidence in support of the theory of biological evolution

Much evidence has been collected to support this theory in the centuries it has been studied and scrutinized.

Increasing complexity in fossils. Multiple examinations of fossils have revealed something extremely important: relative simplicity of organisms increases with age of rocks studied. Study after study has found the remains of ancient organisms to get progressively more complex as one moves from oldest to youngest rocks. This trend broadly supports the timeline of evolution presented in the previous section (Figure 6.6).

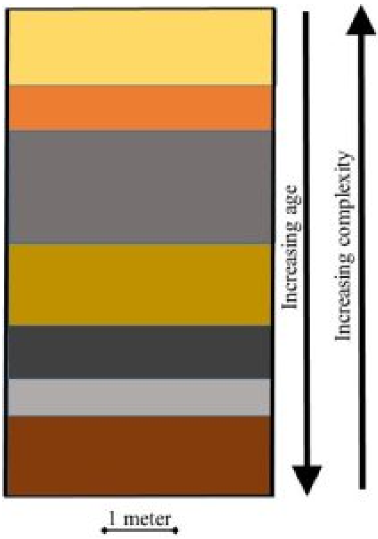


Figure 6.6. *A simplified diagram showing a hypothetical sequence of rocks and the trends in complexity of fossils contained in them: organisms become increasingly complex as age of rock decreases. In cross section. Kelsey,* CC BY-NC-SA.

The universal genetic code. Despite the vast diversity in physical form found in the biosphere, all organisms employ the same basic strategy to store, process, and make use of genetic information. Among other commonalities, every strand of DNA is constructed of the same building blocks, whether it is in the cells of animals, plants, fungi, algae, or bacteria. Scientists view this to be a strong indication that all organisms originated from the same ancestor.

Similar developmental stages in separate organisms. The science of embryology, which studies the development of unborn (or unhatched) organisms, provides some interesting evidence in support of evolutionary theory and a common ancestor. In short, the embryos of related, but distinct, species are very similar, that is, they go through the same developmental stages. For example, a large number of mammal embryos possess gill slits and tails. Humans and other terrestrial organisms lose these structures before they are born, whereas aquatic organisms like fish retain them after they emerge (live birth or from eggs).

Homologous anatomical structures. Distinct organisms living in vastly different environments possess the same basic body plan and bone structure. For example, modern birds, lizards, rabbits, frogs, and humans, as well as some extinct organisms, all have the same three bones in their forelimbs, namely, the humerus, radius, ulna. Of course, the structures have been modified by evolution, but the data suggest these (and many other) animals descended from the same ancestor to have the same basic layout in their arms, wings, legs, and fins (Figure 6.7).

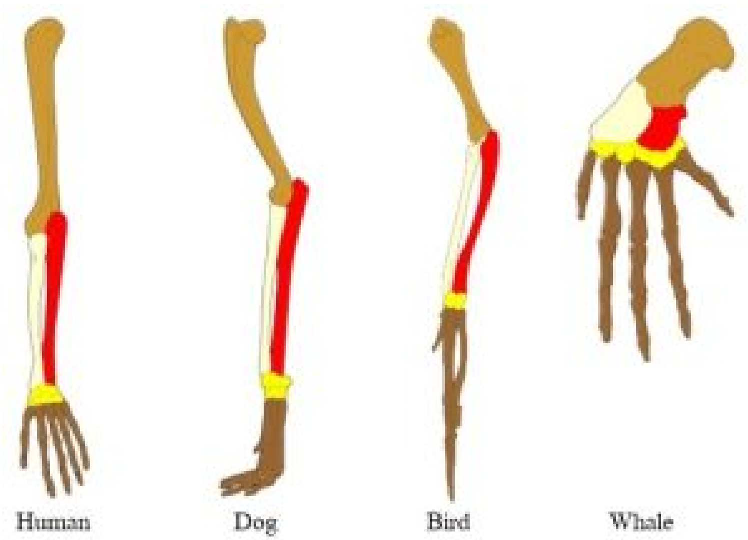


Figure 6.7. *Diagram of four animals suggesting descent from a common ancestor. Each has the same basic forelimb bone structure. Tan, humerus; red, ulna; light yellow, radius.* Волков Владислав Петрович, CC BY.

Existence of transitional fossils. One of the long-standing criticisms of evolutionary theory is that the fossil record lacks so-called transitional fossils, those that show organisms in the middle stages of evolution. However, recent research has uncovered multiple examples of fossils that possess traits of both earlier and later forms of organisms, indicating (but not *proving*, as we learned in Chapter 2!) a link, or transition, between them. One of the most dramatic and well-documented cases is that involving the evolution of whales (including dolphins)[3]. The history of these organisms has been confounding for centuries because of their unique combination of traits: they are confined to water, but, unlike fish, are mammals that have lungs for breathing. Such oddities led scientists to hypothesize that recent ancestors of these aquatic animals were land based, moving into the sea during their evolution. Examinations of fossils during the past century finally connected modern whales to a terrestrial animal that went extinct over 50 million years ago, but evidence of intermediate organisms, those that could document how these mammals transitioned to a permanent life in the ocean, had yet to be discovered. That all changed in the 1980s and 1990s, when paleontologists discovered fossilized remains of organisms that lived between 30 and 50 million years ago, the period during which whales were thought to have made their big move to water. These intermediate forms do indeed possess traits of both earlier and later forms, such as four leg-like limbs (fore and hind), tails, and ears adapted to hear underwater. As time went on, there is clear fossil evidence that bodies grew too big for the small legs that were shrinking but still present, and teeth became better and better for eating fish. The blowhole through which they breathed also moved from the front of the head to the top. Today's whales have flippers for forelimbs (instead of legs), and only internal evidence of pelvic and leg bones (i.e., boney remnants of these former structures attached to their skeletons). Much has been written about whale evolution, and you are encouraged to consult the reference cited in this paragraph for more information.

Observations of evolution. Another common criticism of this theory is that nobody has ever observed evolution occur. Simply put, though, that widely promoted objection is incorrect: we have indeed witnessed natural selection and evolution unfold in real time.

3. Thewissen, JGM., Cooper, LN, George, JC et al. Evo Edu Outreach. 2009. 2:272. doi:10.1007/s12052-009-0135-2. CC BY

Antibiotic resistance

Much about this phenomenon was described in Chapter 5 (review Box 5.6), and was also mentioned earlier in this chapter, so we will not repeat those details here. However, in this current context, you should realize that the development of bacteria able to survive exposure to poisonous substances is a direct result of rapid evolution. Through one of many possible mechanisms (some are explained above), the DNA in a subset of a bacterial population is changed—undergoes a mutation—such that the affected cells can now defend themselves against penicillin or another drug. Put another way, additional, potentially useful, information is added to the gene pool. If environmental conditions change and the bacteria are exposed to a relevant antibiotic, the cells with the new ability have an enormous advantage over all the others; due to natural selection, they reproduce and succeed whereas the sensitive cells perish. In a matter of hours or days, the population is dominated by resistant individuals. As we noted in Chapter 5, antibiotic resistance develops in natural environments, including human bodies. It can also be easily induced under laboratory conditions. If a population known to be sensitive to the effects of an antibiotic is grown in a container containing that substance. The few cells that survive are transferred to a second antibiotic-containing vessel and allowed to reproduce. Within days, a population of cells essentially identical to the original one (except now it is unharmed by the antibiotic) is thriving. In other words, such an experiment allows us to observe evolution directly.

Pesticide resistance

Through a mechanism like the one active in the development of antibiotic resistance in bacteria, a population of insects can become resistant to the toxic effects of compounds that ordinarily kill them. Since these animals are quite a bit more complex than bacteria, the process typically requires a few years to occur (as opposed to a few days as seen in the prokaryotes). But evolution of these multicellular eukaryotes with relatively long and complicated DNA is still easily observable. We will return to this topic in Chapter

9 when we explore the environmental effects of agriculture and learn why, although it is a very interesting phenomenon and a nice illustration of natural selection at work, the development of pesticide resistance is not at all desirable or helpful to humans (it is great for the insects, though!).

Other experimental results

Many experiments have been designed to directly observe evolution. One such study conducted in the late 1970s showed how a population of guppies evolved in response to changing environmental conditions[4]. Since brightly colored male guppies attract both mates and predators in natural streams, the appearance of this fish will vary with specific conditions. In the presence of fewer predators, the colorful form dominates, but duller forms are prevalent in streams subject to heavy predation. Armed with this information, researchers set up artificial ponds to observe how the population would respond to various stressors. Notably, brightly colored males were most common when predation was absent. But when predators were suddenly added to those environments, the guppy population quickly changed to one dominated by dull-colored males. The experiment provides evidence that natural selection favors individuals with the greatest advantages. Here, those that are least likely to be eaten lived long enough to reproduce. Additionally, these results demonstrate how more information is carried around in the genome than is necessarily expressed through the physical appearance of a population (as we noted earlier in this chapter). In other words, these fish are able to quickly adapt because of what we might consider to be extra information in their gene pool.

4. Endler. JA. 1980. Natural Selection on Color Patterns in Poecilia reticulate. Evolution. 34:76-91. Public domain. https://pubmed.ncbi.nlm.nih.gov/28563214/

Different types of evolution in biomes

In this final section on evolution we return to the concept of biomes we first encountered in Chapter 5 to consider three common and well-studied types of evolution.

Convergent evolution. Sometimes organisms in the same biome share many physical characteristics despite the fact that DNA analysis reveals they are only distant relatives. How could organisms with very different genes look similar? Because they live under the same conditions and stressors, distinct species often adopt common survival strategies. Put another way, they take on, or converge into, the same form (Figure 6.8).

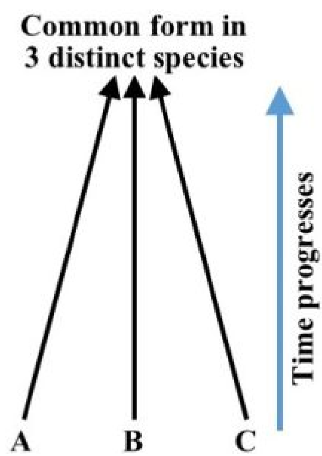


Figure 6.8. *Model of convergent evolution. Three genetically unrelated species, A, B, and C develop similar adaptations and physical forms due to evolution in the same biome.* Kelsey, CC BY-NC-SA.

Examples of this phenomenon are not difficult to find. First, consider killer whales and great white sharks (mentioned above). Although they are both marine animals, they evolved from very different ancestors. Whales, as we know, can be traced back to a 50-million-year-old terrestrial mammal, whereas sharks and their predecessors are fish that have been evolving in the ocean for hundreds of millions of years. Genetic sequencing confirms these different lineages. Because they live in the same type of habitats (even overlapping in many places) and share many of the same food sources, though, they have a lot in common, notably, their size, shape, speed, strength, and sharp teeth. Certainly, there are differences as well, but their gross similarities are unmistakable (Figure 6.9).

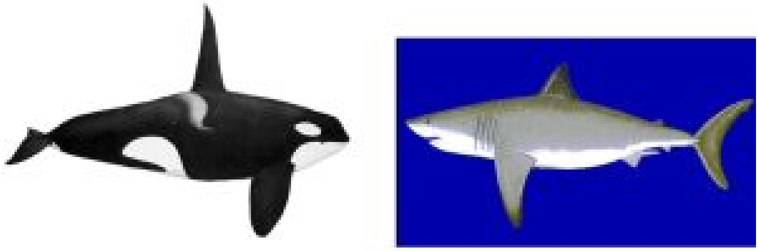


Figure 6.9. *Killer whales, left (a) and great white sharks, right (b) are not genetically related but share many of the same adaptations for survival in their biome due to convergent evolution.* US NOAA, *Public Domain (a); Robbie Cada, Public Domain (b).*

Second, birds and bats are also very distant relatives, yet they both evolved wings for flying to survive in their common habitat. Keep in mind that their wing structures are quite different from each other, though, suggesting they evolved from different ancestors. Multiple pieces of evidence suggest that birds are more closely related to crocodiles than bats, and bats, in turn, are far more like mice than birds. Finally, convergent evolution can even be seen across millions of years. The extinct dinosaur triceratops and the modern rhinoceros have very little in common genetically: they arose from distinct ancestors (i.e., rhinos did *not* descend from dinosaurs). However, to survive the same types of environments, food sources, and other stressors, they took on rather similar physical forms (Figure 6.10).

Figure 6.10. A *comparison of a model of extinct triceratops (based on fossil evidence) (a) and photo of modern rhinoceros (b) suggests convergent evolution. Zachi Evenor,* CC BY *(a); Charles Shapr,* CC BY-SA.

Divergent evolution. In this second case, a single population splits into a few distinct, yet closely related, species living in separate places and biomes. That is, the reason different species resemble each other is their common genetic origin, *not* because they adapted to the same stressors in the same place (Figure 6.11—compare it to that for convergent evolution, Figure 6.8).

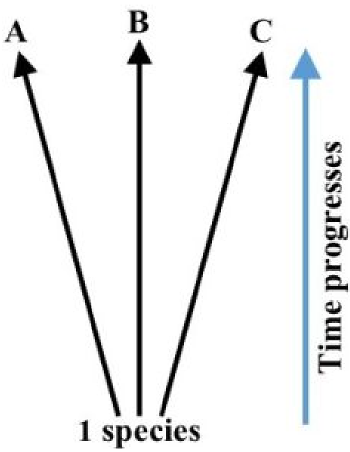


Figure 6.11. *Model of divergent evolution. Three closely related species, A, B, and C developed from a common ancestor because they were separated and then subjected to different environmental stressors. Kelsey,* CC BY-NC-SA.

Migration, volcanic or earthquake activity, or even the breakup of Pangaea and movement of continental masses (Chapter 3), could cause groups within a population to live and evolve under different environmental conditions. Results of divergent evolution can be seen throughout the biosphere. For example, the kit fox, arctic fox, and the red fox are three closely related species that emerged from a single ancestor[5]. The kit evolved in desert environments, developing a sandy color and relatively large ears to help dissipate heat, the arctic developed a white coat in snowy areas, and the red is adapted to life in the forest, where its color allows it to stay hidden among trees (Figure 6.12).

5. Wayne RK , Nash WG , O'Brien SJ. 1987. Evolution of the Canidae. II. Divergence from the primitive carnivore karyotype. 44:134-141. Public domain. https://pubmed.ncbi.nlm.nih.gov/3568762/ Cytogenetics and Cell Genetics

Figure 6.12. *Kit, left (a), arctic, middle, (b) and red foxes, right (c) show evidence of divergent evolution.* US FWS, Pubic Domain (a); Eric Kirby, CC BY-SA (b); Minette Layne, CC BY (c).

The divergence of monkeys and great apes from a common ancestor some 25 million years ago is another good example of this phenomenon. These many species share much of the same DNA sequences (well over 90% in most cases) yet evolved distinct sizes, lifestyles, and other traits in response to differences in the stressors acting on them.

Adaptive radiation. This process is like divergent evolution in that many closely related species arise from a common ancestor. A wheel is a better model for adaptive radiation, though, because competition for resources can cause individuals of a single population to evolve specialized survival strategies and an appreciable increase in the number of related species *in a particular area* (Figure 6.13., left panel). Changes in the gene pool and natural selection lead to distinct new species with the capacity to exploit multiple, narrowly defined **niches** (i.e., each has a particular set of features that allows it to play only a small, distinctive role). Every population uses a slightly different strategy to survive, so the restrictions of the competitive exclusion principle are eliminated (Chapter 5). In short, similar organisms can exist side by side. The most famous example of this phenomenon is, arguably, that reported by Charles Darwin in 1859. While studying in The Galapagos Islands, he saw a large amount of biodiversity. Notably, the many different species of finches (a type of small bird) living there possessed similar, but not identical, physical features. After much research he proposed that such variety was the result of evolution from a common ancestor. Because of intense **intraspecific competition** (Chapter 5), individuals able to find food in novel ways have an advantage over others. He described many small niches occupied by species with different forms. To name just some of the fourteen recognized, beaks specialized for hunting of surface insects are different than those used by species that tear open plants to find food, whereas nut eaters have a thicker, stronger beak to break open hard shells. Again, both visual inspection and gene sequencing reveal that these birds are very close relatives that evolved from a single species (Figure 6.13., right panel).

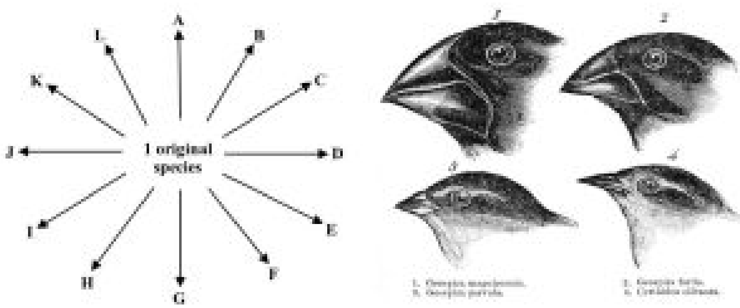


Figure 6.13. *Conceptual diagram of adaptive radiation: one species evolves into many closely related, highly specialized species (indicated by different capital letters), and they all live in the same area, left (a). Sketches of four closely related Finches from the Galapagos Islands that are an example of the concept of adaptive radiation (from the expedition of Charles Darwin), right (b). Kelsey, CC BY-NC-SA (a); John Gould, Public Domain (b).*

> *Worth extra emphasis: **divergent evolution** is the explanation for similar-looking species found in different areas, and **adaptive radiation** is the explanation for similar-looking species living side by side.*

6.3. BIODIVERSITY VARIES WITH SPACE AND TIME

Some areas seem to be obviously teeming with many different forms of life whereas others are dominated by a very small number of species. In addition, careful examination of fossil evidence reveals that species richness has gone up and down during the planet's history. Here we briefly consider some of the explanations for these observed variations in biodiversity.

6.3.1. Factors increasing biodiversity

Processes favoring evolution and therefore speciation tend to increase the number of different species present in an ecosystem. Keep in mind that these are merely generalizations, and multiple factors can interact to bring about unexpected results.

Physically diverse habitats

Anything that increases the number of different types of habitats within an ecosystem will stimulate populations to develop unique survival and life strategies. So, a flat and relatively uniform plain or ocean floor will be likely less diverse than a rocky, hilly area or coral reef zone, respectively.

Moderate disturbance

This factor is tightly linked to the previous one because a relatively minor event such as a small fire or flood can increase habitat diversity and favor new adaptations and species.

Middle successional stages

As we saw in Chapter 5, **ecological succession** changes the dominant species in an area. Generally, early- and late-stage communities tend to be less biodiverse than the middle, transitional ones.

Long-term environmental stability

This factor might not seem to fit our pattern at first, but environments that do not change for millennia or more can encourage the development of increasing numbers of species as they age. Ecologists generally agree that such stability allows for experimentation in the biosphere and a narrowing of niches (as described above). Viewed another way, all available niches can be exploited if evolution is given sufficient time. Diverse tropical rainforests provide a good example of the effect of stability on biodiversity. Since sunlight and other nutrients are abundant near the equator, organisms are productive all year long. Resultant near-constant interactions push competitors to develop new strategies to survive. The adaptive radiation seen in Darwin's finches demonstrates a nice specific case of how stability can lead to specialization.

High diversity in low trophic levels

Through one or more of the mechanisms noted on this list, biodiversity among organisms at the first trophic level can increase. Higher-trophic levels will then likely become more diverse and specialized in response.

6.3.2. Factors decreasing biodiversity

Anything that increases rates of disappearance of existing species relative to those of the evolution of new species will decrease biodiversity.

Extreme environmental conditions

Only a few species are likely to have the capacity to adapt to and thrive in unusually challenging habitats. For example, very low or high average temperatures, a wide range of possible temperatures from lowest to highest, excessive or scarce nutrients or water, or extremes in other properties including those related to chemistry, light, other forms of radiation, or atmospheric pressure tend to create conditions that are at the limits of tolerance for organisms. As a result, biodiversity in such places is typically quite low, with ecosystem functions carried out by the few generalists that can survive under such stress. For example, deserts, the tops of tall mountains, the floors of very deep oceans, and polar regions are simply too harsh for many species to develop the traits necessary for success.

Extreme disturbances

Events like volcanic eruptions, severe flooding, and human construction can disrupt natural systems and vastly reduce biodiversity. If succession is eventually allowed to proceed, diversity could increase. However, when a system undergoes long-term and persistent change, say, a human-induced transformation to a farm, park, golf course, or neighborhood, biodiversity is likely to remain limited.

Invasive species

As we saw in Chapter 5, certain **exotic organisms** known as **invasive species** are better suited for survival than are species endemic to an area. Since they grow more quickly and aggressively, these outsiders outcompete the natives and take over an ecosystem in a short period of time. Both species richness and evenness may decline precipitously after invasives appear.

Extinction

A species that is no longer present on Earth is said to be **extinct.** Now, things get more complicated from here because the circumstances leading to extinction vary, and many forces can contribute to the loss of a species. Adding to the confusion, it can be difficult to know if an organism is extinct or simply rare—many species have seemingly disappeared from forests, prairies, lakes, oceans, and other ecosystems only to be seen alive and well (albeit scarce) after escaping observation for decades. We explore some of the causes and consequences of extinction below.

> A *point worth some emphasis: extinction is a permanent condition. In other words, once an organism is lost from the biosphere, it will not return (current science fiction or future real science notwithstanding).*

Small populations are imperiled. Any phenomenon that reduces numbers can indirectly lead a seemingly robust and successful species to extinction because small **populations** are inherently susceptible to total elimination. Put another way, it is unlikely that a single cause will directly kill off every member of a population, rather, low abundance creates its own risk. Two important reasons small populations go extinct are briefly described here.

Population risk

Numbers may be, in principle, high enough to maintain a healthy population indefinitely. In practice, however, the rate of reproduction is too low to keep up with death rate because individuals simply do not interact regularly enough to produce offspring. It comes down to density: the chances of two fertile individuals finding each other declines as members of a species are spread more and more sparsely within a large habitat (see Figure 6.14). If you picture, say, a handful of whales trying to locate mates in the world's enormous oceans, you can appreciate the way population risk hastens extinction.

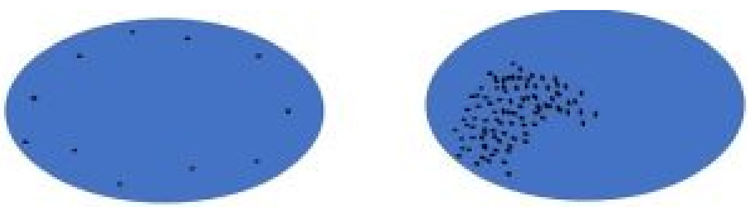


Figure 6.14. *Reduced population size can impede reproduction. This oversimplified diagram captures the essence of the problem: each dot represents a whale in a vast ocean. Low numbers and low density (left) can lead to extinction, whereas high numbers and high density (right) will facilitate success. Kelsey, CC BY-NC-SA.*

Genetic risk

A population may get so small that closely related individuals have little choice but to breed with each other. This phenomenon, known as **inbreeding**, weakens a species because it tends to raise the number of congenital defects passed to young. Why is inbreeding risky? As we learned earlier, DNA mutations do not always lead to adverse physical traits. Realize, though, that for various reasons such errors can stay masked and have no obvious effect on an organism that carries them. A union of two organisms having similar genetic codes (complete with the same unexpressed mutations) is far more likely to yield poorly adapted offspring than is one between unrelated parents (i.e., those with dissimilar DNA). Subsequent generations are increasingly unfit to survive, as mutations become ever more prevalent, numbers drop further, and extinction ultimately occurs. High genetic diversity, therefore, is more beneficial to the long-term survival of a species. Consult Box 6.3 for an example of the effects of inbreeding among humans. Genetic risk can be particularly important for isolated populations such as those on small islands or are otherwise segregated (see Figure 6.15 for one example).

Box 6.3. Risky royal reproduction

To maintain a sort of genetic purity (or at least the perception of one), some royal families have deliberately reproduced with close relatives. Unfortunately, this once widespread practice led to unintentional consequences: many offspring of those unions were born with debilitating congenital diseases. For example, the Spanish Hapsburgs (16th – 17th Century) were a dynasty that died out due to inbreeding. As a result of a few generations of marriages among people with similar genes, they got weaker and weaker until Charles II, mentally and physically disabled, died without leaving an heir in 1707[6]. Some of the descendants of Queen

6. Alvarez G, Ceballos FC, Quinteiro C (2009) The Role of Inbreeding in the Extinction of a European Royal Dynasty. PLoS ONE 4(4): e5174. CC BY. https://doi.org/10.1371/journal.pone.0005174.

Victoria of England (19th Century) experienced a similar fate. The genetic disorder hemophilia (a blood-clotting disease that was largely untreatable a century ago) was prevalent in her family tree. Among those affected was the son of Alexandra (she was a granddaughter of Victoria) and Tsar Nicolas II of Russia (he was one of Alexandra's cousins). You have likely heard of Nicolas, as he was famously murdered along with his hemophilic son and the rest of his family in 1918 by Bolsheviks—but you will have to consult a very different kind of book for more about them.

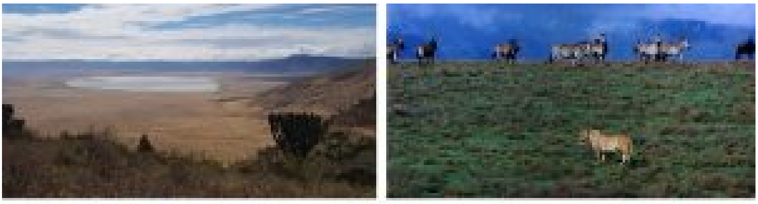


Figure 6.15. *The Ngorongoro Crater in Tanzania (Africa) is an example of an isolated ecosystem. As is evident from the photo on the left (a), movement of animals in and out of this space is hindered by the geometry of the crater. Organisms inside, then, are potentially subject to genetic risk. The small population of lions within this habitat (a lioness with her zebra prey), right (b), is especially vulnerable to the effects of inbreeding, and some scientists worry that they are becoming less and less fit with each generation. Kulwa Kanyoro, CC BY-SA (a); Brocken Inaglory, CC BY-SA (b).*

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0005174

Forces reducing population size. Although humans certainly play an outsize role in driving extinction of organisms these days—we will explore many of the ways they do that below and in upcoming chapters—it is worth noting that natural processes have been responsible for nearly all extinctions during Earth's long history. The demise of the dinosaurs some 66 million years ago, for example, took place at least 64 million years before beings resembling *homo sapiens* even appeared (and some 66 million years before modern humans began to exert their influence). This famous extinction is just one of the hundreds of millions that have occurred since life first arose. It would be fair (and possibly humbling) to say that every organism is subject to extinction sooner or later: data indicate that on the order of 99% of all species *that have ever lived* are now extinct! But it is a complex and nuanced phenomenon, as species persist for different lengths of time, and the rate of loss from the biosphere has ebbed and flowed substantially. Although extinction appears to occur continually, scientists have used fossil evidence to identify five distinct and relatively short-lived periods in the past 600 million years, known as **mass extinction events**, during which an unusually large fraction of species, in the range of 75% or more, living at the time went extinct. The most recent of these is thought to have been the one that ended the 165-million-year reign of the dinosaurs. As we will explore below, a devastating mass extinction event may well be under way today.

Natural environmental change

We have seen multiple times (for example, in Chapters 1, 3, 5, and earlier here) that organisms best adapted to their environments are most likely to survive and reproduce. It should come as no surprise, then, that changes in conditions due to natural processes can jeopardize the ongoing success of previously dominant individuals. Generally, the higher the rate of change the more difficult it is for organisms to develop necessary adaptations for ongoing survival.

Sudden. Some phenomena are just too sudden and powerful for organisms to survive. Volcanic eruptions, earthquakes, tsunamis, infectious diseases, and asteroid impacts are all examples of agents that can devastate populations, communities, or entire ecosystems. Whether an event leads to extinctions depends on the scope of the damage it causes as well as the uniqueness of the affected habitat. Imagine an organism so specialized that it is only found in one small area on Earth. In such a case, a single volcanic eruption, with its lava, poisonous gases, and ash (Chapter 7), could so dramatically reduce numbers that extinction occurs. The global-level effects of a collision with a large asteroid, however, could imperil even abundant, widely distributed species.

Gradual. Slow alterations to an environment may cause a dominant organism to be replaced. Climate change due to natural forces, increased predation or competition (Chapter 5), disease, or other stressors could all drive species that cannot adapt to the brink of extinction.

Anthropogenic activities

Before exploring this topic, a note about the organization of this book is appropriate. You might recall from your reading of Chapter 1 that Part II (Chapters 4 – 6) was advertised as an examination of natural phenomena with little attention devoted to human sources of stress. Here we temporarily deviate from that scheme because a discussion of extinction would be incomplete without some explanation of the roles played by anthropogenic activities.

The numerous and varied ways humans contribute to extinction can reasonably be placed into one of two categories. The first of these, **direct**, involves activities that are *intended* to kill or capture organisms. **Indirect** causes, those in the second category, do not set out to harm organisms, instead, they initiate changes that ultimately endanger survival of a species. The motivation for an activity, in other words, may be used to distinguish between these categories of extinction. We will keep things general for now but explore many specific human activities and their consequences in Chapters 8 – 14.

Direct causes of extinction. Hunting and harvesting are two familiar activities that seek to remove organisms from their habitats. The issue is nuanced, of course, because the killing of organisms that are extremely abundant does not necessarily lead to population risk. For instance, white-tailed deer reproduce so quickly that hunting of these animals does not even maintain a stable population in many areas—their birth rate is just too high for hunting to keep up. Large numbers do not ensure resilience, though, as in the famous case of the passenger pigeon. As recently as the Nineteenth Century, many billions of these birds lived in North America. High-density flocks containing millions of birds—the skies were said to darken when they flew overhead—were such tempting and easy targets that human hunters drove them to extinction a little more than a century ago[7]. The Bengal tiger of India provides one modern-day example of an animal in danger of going extinct. Since its numbers are already quite low, laws have been passed to protect and conserve the cat. Ongoing hunting (or, more appropriately, poaching), along with other stressors described in the following sections, continue to affect this and similarly threatened organisms.

Indirect causes of extinction.

- Habitat destruction. By modifying or eliminating natural ecosystems to make space for housing, recreation, industry, commerce, and farms, humans can introduce sufficient stress to drive species to extinction. The problem is more complicated than it may appear because the effects of development depend on factors such as whether an organism is a generalist or specialist and how unique is the affected habitat. An organism with a broad niche (Chapters 5 and

7. U.S. EPA. Endangered species. 2019. https://www.epa.gov/endangered-species/learn-more-about-threatened-and-endangered-species

above)—say, rats (clearly, they can live about anywhere humans live)—will be far less vulnerable to alterations than would one with a narrow niche—say, giant pandas (an animal having very quirky habitat needs and reproductive habits that, even absent human activities, hinder their long-term survival). For our purposes, "destruction" can refer to dramatic replacements of natural systems by human ones; think clearing a dense forest and replacing it with a shopping mall or agricultural field.

- Habitat fragmentation. Seemingly subtle changes and modest losses of natural habitat could profoundly affect vulnerable organisms. Although a construction project might convert just 10% of land into a highway or pipeline (Chapter 10), the presence of this unnatural barrier could interrupt migrations, reduce genetic diversity in a small region, disrupt seed dispersal and pollination, or otherwise impede behaviors necessary for a species' survival. For example, a single forest split into disconnected sections by roads, railroad tracks, parking lots, recreational areas, or other anthropogenic features could lead to the extinction of an organism that requires large areas of continuous space to be successful (Figure 6.16). Dammed rivers and divided wetlands can similarly harm an organism.

Figure 6.16. *Human development can lead to habitat fragmentation and reduce the success of a species. The addition of roads and houses creates more and more barriers with time (top diagram, in map view). These highways in Indiana have disrupted the free movement of animals (bottom photo, map view). Kelsey, CC BY-NC-SA (top); USGS, Public Domain (bottom).*

- Habitat degradation. Humans may not appreciably alter the appearance or continuity of a habitat, but they still can decrease its suitability for organisms that depend on it for survival. Water pollutants (Chapter 11), direct dumping of toxins (Chapter 13), or release of air pollutants (Chapter 14) can stress organisms living in a region. Somewhat related to habitat degradation: actions that have a direct impact on one **trophic level** could indirectly harm other organisms in a community. For example, by killing members of a prey species, humans might decrease the amount of necessary nutrients available to a high-level predator.

Consequences of extinction. The disappearance of an organism from Earth can trigger consequences ranging from subtle to severe.

Biodiversity decreases

As stated above, each extinction reduces the biodiversity in an ecosystem (and, for that matter, in Earth's biosphere). The relevance of such changes is addressed later in this chapter.

Ecosystem structure changes

To understand this effect of extinction, recall that we defined structure as the identity of the organisms present in an ecosystem (Chapter 5). Clearly, when a population disappears, the inventory of species is modified. Effects of such changes can vary, as we will see in the next item.

> *Of course, whether biodiversity increases or decreases is more complicated than the simple statement in the main text implies. We would need to use systems analysis to fully answer the question: if the rate of output (i.e., extinction) outpaces the rate of input (i.e., evolution of new species), then biodiversity* ***will*** *decline.*

Ecosystem function may change

Extinction can alter this second ecosystem characteristic, although its effects are not always dramatic or obvious. On the one hand, a loss (i.e., altered structure), could bring on catastrophic changes if members of the affected population carried out a critical in an irreplaceable way. Such an organism is often called a **keystone species**, one that has a disproportionate effect on the survival of an ecosystem (relative to its abundance). Consider, for instance, the potential effects of extinction of a high-level predator: among other things, the growth of prey species could accelerate and lead to habitat degradation. The sea otter in waters off the Pacific northwest of the U.S. is a keystone species because it preys upon sea urchins that would otherwise destroy kelp forests[8]. The white nose bat is also a keystone species because it plays such crucial roles as both predator and prey in its ecosystems (largely northeastern United States). Recent declines in bat numbers are of great concern because extinction could devastate the ecosystems it inhabits[9]. Alligators in swamps in southeastern U.S.[10], gorillas in Africa[11], and oak and hickory trees in the central hardwood forest (east-central U.S., including Tennessee)[12], are all examples of keystone species. On the other hand, some organisms are less crucial for the survival of entire ecosystem because their function, although necessary and important, can be carried out by another in largely the same way. In principle, removal of such an organism could enable a second species to grow and expand to ensure the overall survival of an entire ecosystem. Commercial fishing (Chapter 12) is one activity with the potential to exert so much pressure on one species that others having less economic and nutritional value for humans could take over. In such a case the whole ecosystem would suffer minimal changes, yet *people* would

8. National Parks Service. 2016. https://www.nps.gov/glba/blogs/a-keystone-species-the-sea-otter-colonizes-glacier-bay.htm
9. Fish and Wildlife Service. 2017. https://www.nps.gov/articles/what-is-white-nose-syndrome.htm
10. U.S. EPA. 2020. Watershed Academy. cfpub.epa.gov/watertrain/moduleFrame.cfm?parent_object_id=540
11. U.S. Fish and Wildlife Service. ecos.fws.gov/ecp/species/4080
12. Fralish, J.S. 2004. The Keystone Role of Oak and Hickory in the Central Hardwood Forest. U.S. Forest Service. fs.usda.gov/treesearch/pubs/6500

be harmed due to the loss of a valued commodity. We will see more about the effects of extinction on humans below.

Human civilization can be affected

Loss of important ecosystem services. Humans depend on many processes active within ecosystems, including the release of dioxygen by primary producers, nutrient and water recycling, decomposition, and trophic interactions. Extinctions can disrupt these services by changing **ecosystem structure** and function.

Loss of economically valuable products. Since extinction is an irreversible process, it could lead to the permanent disappearance of organisms having medical, industrial, or commercial uses. The Pacific yew tree is one example of the way extinction could bring on an undesirable outcome from an anthropogenic perspective. In the 1980s, scientists discovered that paclitaxel (known as Taxol commercially), a chemical substance derived from the tree's bark, could treat human cancers. Subsequent high demand for this drug, as well as other stressors from human development, have reduced the size of the yew's population. Although numbers have not yet dwindled to critically low levels, the potential adverse outcome of its extinction is likely obvious: loss of this tree would mean no more Taxol[13]. Unsustainable harvesting of fish (known as **overfishing**, a phenomenon we will revisit in Chapter 12) could similarly lead to the loss of a source of both income and food. Reductions in the availability of valuable organisms could also lead to conflicts among nations and peoples. Finally, high biodiversity attracts tourists (and their money) to areas having unique ecosystems. For nations that are heavily reliant on income from visitors seeking exotic wildlife, extinctions might equate to job losses, increased poverty, and crime.

Subjective considerations and losses. Obviously beyond the scope of science, these are important to many people. Put simply,

13. U.S. National Institutes of Health. dtp.cancer.gov/timeline/flash/success_stories/S2_Taxol.htm.

human-driven extinction is seen by some as morally unacceptable. Although not objectively measurable, concerns about the rights of non-human organisms to survive as well as aesthetic and cultural desires to maintain Earth's current biodiversity inspire interest in preserving endangered organisms.

Steps can be taken to slow extinction. Many laws enacted during the past several decades are intended to counteract or reverse human contributions to extinction. In the United States, for example, President Theodore Roosevelt helped establish the first national wildlife refuge to protect threatened water birds in 1903. In the years that followed, numerous other related actions were taken by the federal government, including the passage of the Endangered Species Act (ESA) in 1973 (with amendments in subsequent years). In short, this law uses scientific data to identify at-risk species and empowers the Fish and Wildlife Service (FWS) to formally designate them as **endangered** or **threatened**. These two categories reflect a matter of degree: organisms in the former are in imminent danger of complete disappearance and those in the latter have dwindled enough that they could become endangered. Listed species are protected by restrictions on hunting, harvesting, or any actions that directly harm them (referred to as **take** in the language of the law), as well as human activities that would damage their habitat or otherwise indirectly threaten them. Many plants and animals have been added to the list of endangered and threatened species since the ESA was first enacted, including the American alligator, American peregrine falcon, and the San Clemente Island Indian paintbrush[14]. The alligator and falcon have since recovered sufficiently to be delisted, but the paintbrush shrub is still considered threatened by the FWS (we will see other examples

14. Fish and Wildlife Service. 2023. Endangered species. www.fws.gov/program/endangered-species

shortly). State and local agencies in the U.S., as well as international organizations, also pass laws to protect imperiled organisms.

Protecting species poses multiple challenges, especially when such efforts are perceived to conflict with human quality of life. Should we restrict logging in a unique ecosystem because it endangers the survival of a (seemingly) obscure insect or bird? The people who earn their living by cutting down trees often answer with an unequivocal "no" (more in Chapter 1)! Similar questions can be asked about many other potentially threatening endeavors such as development of new housing, expansion of agricultural land (Chapter 9), exploitation of fossil fuel resources (Chapter 10), and commercial fishing (Chapter 12).

Here we take a generalized look at the protection and recovery of endangered and other at-risk species. We will see some important example organisms and the specific measures taken to protect them in the next section.

Identifying a change in population status

Scientists working for various governmental agencies conduct surveys of organisms and ecosystems to monitor trends in population sizes and stability. If a species appears to be declining, a decision is made about changing the way it is regulated, that is, whether it should be added to a formal list of endangered or threatened species (compiled by U.S. state and federal agencies and at the international level). A species on one of those lists gains special protections against measures that could further imperil it (including hunting, commerce of products from it, and land use decisions that adversely affect its habitat).

Determining the cause of decline of a species

We need to determine what is limiting success—lack of suitable habitat, new competitors, moisture stress, pollution, excessive hunting, to name just some of the possible stressors—before we can do anything to reverse an observed decline.

Devising a recovery plan

Identifying the underlying problem is often a challenge in its own right, but taking the necessary next steps can be even more difficult. Often, unpopular steps are required, including the imposition of restrictions on harvesting and hunting, the establishment of protected areas, and the tightening of laws regarding new construction. In any case, the development of a plan to help an endangered species requires answers to several questions, including those described here.

What is our objective? This seemingly simple question can be, in practice, very difficult to answer. Of course we want the size of a target population to increase to healthier levels, but determining what that number might be presents some challenges. How big is big enough? Because of divergent points of view and agendas, different people will come to different conclusions from the same data about a target organism. In many cases, it is impractical for a population to rebound to some historical size because humans now live in its former range. Do we really want gray wolves to return to the geographic locations that are currently villages and cities? Assuming most people would opt out of sharing space with dangerous predators, we need to settle on a less-ambitious goal like small but sustainable populations living in remote areas (more about wolves follows shortly).

What steps should we take? The answer to this question is highly dependent on the species of interest, its habitat, and the causes of its decline. If it is a commercially important fish, restrictions on harvests are clearly in order, whereas protection of a marine mammal that is not deliberately caught but gets killed inadvertently by gear designed to capture fish would likely involve changes to fishing practices (Chapter 12). Elephants, rhinoceroses, lions, and bison benefit from restrictions on hunting and poaching as well as habitat protection. Banning the trade of things like ivory is also intended to enhance survival of protected species. Endangered insects, worms, plants, and other types or organisms obviously require varying species-specific approaches. Wildlife refuges and zoos can be parts of recovery plans, although their utility and appropriateness are not universally recognized (see Box 6.4 for more).

Box 6.4. What about captive animals?

Animals (and, to a lesser extent, other organisms) are often taken from the wild and kept in zoos and aquaria. A related, but clearly different approach, involves large areas commonly called wildlife refuges in which organisms are protected and live under conditions that are very similar to their natural habitats.

1. **Zoos**. Animals are put on display in enclosures that may bear some or little resemblance to their natural habitats. Arguments are made in favor of and against their usefulness. On the plus side, their captive breeding programs can be important components of recovery plans for listed organisms. For example, the U.S. Department of Fish and Wildlife (FWS) has a relationship with the American Association of Zoos and Aquariums that, among other things, takes advantage of the capacity of zoos to increase reproduction of endangered species[15]. Furthermore, many hold that zoos educate and inspire people to contribute to conservation efforts. Detractors contend that zoos are unrealistic displays of inhumanely imprisoned animals and teach little about the value of conservation (e.g., African zebras in a snowy pen in Pennsylvania,

15. Association of zoos and aquariums memorandum of understanding. U.S. Fish and Wildlife Service. www.fws.gov/program/endangered-species/aza-memorandum-of-understanding

U.S.A.???). Some see zoos as an inappropriate allocation of resources that could otherwise be spent on more meaningful conservation efforts of organisms in their natural environments.

2. **Wildlife refuges.** These areas are intended to keep organisms in natural conditions. Put another way, rather than removing an animal from an African savannah and putting it into a glass-enclosed cage in Washington, D.C., efforts are made to protect organisms in their native habitats and enhance their ability to reproduce and thrive in place. Refuges have been set up in many countries. The U.S. FWS oversees several hundred such places throughout the United States and its territories. Since they address site-specific needs, the details of their operations vary considerably. In general terms, they are maintained to conserve natural species and habitats, including those that are endangered or threatened. You may be surprised to read that refuges often are not closed to the public. Hiking, swimming, picnicking, hunting, fishing, and harvesting tend to be allowed, with regulations and controls, in these protected spaces. Other examples of refuges are seen in multiple African nations stiving to strike a balance among hunting, farming, tourism, and conservation. The U.S. FWS is one of the partners that support the work carried out by the Central African Regional Program for the Environment (CARPE) to maintain and protect animals, plants, and ecosystems that are threatened by hunting for both trophies (e.g., elephant ivory and exotic pets) and food (e.g., the flesh of great apes and other large animals or **bushmeat**),

grazing by cattle, mining, logging, and other anthropogenic activities[16]. Their focus on in-place conservation and the needs of native peoples distinguish refuges from zoos, although the potential for abuse and corruption, among other weaknesses, make them an imperfect solution.

Important final thoughts. Before we move on, two additional comments are in order. First, our approach to protection has changed during the past several decades. Historically, efforts focused on a single species but, given our understanding of how organisms interact with and rely upon their surroundings, modern scientists often target whole ecosystems (more below). Rather than exclusively tracking and helping, say, a particular endangered turtle, we strive to understand the living and non-living factors upon which that turtle relies and conserve them. Many factors must be studied, such as whether (and how) anthropogenic activity affects the size of its wetland habitat, the chemical properties of the water in which it lives, the success of other organisms (e.g., competitors, parasites, predators, and prey), and so forth. Sometimes the trouble is relatively localized—say, runoff from a nearby farm (Chapter 9) or habitat fragmentation—but it could also be linked to larger issues such as global climate change (Chapter 14). Second, protection of organisms and natural systems can be accomplished by two distinct, if similarly named, approaches: conservation and preservation.

16. U.S. Fish and Wildlife Service. International affairs. www.federalgrants.com/Wildlife-Without-Borders-Africa-Program-53680.html

- Conservation. This strategy is predominantly driven by a desire to directly improve human quality of life. In other words, conservation is *management* of ecosystems and organisms so they can provide natural resources. Organisms and biodiversity are certainly protected—as the word suggests, conservationists seek to maintain systems in their current condition indefinitely—but growth of organisms is optimized to meet the needs of people. For instance, regulations on the rates of logging, fishing, and hunting are intended to maximize harvests and profit while conserving valued commodities in perpetuity. The designation and maintenance of open spaces, forests, and parks to provide necessary habitats and for recreational purposes also are conservation efforts. We will see examples of conservation of valuable resources in Chapter 12.
- Preservation. Unlike conservation, preservation seeks to protect natural systems because of their ecological importance and / or intrinsic value. The goal here, then, is not to directly enhance human standards of living but to allow ecosystems to flourish without any interference. Organisms are decidedly *not* viewed as resources. In principle, they are shielded from not just harvesting and the like but hiking, camping, and any anthropogenic presence or activity. To be a bit flip about it, picture a forest with a large KEEP OUT sign on its border and you will get the idea. Pure preservation is, as you might imagine, far more difficult to accomplish and is becoming more and more rare as the size and standards of living of the human population continue to increase. In fact, it is not uncommon for so-called preserves to simultaneously serve as both havens for organisms and areas for human enjoyment, recreation, and similar activities.

Examples of protected species

One could find many examples of conservation of terrestrial and aquatic organisms and systems around the world. Due to the limited amount of space in this book, we will consider just a small number that demonstrate some important principles and problems.

Giant panda. The case of the panda merits some consideration because of its many complexities and the difficult questions it raises. This organism has been on the brink of extinction for decades due to poaching and habitat destruction. Their very inefficient reproductive strategy, narrow **niche**, and other unhelpful traits make them particularly vulnerable. What pandas *do* have going for them is fantastic public relations and passionate advocates. It is safe to say that people tend to find these creatures irresistible, adorable, and worthy of financial and political support (Figure 6.17).

Figure 6.17. *The famous and photogenic giant panda. Rob, CC BY.*

And support they get: billions of dollars have been spent in China (their only native habitat) and other countries to study, breed, display, and protect pandas. The population of wild pandas is very small—a few thousand individuals at best—and captive breeding programs have added some hundreds to that number.

This textbook is not the proper venue for a lengthy examination of the costs and benefits of panda conservation. In brief, there is hardly universal agreement that resources and efforts should be expended to save such an organism, although proponents make several arguments to support their point of view. First, all efforts to preserve biodiversity are valid, including those associated with pandas. Second, humans should fix this problem because they are largely responsible for its creation. Third, the panda is a powerful symbol for broad conservation and inspires protection of other endangered species. Finally, they are culturally and economically valuable. You are encouraged to read more about the complex financial aspects of panda conservation at the following link: The

Value of Ecosystem Services from Giant Panda Reserves – PubMed. Some people, though, think panda conservation is a waste of money that could otherwise be spent on more ecologically relevant systems or even improving the conditions for impoverished humans. They argue that artificially propping up a poorly adapted organism is ultimately doomed to fail. Furthermore, recent research indicates that efforts to protect pandas in the wild have led to losses of large predators from the same habitats, including leopards, snow leopards, and wolves[17]. Importantly, data support the contention that focusing on a single species is likely to be less broadly successful than working to protect entire ecosystems. Another sticky yet crucial question is raised by the debate over pandas: is it appropriate for any species to receive a disproportionate share of attention (and help) because they are so aesthetically and emotionally appealing while organisms that are not so adorable (or, even subjectively ugly to humans) are ignored? In other words, in decision making about resource allocation and conservation efforts, what relative weight should be given to science, ethics, and popular opinion?

Aye-Aye[18]. Unlike the giant panda, this small, nocturnal mammal living on the Island of Madagascar (off southeast Africa) is decidedly *not* considered to be aesthetically (or otherwise) appealing by most people and has not garnered an enormous following of protectors and supporters (Figure 6.18).

17. Li, S., McShea, W.J., Wang, D. et al. 2020. Retreat of large carnivores across the giant panda distribution range. Nature Ecology and Evolution 4:1327–1331

18. Fish and Wildlife service. ecos.fws.gov/ecp/species/7643

Figure 6.18. *The aye-aye. Arguably, this organism is not objectively attractive (apologies to aye-aye lovers). Appearance is irrelevant to ecology, but it can influence conservation efforts. Vassen F., CC BY.*

Should it be protected with rigor? The aye-aye has roles to play in its ecosystem, just like its more attractive animal kin.

Gray wolf [19]. This large predator once roamed through most of the lower U.S. states in numbers that approached hundreds of thousands of individuals (Figure 6.19). As human population density increased, so did conflicts between people and wolves: farmers, ranchers, parents of small children, tourists, and others sought to eliminate wolves to protect their various interests.

19. Data about wolves are derived from the following two sources. (1) Fish and Wildlife Service. 2020. Endangered and Threatened Wildlife and Plants; Removing the Gray Wolf (Canis lupus). From the List of Endangered and Threatened Wildlife. www.federalregister.gov/documents/2020/11/03/2020-24171/endangered-and-threatened-wildlife-and-plants-removing-the-gray-wolf-canis-lupus-from-the-list-of (2) Wolf restoration. National Parks Service. nps.gov/yell/learn/nature/wolf-restoration.htm

Starting in the late 1800s, these animals were hunted, poisoned, and driven out of nearly all their former ranges. Among the areas that saw populations drop to near zero by the 1920s was Yellowstone National Park in Montana, Idaho, and Wyoming (U.S.A.), and we will focus our attention on this important example. The complicated story of wolves in Yellowstone is rife with science, politics, advocacy, and competing interests. Even as some people did all they could to kill off these potentially threatening animals, others worked to protect them. In the early 1970s the gray wolf was added to the endangered species list, but its status was debated and changed during the decades that followed. The recovery plan for this organism included unsurprising protections against hunting as well as a controversial effort to re-introduce it into the park.

Figure 6.19. *This grey wolf stands in the middle of a road in Yellowstone national Park. US NPS, Public Domain.*

After many years of passionate debate and legal wrangling, a handful of wolves were moved from Canada into Yellowstone in the middle 1990s. In the following decades, the populations expanded, and the arguments between pro- and anti-restoration camps raged on. Keeping things brief, as the number of wolves increased in the 1990s, 2000s, and beyond, they were taken off and re-added to the endangered species list in the three relevant U.S. states several times. Data indicate the population has grown to a few hundred individuals, a level scientists consider to be healthy under the complex circumstances. But the issue continues to divide conservationists from ranchers and others, with scientists stuck somewhere in the middle trying to provide objective assessments of the effects of wolves on other predators (like coyotes) and prey (like elk).

Holmgren milk-vetch[20]. Our final example is a plant found in specialized environments in Utah and Arizona (U. S.). It is likely not one that is known to you, and the details of its history and threats to it are not necessary for our purposes. It is included here simply as a reminder that the list of endangered species includes not just animals, but also plants, fungi, and others (Figure 6.20).

Figure 6.20. *It may look common enough, but this Holmgren milk-vetch is endangered. It and many other plants face uncertain futures due to anthropogenic and natural stressors. US FWS, Public Domain.*

What about resuscitation of extinct organisms?

As our understanding of genetics grows, additional strategies to

20. U.S. Fish and Wildlife Service. Environmental conservation online system, ECOS online. ecos.fws.gov/ecp/species/4590

protect endangered species, or even bring back some that are lost, may become available. However, considering what we know now, extinction is a permanent condition, one that cannot be reversed through methods that likely will remain in the realm of fiction. Consult Box 6.5 for a bit more about this topic.

Box 6.5. Could we bring back dinosaurs?

The short answer to this question raised in multiple books and films is: sorry, no. Why not? First, DNA does not persist for more than tens of thousands of years, not even close to the 65 million (or more) that have elapsed since dinosaurs went extinct. In other words, the genetic information needed to pull off this nifty trick is gone. Second, even if we could clone dinosaurs, the likelihood they would survive is practically nil. Think like an environmental scientist: atmospheric chemistry (including the relative proportion of gases like dioxygen and carbon dioxide) is quite different today than it was back in the Jurassic Period. The properties of the solar radiation striking the planet's surface have changed as well. Third, most of the plants and animals they ate are extinct today. While we are at it, many of the microorganisms that interacted with dinosaurs also are gone, meaning that digestion, immunity, and decomposition of their waste would likely be seriously hampered. Arguably, it is best that dinosaurs are relegated to history in any case—imagine the environmental changes *we* would experience if they came back.

Are we causing another mass extinction? As noted earlier, Earth's biosphere has been affected by at least five previous periods of widespread extinction during the past 600 million years. Much evidence indicates that human activity is currently driving a sixth. This time around, though, the pace of loss is hundreds or thousands of times faster than we would expect under natural conditions. Researchers estimate that at least 400 **vertebrates**, including reptiles, amphibians, birds, and mammals, have gone extinct in the past century—losses that likely would have taken 10,000 years in the absence of humans; many invertebrates and plants have disappeared as well [21]. Projections of future human expansion during the coming century (Chapter 8) suggest that extinctions will continue at an ever-increasing rate in the future.

6.4. WHAT ARE THE BENEFITS OF BIODIVERSITY?

Before we close this chapter, we should address a lingering question: when it comes to biodiversity, is more always preferable to less? As usual, the answer depends on who you ask and under what circumstances you pose the question.

21. Cebellos, G., Ehrlich, P.R., Raven, P.H. 2020. Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. Proceedings of the National Academy of Science of the United States of America. doi.org/10.1073/pnas.1922686117. CC BY

6.4.1. Conflicts can arise over it

We often face choices that pit human development against the protection of ecosystems and biodiversity. The economic and cultural benefits of, say, a new shopping mall, apartment complex, park, or farm must be compared to the value of maintaining a unique ecosystem that is one of the few remaining habitats of an endangered butterfly species. Priorities vary, as is well known.

6.4.2. Preserve it?

Analyses of the costs and benefits of preserving biodiversity, as well as the conflicts that tend to arise in these cases, are informed by both subjective, values-based reasons as well as objective, science-based evidence.

Subjective reasoning

We have noted before that the focus of this book is environmental science, not environmentalism. So, although we will acknowledge the passionate beliefs people hold about biodiversity, we will not be overly swayed by them. Instead, we will approach the problem as scientists.

Objective reasoning

The causes and relevance of biodiversity have been studied, but we still have much to learn about them. Two of the many questions we ask are briefly discussed here.

Does high genetic diversity provide benefits? We saw earlier that a single **population** with a large and diverse gene pool is more likely to adapt to and survive environmental change than one lacking such an advantage. Although it can help the affected species, is this diversity always beneficial to humans? If the organism is valuable to us (e.g., for economic, medical, environmental, or aesthetic reasons), we clearly appreciate genetic diversity. However, when it increases the fitness of disease-causing bacteria or crop-eating pests, outcomes we noted earlier, high genetic diversity is problematic and dangerous.

Is high species richness always best? This question may be more nuanced and controversial than you expect. Briefly, its answer is affected by many factors, including the type of ecosystem under consideration.

Artificially maintained environments

High species richness is often antithetical to accomplishing the goals of these systems. On a farm, for example, *low* biodiversity is generally preferable. As we will see in great detail in Chapter 9, food is usually produced in ecosystems manipulated by humans to maximize the growth of only one or a few organisms. In other words, high species richness is detrimental, and many steps are taken to reduce it. Public parks, golf courses, lawns, baseball fields, sports fisheries, and the like similarly strive to encourage the growth of a few desired organisms, that is, biodiversity in those places is intentionally kept low.

Natural environments

A widely held assumption is that high species richness is necessary to ensure the on-going health and success of all ecosystems on Earth. The larger the number of different species present, the argument goes, the more stable and able to withstand environmental change a system will be. High biodiversity is seen as a kind of insurance policy, if you will, analogous to high genetic richness within a single population. However, some ecologists question this view. They point out that since high biodiversity results in increased specialization and narrow niches, it could actually *increase* a system's vulnerability to change. How could this be? Keep in mind that ecosystems with low diversity are dominated by a relatively small number of species that each occupy a broad niche. In other words, the **functions** active in them are carried out by generalists. In principle, such an organism would have an easier time adapting to change than would a specialist because the former already operates under a wider range of conditions than does the latter. So, even subtle alterations in average temperature, water availability, and prey identity have the potential to imperil many specialists. The same changes would also exert pressure on generalists, but they are more likely to endure and continue to dominate. Proponents of this idea note that, although environmental stability tends to induce higher species richness (as we learned above), the reverse is not necessarily true. Keep in mind that this logic should not be confused with advocacy to induce extinctions or otherwise reduce the number of different species on Earth! It is simply an objective evaluation of the science, one that is certainly not universally held. In any case, high biodiversity is a feature we value for multiple reasons, and declines in it would be taken as a sign of increased stress. Furthermore, although we know it can indicate healthy ecosystems, there is a great deal about the role and importance of biodiversity that has yet to be learned.

THE CHAPTER ESSENCE IN BRIEF [22]

However expressed, biodiversity is related to the number of species

22. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 6 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

present in a defined space. Many natural and anthropogenic activities affect biodiversity. Two particularly important phenomena, biological evolution and extinction, can be understood as opposing forces that increase and decrease biodiversity, respectively.

Think about it some more…[23]

What do you think is the most useful way to express biodiversity?

How is the rise of a population of antibiotic-resistant bacteria a good example of evolution?

Is the well-known phrase "survival of the fittest" always consistent with evolutionary theory? If not, why not?

Are organisms that look and act like each other necessarily close relatives? Why or why not?

Why are small populations at elevated risk of extinction?

23. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

If extinction can be a natural process, why should we worry about it?

PART III

7. Earth's Natural Hazards

JASON KELSEY

The living and non-living systems we studied in Part Two are regularly subjected to stress. As we will see in detail in Part Three, although much of this pressure comes from human activity, powerful and disruptive natural phenomena also can bring about substantial changes to landforms, waterways, organisms, and ecosystems. The bottom-line message is: Earth has always been a risky place to live. Even before people and their technology started to spread across the planet, members of the biosphere were in near constant danger. In this chapter, we will briefly consider the origins and consequences of some of those natural hazards.

Key concepts

After reading Chapter 7, you should understand the following:

- How anthropogenic and natural hazards differ and why this distinction is important to environmental scientists
- The causes (review Chapter 3), consequences, and relevance of earthquakes and volcanic eruptions
- The many natural and anthropogenic phenomena that contribute to mass wasting

- How floods occur, can be predicted, and how human land-use decisions can make them more destructive
- How human activity may worsen the effects of biological stressors like pathogenic bacteria

Chapter Contents

7.1. CLARIFYING THE MEANING OF OUR TERMS

Stressors are assigned to one of two groups: those that are brought about by humans are said to be **anthropogenic**, and those that occur in the absence of humans are **natural**. This scheme is not perfect, and it might even induce an argument about the meaning of the word "natural" and how it applies to humans. However, it is a useful approach because it allows us to distinguish between hazards we can control and regulate and those we cannot. Any lingering values-based questions about our terminology will need to be addressed in venues more appropriate for such discussions than this book.

7.2. PHYSICAL STRESSORS

Here we take a quick look at some non-biological forces that shape the Earth and its organisms. Note that our coverage of these subjects is not meant to be exhaustive but rather is intended to provide an overview of and some background about important stressors and the effects they can induce. Many detailed sources of information about these topics are available, and you are encouraged to consult them for additional information (start with the web link provided at the end of this chapter).

7.2.1. Earthquakes: the first of two hazards caused by moving tectonic plates

Vertical shaking

Earthquakes are vibrations of the surface caused by movement along faults. A **fault** is often visualized as a crack in the crust, although it really is a surface, or plane, along which rocks can move relative to each other (Figure 7.1).

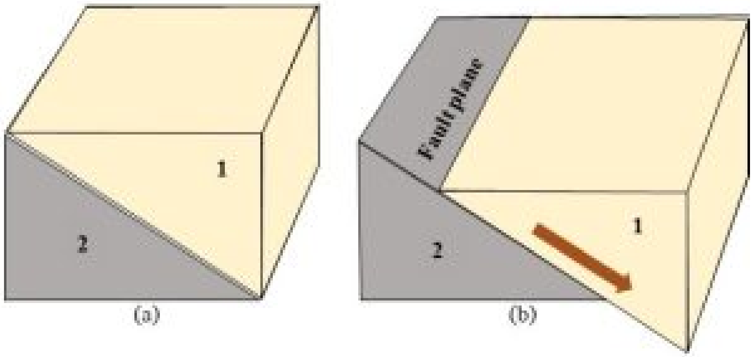


Figure 7.1. *Movement along a fault causes vibrations at the surface: a, prior to motion; b, block 1 has moved downward relative to block 2. Kelsey, CC BY-NC-SA.*

What causes faults to form? Recall that lithospheric plates are constantly in motion (Chapter 3). Movement in the vicinity of divergent, convergent, and transform boundaries applies gradual and persistent stress to the crust. Importantly, though, the rocks in such zones do *not* move constantly; instead, they absorb some energy while remaining stationary. When the accumulated stress eventually overcomes the strength of the materials holding the rocks together, sudden slippage occurs (known as rock strain). The amount of movement can range from nearly imperceptible to many meters in seconds or minutes. By and large, the severity of vertical shaking felt on the surface near a fault is proportional to the distance of displacement along that fault.

Three terms are applicable to any materials, including rocks. ***Stress*** *is related to the force applied to an object,* ***strength*** *is that object's ability to remain intact under stress, and* ***strain*** *is the change that occurs to the object* *once its strength is overcome.*

Effects

Damage at the surface. Vibrations have the potential to disrupt, damage, or destroy anything on the surface. We might be most familiar with the ways earthquakes affect buildings, roads, and other human structures (Figure 7.2), but they can also influence natural ecosystems. For example, trees growing on a slope could lose their support and fall over during an earthquake; as we know, secondary succession would likely follow such an event (Chapter 5).

Figure 7.2. *Buildings destroyed by an earthquake in West Sumatra, 2009. Ramzy Muliawan, CC BY.*

Landslides. As is described in the section on mass wasting (see below, after the section on volcanoes) earthquakes can trigger these important events.

Tsunamis. These are ocean waves induced by the sudden displacement of a large amount of water. They can be brought on by many forces, including a large asteroid or piece of continental crust splashing into the ocean or even localized high winds. An earthquake that occurs on the ocean floor can also start the process of tsunami formation. In short, movement along a submarine fault can cause water to quickly shift position and form a temporary bulge on the surface. It is worth noting that in the open ocean, the wave created by an earthquake is very difficult to detect—even passengers in a boat directly above a fault would likely be unaware of what is happening beneath them. However, as the bulge moves toward coastlines it changes from a subtle, horizontally oriented wave, to a predominantly vertical wall of water. By the time it strikes the beach, it can range from a few centimeters to 10 meters in height (Figure 7.3).

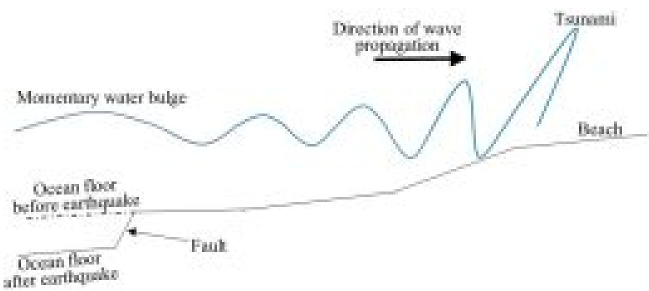


***Figure* 7.3.** *Simplified diagram of tsunami formation. Movement along the fault causes a sudden drop of the ocean floor and displacement of a large amount of water. Propagation of the bulge of water can lead to a tall wave striking the coast. Not to scale, in cross section. Kelsey*, CC BY-NC-SA.

Clearly, a very tall tsunami could be extremely destructive, affecting both human and natural systems. In 2004, for example, a powerful earthquake in the Indian Ocean led to a tsunami large enough to kill well over 200,000 people as well as pummel coastal ecosystems (Figure 7.4). See Box 7.1 for more about tsunamis, including how humans can detect and respond to them.

Figure 7.4. *Damage caused by two large tsunamis: left, a, ship and truck lifted by the wave, Alaska, U.S.A., right, b, devastation in southeast Okushiri Island, Pacific Ocean. US Dept of Interior, NOAA/NGDC, Public Domain (a and b).*

Box 7.1. Tsunamis: fascinating and deadly

Contrary to what you might expect, people often walk *toward* instead of *away from* the beach in the moments before a tsunami occurs. What accounts for this seemingly illogical behavior? In short, tsunamis are preceded by very rare and interesting phenomena that few people recognize as a warning. It starts when water suddenly and dramatically moves off the beach, straight out to sea. As a result, fish and other tasty organisms that ordinarily live under water in that zone are revealed. Many observers cannot resist the temptation to scoop up the fresh seafood just lying in the wet sand, so they walk into danger instead of running inland as fast as they can toward higher ground. Unfortunately, the propagating waves quickly bring back all of that missing water—plus much more—to the coastline in a potentially tall and deadly wave. Governments and scientists alike are working to minimize the numbers of deaths by improving tsunami prediction, preparedness, and awareness. It has been difficult to achieve these goals, though. Since tsunamis are difficult to spot in the open ocean, we often do not know where and when they will strike until the last minute. Technologically developed countries have employed specialized buoys off their coastlines that are sensitive enough to detect the signs of impending tsunamis. The United States, for example, has an array of warning buoys deployed in the Pacific Ocean. In case of emergency, notification and evacuation plans are in place. A public awareness campaign has been launched as well, with signs such as the one in Figure 7.5 posted in risky areas. The level of preparation in developing countries, though, is inconsistent and often insufficient.

Figure 7.5. *As part of a public awareness campaign in the U.S.A., signs such as this one are posted in coastal communities subject to tsunamis.* US NOAA, *Public Domain.*

Susceptible locations

By studying plate tectonics and historical records, geologists have been able to identify regions most susceptible to earthquakes. Generally, they occur at or near plate boundaries (review Figure 3.3 in Chapter 3). So, places like southern California, U.S.A., Iran, and Tibet, experience more earthquakes than, say, those in central North America, southern Australia, or central Africa.

Prediction

Given the relationship between earthquake occurrence and tectonic plate boundaries noted in the previous paragraph, it is fair to say we know *where* they are likely to occur. However, it is extremely difficult to predict *when* they will strike a particular location. Geologists use two strategies to try to evaluate the risk for future earthquakes.

Historical data and long-term predictions. An evaluation of previous earthquake activity combined with knowledge of the properties of the rocks in a region can allow us to calculate the likelihood that an earthquake of a given magnitude will occur within a defined time frame. Unfortunately, the best we typically can do is narrow it down to years or even decades. Information provided is still useful, though, because it allows people to make long-term plans and modifications to vulnerable structures.

Precursor phenomena and short-term predictions. Earthquakes can occur without any notice, bringing about massive death and destruction to unsuspecting people. Sometimes, though, earthquakes are preceded by warning signs evident just before they occur (minutes, hours or days). These precursor phenomena include visible changes to the surface in an existing fault zone, sudden production of radon gas near a fault, abnormal behavior of waves artificially sent by geologists into a fault, and even changes to animal behavior (although that last sign is not universally embraced as valid or reliable). If detected and properly interpreted, human communities can make life-saving preparations.

7.2.2. Volcanic eruptions: the second of our two tectonic hazards

Explosions

Volcanoes are vents from which molten and solid rock, ash, and gases from the interior of the Earth are ejected to the surface and into the atmosphere. They vary in size and shape and occur both above and below sea level. Figure 7.6 shows two common types of volcanoes, relatively explosive **cones** (also called stratovolcanoes) and **shields**, from which lava tends to flow more gently. The details are unnecessary for us, but you should note that differences in chemistry of the molten material, which depends on the type of plate boundary involved, determine the shape and behavior of the resulting volcanoes.

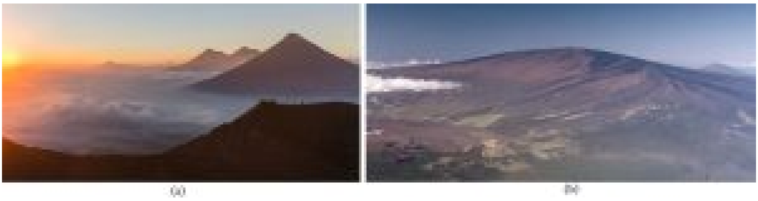


Figure 7.6. *Two types of volcanoes: left (a), stratovolcanoes are cone shaped; right (b), shield volcanoes are relatively flat. Christopher Crouzet, CC BY (a); USGS, Public Domain (b).*

Effects

Volcanic eruptions can bring on many consequences to human and environmental health.

Death of organisms. Volcanoes are somewhat famous for their lethality. Among the many mechanisms by which they can kill people and other organisms is through molten rock (temperatures can reach 1000 °C or higher). Few common materials will resist incineration when they encounter lava. So, a volcano could destroy a large climax forest, leaving behind a barren landscape ripe for primary succession (Chapter 5). But it doesn't stop there: their high-energy explosions can destroy nearby structures, and the large amounts of ash (essentially, tiny shards of glass that can injure lung tissues), and gases they emit can be deadly[1]. In other words, even if it does not bring on a fiery death, a volcano has other strategies it can use to kill you.

Change of climate. The tons of fine ash and aerosols ejected during some eruptions can disperse throughout the atmosphere and partially block incoming solar radiation (Figure 7.7).

1. Some of those gases are directly toxic; carbon dioxide, which also can be emitted, kills because it displaces dioxygen and leads to the asphyxiation of aerobic organisms.

Figure 7.7. *Ash is projected into the atmosphere during an eruption, upper (a) and settles on the ground, lower (b). USGS, Public Domain (a and b).*

As a result, average global temperature may decline after a large eruption. The eruption of Mount Pinatubo in the Philippines in 1991 is one example of this phenomenon: measurable cooling of the entire planet occurred between 1991 – 1993. We will see more about the many factors influencing Earth's climate in Chapter 14. Ash can also directly affect human civilization. Major eruptions in Iceland in 2010 halted air travel in Europe and North America because of the dangers posed by the materials they released into the atmosphere.

Contributions to acid precipitation formation. Some of the chemical compounds released through volcanoes can be transformed into acids in the atmosphere. So, following an eruption, both sulfuric and hydrochloric acids can be produced and change the chemistry of water falling to the surface. Acid precipitation can also be caused by anthropogenic activities, a topic we will return to in Chapter 14.

Susceptible locations

Volcanoes are usually found near divergent and certain convergent (in **subduction zones**) tectonic plate boundaries. For example, the large rim of the Pacific Ocean including Alaska, Japan, Indonesia, and Peru is sometimes called the "Ring of Fire" because it is defined by multiple plate edges and is the location of concentrated volcanic (as well as earthquake) activity (Figure 7.8).

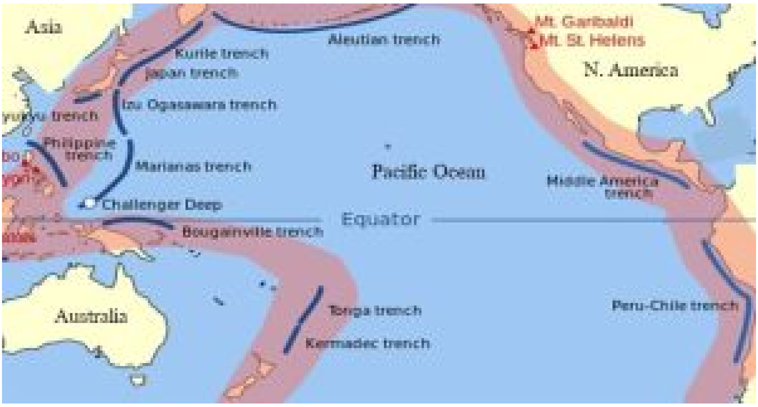


Figure 7.8. *The Pacific "Ring of Fire" is a zone of concentrated volcano and earthquake activity. Gringer (talk), Public Domain (modified by Kelsey).*

Volcanoes in places such as Iceland, Italy, and Tanzania also result from interactions among plates. In rare but notable cases such as the Hawaiian Islands, they occur far from plate boundaries.

Prediction

Volcanic eruptions have the potential to cause wide-spread death and destruction, so we are keenly interested in predicting and preparing for them. Like we see for earthquakes, both long-and short-term predictions are possible. Their locations near relevant plate boundaries, as well as their generally prominent appearance (e.g., a large, cone-shaped mountain) facilitate our ability to identify volcano-prone regions. In the near term, a volcano will provide many clues that an explosion is imminent, allowing for appropriate evacuations and other measures. Geologists (specifically, people known as volcanologists) study the sites of past activity and monitor gas emissions, changes in the shape of the surface, and other properties to assess the risk of impending eruptions.

7.2.3. Mass wasting

Downhill transport

Defined succinctly, this term refers to the gravity-driven movement of Earth's materials from high to low elevation. Multiple variables are used to distinguish among the many specific types of phenomena grouped together here, including the rate of movement (ranging from nearly instantaneous to very slow), the role of water (liquid and frozen), and how dramatically the materials change form during their transport. So, for example, a **landslide** involves the rapid movement of relatively dry and intact rocks and soils down a sloped surface, whereas during a **fall,** rocks are not in contact with the slope as they move downward. **Creep**, as the name vividly suggests, is the very slow, all but imperceptible movement of generally intact soils, rocks, and even the plants growing on them along a slope. Figure 7.9 provides a simplified diagram of these three movements.

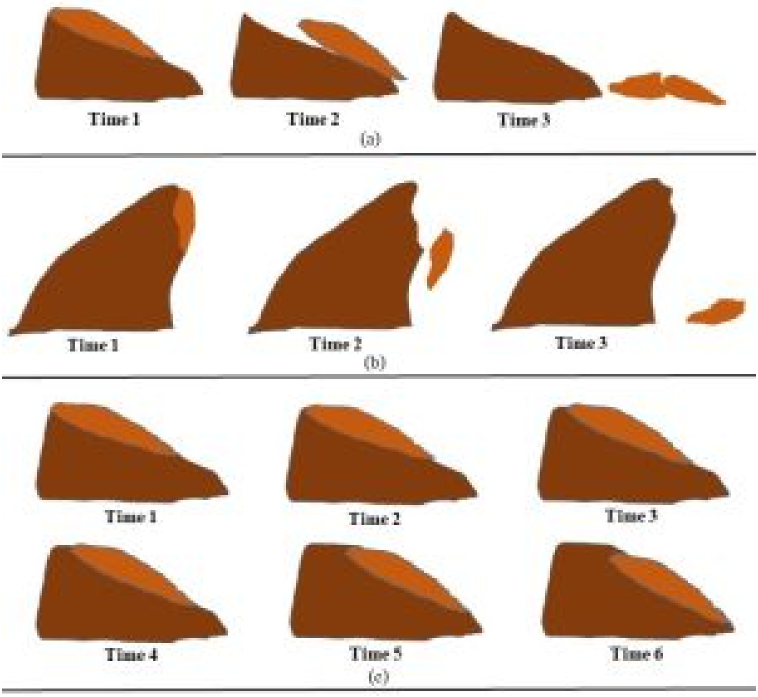


Figure 7.9. *Simple models of three common types of mass wasting: slide (a); fall (b); creep (c). In all, the light-colored piece of the landscape moves downhill due to gravity. The time elapsed between Time 1 and Time 3 in the slide and fall (a, b) is seconds or minutes. The time elapsed between Time 1 and Time 6 in creep (c) could be decades or more. In cross section. Kelsey,* CC BY-NC-SA.

Note that several other types of mass wasting phenomena are possible, driven by different environmental conditions and affected by the properties of the rocks and soils present.

More on factors affecting mass wasting

Mass wasting occurs due to gravity—everything is pulled downward by this relentless force. Several other factors also play roles, influencing how readily materials will move downhill.

Presence and form of water. Liquid water can act as a lubricant and facilitate mass movements. So, a large and intense rainfall event can trigger a mud- or landslide. An icy layer under loose soil, common in polar regions, can also increase the likelihood of downward movement.

Presence of vegetation. Plants can help keep soil intact and therefore stabilize a slope for two reasons. First, the network created by their roots holds soil in place against the forces that would remove it (e.g., weathering by wind and water—see Chapter 3). We will return to this topic when we encounter soil erosion in Chapter 9. Second, their above-ground tissues like leaves block incoming rainfall, protecting the soil beneath from direct impact. As a result, the soil tends to stay in place. Natural fire and wind that remove vegetation, as well as human activities such as land clearing for farms (Chapter 9) and logging (Chapter 12) can bring on mass wasting. If the change occurs rapidly, as when an entire stand of trees is removed at one time (i.e., **clear cutting**), the risk of movement increases.

Steepness. A gentle hill will be less susceptible to mass wasting than will a cliff or steep slope. Natural erosion and human forces such as construction of roads and buildings can change the shape and angle of a slope (Figure 7.10).

Figure 7.10. *Construction and erosion have combined to steepen this slope, making it increasingly unstable with time. Obviously, the house is in serious trouble. Wilson44691, Public Domain.*

Triggering event. An unstable slope might maintain its integrity for an extended period. However, the vibrations of an earthquake might be just enough to dislodge the materials and send them downhill.

Susceptible locations

Mass wasting occurs anywhere there is a difference in elevation and gravity is active. In other words, it is not restricted to plate boundaries like we saw for the tectonic stressors—it can shape ecosystems all over the world. Now, it is important to realize that the specific type of event that occurs varies from place to place as a function of prevailing conditions. Movements in polar regions, for instance, are often affected by the presence of ice. Those in tropical regions are likely hastened by flowing water whereas desert landslides occur on slopes populated by minimal vegetation.

Effects

As we saw for the tectonic hazards above, both natural and human systems are affected by mass wasting. A common ecological consequence is seen in trees that develop curved trunks when they try to grow on a slope subject to creep. As their roots move downhill, they compensate by curling their crowns (i.e., their tops) toward the sun (Figure 7.11).

Figure 7.11. *The effect of creep: note how the bases of some trees are curved (highlighted by red arrows). Kent G. Budge, CC BY.*

In fact, the presence of J-shaped trees on the side of a hill provides a clue that the slope is not as stable as it might appear. Eventually, gravity always wins the battle, and a climax community will be damaged or destroyed, making way for secondary succession. Death and destruction of human communities can also be the result of rapid events such as a sudden landslide or fall (Figure 7.12).

Prediction and prevention

Figure 7.12. *Sudden mass movements can be devastating: houses are buried by a slide.* USGS, *Public Domain.*

An examination of the properties of the rocks and soils in an area as well as factors like slope angle, climate, and land use by people can help geologists calculate the risk of mass wasting in a location. Unlike earthquakes and volcanic eruptions, humans also have some degree of control over the occurrence of mass wasting. Clearly we cannot change gravity or weather, but we do have the ability to minimize activities that reduce slope stability (land clearing and poor construction choices, as noted previously). Alternatively, humans often attempt to minimize mass wasting *after* a poor land-use decision has been made. Retaining walls are one strategy used to slow the onset or effects of a landslide after, say, a road has been built on the side of an unstable hill. Unfortunately, they do not always work as hoped (Figure 7.13).

Figure 7.13. *A failed retaining wall. Ray Folwell,* CC BY-SA..

7.2.4. Extreme weather

Both human and natural systems are affected by the normal climatic conditions of their environments, which can include large seasonal fluctuations in temperature, moisture, and other conditions. As we know, organisms must adapt to these and other properties if they are to survive. However, relatively rare events such as hurricanes, thunderstorms, tornadoes, heat waves, and droughts can be sources of severe stress capable of damaging or destroying existing communities. Consider, for instance, that high winds associated with any of these phenomena could bring down trees, potentially setting the stage for secondary succession in a climax forest (Chapter 5).

7.2.5. Water-related hazards

Water is a powerful agent of change that can affect living and non-living systems in two distinct ways.

Weathering and erosion

As we learned in Chapter 3, water plays very important roles in the shaping of Earth's surface. First, it is instrumental in the **weathering** of rocks and soils. Second, water can contribute to **erosion**, carrying away the products of weathering. These two processes can be accelerated when humans disrupt and destabilize soil, for example through agriculture (Chapter 9) and logging (Chapter 12). Although weathering and erosion do bring on some negative consequences in those cases, they generally do not create immediate hazards. The rapid removal of sand and soil beneath buildings and bridges, however, can lead to loss of both life and property. Structures on beaches and barrier islands are particularly vulnerable because those sandy features are inherently unstable. Day-to-day wind and waves along with periodic intense storms act to change the shape of coastlines and the human communities built there. In fact, hurricanes and similar events can remove entire beaches in a matter of days. Natural ecosystems in affected areas also can undergo dramatic transformations, stimulating succession and recovery.

Flooding

Input vs. output. There is not much mystery here: water will accumulate in a system if it enters faster than it is removed (Chapter 2). At some point, a reservoir's capacity will be exceeded, and surrounding areas that are normally dry will be inundated. Consult Chapter 4 for more about pathways and reservoirs of water.

Effects.

Death of terrestrial organisms

Ecosystems characterized by predominantly dry conditions can undergo severe changes if water remains in them for an extended period. The large trees in a climax forest, for example, will likely die if they sit under water following a flood. These stands of dead trees will ultimately be replaced by **secondary succession**. See Figure 7.14.

Figure 7.14. *The trees in this forest were killed by prolonged submersion in water. Compare to Figure 12.2. Fredlyfish4, CC BY.*

Spread of water-borne diseases

Certain pathogenic bacteria and protozoa (Chapters 3) from polluted waters can mix into human drinking water after flooding. Somewhat ironically, one of the outcomes of flooding is a profound lack of useable drinking water. If standing water persists for several days, people living in affected areas are at risk of contracting cholera and other diseases (Figure 7.15).

Figure 7.15. *Diseases like cholera spread easily in flood waters like those pictured here. Donavanik, CC BY.*

Property damage

Although clearly not an ecological concern, flooding can lead to massive loss of housing and other buildings as well as damage to crop and other valued lands.

Susceptible locations. Many factors influence the likelihood of flooding in an area.

Topography and climate

Perhaps the most obviously relevant factors affecting flooding are elevation, proximity to surface reservoirs, and amount of precipitation received. Low-lying homes near a river or ocean in areas subject to heavy rainfall are more likely to suffer flood damage than those on the tops of mountains in dry areas.

Surface properties

In addition to the important factors noted in the previous paragraph, severity and frequency of flooding is influenced by the types of soils and rocks present in an area. Permeability to water varies enormously among materials, so even if topography, climate, and other environmental factors are equal, the likelihood that rain will lead to a flood could still differ in two locations. Why? Recall from Chapter 4 that precipitation can experience two fates, infiltration (vertical downward movement) and runoff (essentially horizontal movement across the surface). Remember also that the relative importance of these two pathways is inversely related, meaning as one carries more water, the other carries less. If water cannot penetrate the surface (for example, if it encounters solid rock or tightly packed clay), a larger percentage of it will flow horizontally than vertically. Overland flow (i.e., runoff) tends to be more direct and efficient than is percolation underground, so if it dominates, nearby waterways will flood more readily than if infiltration is comparatively important. How much water gets to a reservoir within a given period is critical. If a fixed volume moves quickly to a stream, reaching its destination essentially all at once, flooding is more likely than if the arrival of that same volume is spread out over more time (again, using systems analysis, when rates of input exceed rates of output, flooding is the result). Both natural and human activities (described below) affect surface permeability and the speed of water movement downhill.

Vegetation present

Put simply, plants tend to slow the rate of runoff, so the risk of flooding is relatively high in barren areas. As is the case with surface properties, the amount of vegetation can change for many reasons, including those related to human land use (see next item).

Human activity and land use

It would not surprise anyone to hear that people modify what happens on Earth; supporting evidence is not difficult to see if you step outside. In the case of water storage and movement, two modern phenomena are particularly important. First, consider a river that is downhill from a suburban neighborhood. You might imagine that in the days before humans, the area was forested and dominated by permeable soils and rocks (Figure 7.16., left). After the impermeable roads and buildings appeared, runoff become far more important than it once was, and flooding of the river and downstream regions became more frequent (Figure 7.16., right).

Figure 7.16. *Effect of land use on surface properties. An unpaved hillside provides many obstacles to flowing water and slows the rate of runoff and flooding, left (a); land clearing and paving vastly decrease infiltration and increase the rate of flooding downhill, right (b). Nigel Chadwick* CC BY-SA *(a); Mick Lobb* CC BY-SA *(b).*

This common phenomenon has had a profound impact on human and natural communities alike. Second, the clearing of land for agriculture (Chapter 9), forestry (Chapter 12), and other purposes has also increased flood risk along many waterways. If plants in a field or forest are removed, water will move more rapidly along the surface than it once did. In short, although flooding is clearly a natural hazard, humans can increase its effect.

Prediction. As we saw for the other natural sources of stress on this list, we can assess the likelihood of flooding in both the near and far term. Weather forecasts of heavy rainfall can provide information about the immediate threat (i.e., days or weeks) of flooding, and appropriate preparations can be made in response. In the long run, analysis of the history of a region can be used to generate a flood zone map showing the probability that areas will flood. Typically, such a map is divided into 10-, 100-, and 1000-year flood zones, in essence, the zones with a 1 / 10, 1 / 100, and 1 / 1000 chance of flooding each year (Figure 7.17).

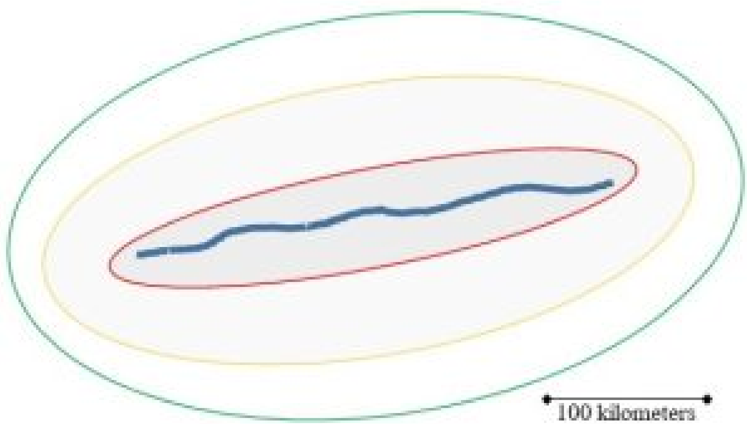


***Figure* 7.17.** *A theoretical map of a river and its flood zones. Elevation generally increases in all directions away from the center (in blue, the river valley). Areas enclosed by the red, yellow, and green ovals are the 10-year, 100-year, and 1000-year flood zones, respectively. The scale is hypothetical. Kelsey*, CC BY-NC-SA.

As these maps are based on statistics and probability, they are subject to some uncertainty. Furthermore, and very importantly, they predict events for *each year*. An occurrence of a large flood now does not guarantee anything about the future. So, for example, just because houses in the 100-year flood zone are under water in 2023 their owners should not assume they are free and clear until 2123! The chance that a flood of the same size will occur in 2024 is still 1 / 100.

Prevention. Clearly, we cannot change the weather and other natural processes contributing to flooding. However, we have the power to minimize both the occurrence of and damage from flooding in several ways.

Land use choices

The seemingly short and simple message here is: do not live in flood zones. The maps noted in the paragraph above are intended to inform development decisions, although people still build in risky areas (Box 7.2 explores "why??"). Generally, property owners in flood-prone regions are subject to laws requiring the purchase of additional (and costly) hazard insurance. Still, since public resources (i.e., taxpayer dollars) are often used to clean up after floods, there is considerable debate about how or even whether any kind of development should be allowed in such zones.

Box 7.2. Why build in a flood zone?

You might wonder what would motivate someone to build a home or other structure in an area that is likely to flood in the foreseeable future. Well, some of those locations are attractive enough for people to take the risk, although it is very often a move that leads to loss of property, money, and life. Coastal zones, for example, have obvious appeal for those who like to live or visit the ocean. As we saw in the section on erosion above, though, beaches and barrier islands, popular places to construct hotels and houses, are very unstable. They are also vulnerable to flooding. Floods of inland rivers can cause massive damage and destruction as well, yet people live in these risky areas nonetheless. To understand why, it is helpful to examine a diagram of a river and its immediate surroundings (Figure 7.18).

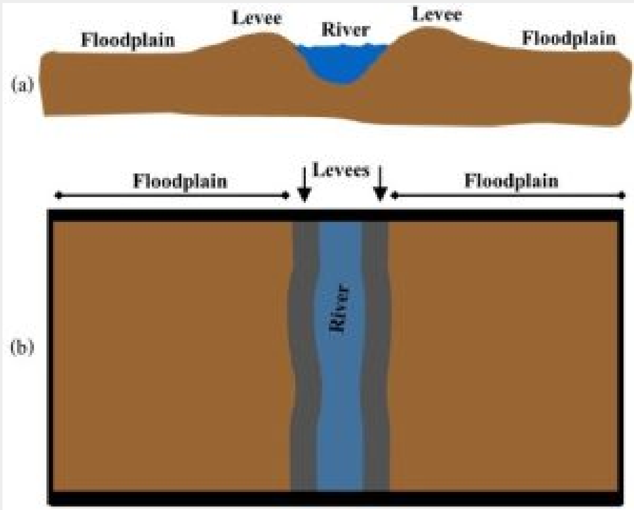


Figure 7.18. *Two views of a portion of a theoretical river, levees, and flood plains: top (a), cross section and bottom (b), map view. Once water rises above the height of the levees, it can easily flow across the flat floodplain. Here, the river is shown at normal height, that is, the water is confined to its channel. Kelsey, CC BY-NC-SA.*

Under normal conditions, a river is confined to its channel. However, when more water enters the system than can be contained, there is overflow onto the adjacent land. This area on either side of a river is called a **flood plain** and is typically very flat and fertile. How did it get that way? The short answer is: repeated flooding. Each time the water exceeds the capacity of its channel it flows across the plain, depositing layers of sediments that were previously carried in the river. So, flooding creates a zone that is great for building houses because it is flat and for farming because its soil is full of nutrients (more about factors affecting farming is described in Chapter 9). Note that in addition to creating a flat plain, flooding also produces **natural levees** on either side of a river, structures that run parallel to the channel. These are made of coarse sediments that get piled up on the banks by flowing water. Each time there is a flood, levees—essentially walls of soil and rock—get larger and the height to which the water must rise to enter the flood plain increases. Humans can modify levees to make them even higher, although under the intense stress of a strong hurricane or other event they can fail, allowing a large amount of water to quickly enter the flood plain (this is a common cause of flood-related damage to humans and their structures).

Build structures to protect

Sometimes, communities try to reduce risks by constructing walls and similar structures. For example, an artificial levee or a dike can be built to keep water from damaging homes in flood-prone areas. Unfortunately, these walls tend to fail under extreme stress from very high and fast-moving water, in other words, when they are most needed. Additionally, their presence can encourage people to live in risky flood zones that might otherwise be considered uninhabitable. For example, New Orleans, Louisiana, U.S.A. suffered severe damage in 2005 when rising waters associated with the very powerful Hurricane Katrina broke through levees built to only withstand milder storms.

7.3. BIOLOGICAL STRESSORS

Stress from biological entities adds to the pressures exerted by non-living hazards. Chapter 5 describes the details of the three deleterious interactions among organisms listed below, so we will not dwell on them again here. You should keep in mind that, as we saw with the physical stressors above, living hazards are extremely important to members of today's biosphere and also challenged and shaped life on Earth long before humans arrived.

7.3.1. Competition

Competitors are denied easy access to limited resources by the presence of each other. Ultimately, the weakest individuals die as a result of this phenomenon. Both participants are harmed, or endure stress, during competition (Chapter 6).

7.3.2. Predation

Here, a predator receives benefits from eating its prey. This type of stress affects both participants, though, because only the most fit individuals—that is, relative to both catching food and surviving predation—will endure long enough to produce offspring (Chapter 5).

7.3.3. Parasitism

Like predation, this interaction is characterized by a winner and a loser: a smaller parasite harms a larger host by causing disease or other adverse outcomes. Disease can kill directly by bringing on lethal changes or indirectly by weakening a host. In the latter case, since its fitness is reduced, the affected host is less likely to reproduce. See Chapter 5 for more details about this important stressor.

THE CHAPTER ESSENCE IN BRIEF [2]

The distinction between natural and anthropogenic stressors is important because humans have little or no control over the former but a great deal of control over the latter. Earth's hazards may threaten modern civilization, but informed decision-making can minimize the potential damage caused by tectonic, water-based, and biological stressors.

---

2. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 7 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[3]

Why do we distinguish between anthropogenic and natural hazards?

Which of our natural hazards worries you the most? What variables might affect your answer?

How does human land-use planning influence the severity of the damage caused by the natural hazards in Chapter 7?

Additional suggested reading on physical stressors

The U.S. Geological Survey has an excellent web page you can explore for a lot more information about Earth's hazards.

3. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

8. Human Population as Environmental Stressor

JASON KELSEY

Chapter 7 described some natural sources of stress. Now we begin to study ways humans can adversely affect organisms and ecosystems (i.e., anthropogenic sources of stress). In Chapter 8 we consider the causes and consequences of rapid human population growth and increased demand for resources. Upcoming chapters will build on those foundations, and we will see how limitations on available materials like fertile soil, space, and clean water and air are linked to the number of people on Earth and their standards of living.

Key concepts

After reading Chapter 8, you should understand the following:

- How and why the current rate of growth of the human population is much higher than that seen in the past
- What is meant by exponential growth
- Different ways to express the growth of a population

- Why demographic shifts in populations are important
- The factors that influence different rates of growth and resource use in different countries
- What is likely to happen to the size of the human population in coming decades
- How we might adapt to or mitigate the effects of a growing population and increasing expectations of resource availability

Chapter contents

8.1. A BRIEF HISTORY OF HUMAN POPULATION GROWTH

8.1.1. Numbers and growth rate were low until the middle 1800s

Evidence suggests modern humans have been present on Earth for about 200,000 years. For nearly all that time, the size of the population was low relative to that seen today: most researchers hold that the number of people in the earliest days was only in the tens of thousands. By about 10,000 BC, it had likely increased to a few million. The size of the human population then grew slowly, up to 1 billion by the year 1800[1].

8.1.2. Population size has increased dramatically since the 1800s

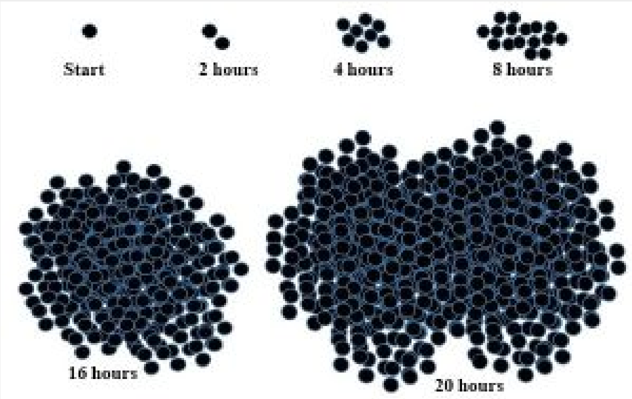
In about 1850, the rate of growth began to rise. By 1910, it became constant, 1 to 2% each year (the 2022 rate was 1.03%[2]). Put another way, the human population has grown **exponentially** during the past century or so (see Box 8.1 for more about exponential growth). This is a concept we encountered during our discussion of positive feedback (Chapter 2).

Box 8.1. Exponential growth: filling up in a hurry

As a population grows, **positive feedback** will cause its size to increase at an ever-increasing rate. **Exponential growth** is one result of positive feedback, a phenomenon that has driven the human population to

1. Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina and Lucas Rodés-Guirao. 2019. World Population Growth. Ourworldindata.org
2. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

expand rapidly during the past century. **Bacteria** provide an even more dramatic demonstration of the power of exponential growth than do humans. Imagine we add a single microscopic cell to the middle of a **petri dish** and then watch what happens. If our hypothetical species doubles in size every two hours (a reasonable rate for bacteria), it would not take long for it to grow into a population large enough to visibly fill the entire surface of its artificial home. How can this be? Well, one cell splits into two, then two become four, then eight, and so forth. In just 20 hours one cell can become 1024 cells, and over one million cells would be present after 40 hours! The figure below models how fast all of this occurs. Note that each dot represents one cell and that all doublings are not pictured. You should also realize a petri dish would contain a fixed amount of space and nutrients, meaning the bacteria would eventually outgrow their resources and die.



Putting it rather informally, with growth, more people become capable of producing more people, and the population gets ever larger. Figure 8.1 is a graph of human population with time. Among other important trends, it shows how different were growth rates before and after 1850. The number rose from 300 million in the year 1 AD to 1 billion some 1800 years later[3], but then another 7 billion were added in only the next 222 years.

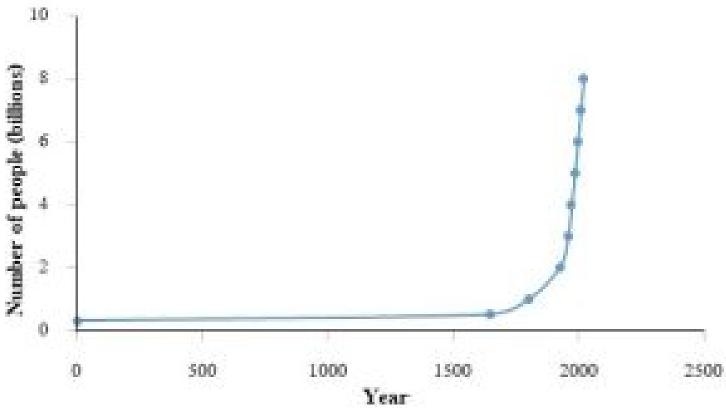


Figure 8.1. *Human population growth from 1 AD to the present. Note how the shape of the curve has changed dramatically during the past century. Kelsey,* CC BY-NC-SA.

Table 8.1 provides another way to envision growth: it quantifies the increase from 1804 to the present in increments of 1 billion people and includes projections to the end of this century (data for the years 2050 and 2100 come from the United Nations[4]). Note that,

3. Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina and Lucas Rodés-Guirao. 2019. World Population Growth. Ourworldindata.org

4. United Nations Department of Economic and Social Affairs,

although the overall number of people is expected to go up until at least 2100, the rate at which that change occurs will likely slow. Box 8.2 summarizes some important facts gleaned from Figure 8.1 and Table 8.1.

Table 8.1. *Growth in increments of 1 billion people.*

Year	Population (billions)
1804	1
1927	2
1960	3
1974	4
1987	5
1999	6
2011	7
2022	8
2050	9.7
2100	10.4

Box 8.2. Some important facts about the human population

The figure and table above contain a lot of data about the size and growth of Earth's human population. Some

Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

of the most noteworthy conclusions and trends are further summarized here.

1. Total population in fall of 2022: 8 billion

2. Annual growth rate in 2022: 1.03% (a net addition of over 80 million people per year)

3. Annual growth rate during the 20th Century: a range from 0.8% (1900) to 2.1% (1962)

4. Net number added during the 20th Century: 4.5 billion (i.e. an increase from 1.6 to 6.1 billion)

5. Years to double in size from 1 to 2 billion: 123 (the year 1804 to the year 1927)

6. Years to double from 4 to 8 billion: 48 (the year 1974 to the year 2022)

8.2. MEASURING GROWTH

Demographers, scientists who study the characteristics of populations, can calculate changes to the size of the human population using the two methods described here.

8.2.1. Growth rate

Growth rate is an expression of the net change in the total number of living individuals within a given period (generally per year). By convention, this information is expressed as number per thousand people, in a fraction such as 9/1000 or -2/1000. Growth rate is calculated using simple mathematics:

Growth rate = birth rate – death rate.

For instance, in 2020, the global birth rate was 18.1 / 1000 and the death rate was 7.7 / 1000[5]. Put another way, for every 1000 people in the world on January 1, 2020, 18.1 were added by December 31, 2020; 7.7 of every 1000 died during the same period. So, net growth for 2020 =

18.1 / 1000 – 7.7 / 1000 = 10.4 / 1000.

What does this answer mean? For every 1000 people at the beginning of the year, there were 1010.4 people at the end of the year. Note that net growth would be affected even if only *one* of the terms changed. Imagine if birth rate remained at 18.5 / 1000, but death rate declined to 6.7 / 1000. In such a case, growth rate would increase to 11.8 / 1000. Similarly, if inputs and outputs changed by the same amount, say, each was reduced by 1, then growth rate would be constant (i.e., 10.4 / 1000). Of course, the relative size of different age groups would change even as the total number was unaffected. In our example involving constant birth rate with declining death rate, there would be more older people and fewer younger ones (a situation that is increasingly affecting many nations, as we will see later). Further complicating matters, since rates of birth and death, as well as immigration and emigration, are not the same everywhere, growth rates in some nations are higher than the global average and some are lower (even below 0 where populations are contracting). We will return to a discussion of Earth's heterogeneous population growth shortly.

5. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

8.2.2. Total fertility rate (TFR)

Unlike growth rate, which applies to the entire human population, TFR is the number of children born to an average woman during her childbearing years (formally defined for the purposes of this calculation as ages 15 – 49). The global fertility rate for 2020 was 2.42, although, as we noted for growth rate, it varied considerably among nations[6]. Three additional points merit brief attention here. First, unlike growth rate, this measurement is a direct way to quantify the number of births per woman and does not consider the number of men. Second, TFR in a year can be used to project future growth of a population. For example, since the fertility rate in the United States was 3.5 in the 1960[7], an average woman born that year gave birth to 3.5 children by the time she turned 49; the total growth of the population during that period could therefore be predicted ahead of time. The TFR for a woman born in the U.S. was 1.4 in 2020[8], suggesting a lower rate of growth in the future. Finally, since two parents are required to produce each child, a fertility rate of 2.0, known as the **replacement rate**, would in principle yield a stable population. Each set of two parents is exactly replaced by their two offspring, so the total number of individuals does not change. In practice, the actual global replacement rate is generally understood to be approximately 2.1 for several reasons, including the fact that some people die before they produce any offspring.

6. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/
7. Max Roser. 2017. Fertility rate. Published online at ourworldindata.org
8. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

Due to differences in life expectancy and other factors (more below), replacement rates in individual countries can be higher or lower than that average of 2.1.

8.3. FACTORS AFFECTING POPULATION SIZE

Whether the size of the human or any population increases, decreases, or remains constant ultimately depends on the relationship between rates of input and output of individuals. Those additions and subtractions are, in turn, influenced by a number of complex forces.

8.3.1. Inputs

Globally, births are the only source of human beings. Locally, though, immigration also adds people to the population of a country. Rates of both births and immigrations, therefore, should be added together to calculate total growth rate in any area.

Birth rates

The annual number of live births per thousand people has varied with time, increasing and decreasing, and differs among regions and countries for several reasons.

> *Recalling systems analysis (Chapter 2), we can say that*

the number of people in the system (Earth) has increased during the past centuries because inputs (births) have outpaced outputs (deaths).

Infant mortality. Defined simply, this term relates to the rate at which pregnancies end in a stillborn baby; that is, lower numbers of infant mortality indicate higher numbers of live births. Globally, this rate has declined dramatically during recent decades from 93 / 1000 births in 1990 down to 37 / 1000 births in 2020[9]. Individual countries have changed to different extents, though. For example, per 1000 births, it dropped from 11 to 6.3 in the United States, 6.3 to 2.5 in Japan, 7.7 to 3.0 in Monaco, and 146 to 29 in Bangladesh during that period[10]. This decline in mortality largely has been the result of increased access to quality medical care.

Cultural considerations. A detailed exploration of this topic is beyond our scope, so we will only make a few brief comments here. Broadly, it is fair to say that countries with a tradition of large families tend to continue that trend with each subsequent generation, but many factors affect attitudes and fertility. Differences in values, social structure, religion, education, and level of development are among the variables influencing birth rate. For example, as rural, agricultural economies transition to urban, industrial ones, birth rates generally decline. Also important are levels of education and empowerment of women: as these increase, sizes of families very often decrease.

9. 2022. Child Mortality of the Past. ourworldindata.org

10. 2022. Child Mortality of the Past. ourworldindata.org

Immigration

The number of new individuals present in a country increases each time a new person arrives from an outside place. Like birth rates, immigration rates are unequal among Earth's nations and are influenced by many forces including the perception that entering a host country will improve economic opportunities, safety, health, and freedom. In other words, countries most affected by immigration are those that are seen to enhance the lives of people who move to them.

8.3.2. Outputs

Death is ultimately responsible for removing human beings from the global system, but both mortality within and emigration from individual countries must be considered to determine local rates of output.

Death rates

Each human being born will die eventually. However, the length of time a typical person lives—commonly referred to as **average life expectancy**—is an important variable affecting overall population size and its rate of growth. Notably, if the rate at which people die goes *down*, the total number of people alive at any one time goes *up* (review equation 1 above to visualize the effect of declining death rate on total growth rate). The difference in growth rates before and after 1800 are in large part a consequence of declining death rates with time (review Figure 8.1 and Table 8.1).

Many factors contributed to the slower output of humans from Earth's system during the past 150 years. Although these changes certainly did not affect all regions or people equally, they still led to improvements in average survival, and life expectancy and can be linked to observed increases in population size. Of course, unexpected stressors and events can complicate our calculations. See Box 8.3 for information about the way the COVID-19 pandemic affected human population.

Box 8.3. What about COVID?

The effects of the pandemic of 2020 (and beyond!!) on population growth have yet to be completely understood, but some tentative conclusions can be drawn. First, the global growth rate of 2020 was less than 1 % / year for the first time since 1950[11]. Second, average global life expectancy dropped between 2019 and 2021, although the story is nuanced. Some areas—for example, Mexico, Russia, and Lebanon—saw a 4-year-decline in predicted lifespans of babies born in 2021, whereas others—for example New Zealand and Australia—actually saw an increase in life expectancy of about 1 year[12]. Third, although the numbers are difficult to determine with a high degree of certainty, The World Health Organization (WHO) estimates that the pandemic

11. United Nations Department of Economic and Social Affairs, Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

12. United Nations Department of Economic and Social Affairs, Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

led to about 15 million excess deaths (i.e., deaths that would not have occurred without COVID-19) between January 2020 and December 2021. Note this number includes deaths directly and indirectly attributed to COVID and is therefore higher than the 6.6 million officially reported to the WHO[13]. Fourth, the way the pandemic has influenced fertility is rather difficult to assess. On the one hand, isolation can limit access to contraception, but, on the other, economic stress tends to depress fertility. At this point, no obvious trend can be identified[14]. Demographers continue to study the problem.

Improvements in medicine. Human healthcare has dramatically improved during the past century. For example, advances in the treatment of cancers, infections, heart disease and numerous other life-limiting conditions have contributed to accelerated population growth by slowing the rate of death. Three other critical changes are briefly described here.

Knowledge about causes of infections

As noted in Chapter 3 (especially Box 3.5), human understanding of infectious disease grew substantially in the mid and late 1800s. Up until that time, the role of microorganisms in causing conditions

13. World Health Organization, 2022. Global Excess deaths associated with COVID-19, January 2020-December 2021

14. United Nations Department of Economic and Social Affairs, Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

such as cholera and tuberculosis, to name just two important historic and deadly diseases, was not known. Consequently, simple precautions we take for granted today such as bathing, disinfection of medical instruments, and other steps that slow the spread of germs from person to person were not employed. As remarkable as it might seem to us, with few exceptions, physicians did not even routinely wash their hands between patients until at least the late 1800s! The study of germs and ways to combat them led to appreciable changes in the health (and likelihood of survival) of both sick and healthy people.

Antibiotics

We learned in Chapter 5 that the discovery of antibiotics in the 1920s profoundly changed the way we treat cholera, strep throat, syphilis, and numerous other diseases caused by bacteria. Drugs such as penicillin cured many people who would have likely died in previous years and contributed to an increase in the size of the human population. As we know, though, there are risks with antibiotic use, notably the development of antibiotic-resistant bacteria (Chapter 5).

Vaccinations

The administration of vaccines to protect against diseases including smallpox, polio, measles, rubella, and tetanus made a profound contribution to life expectancy during the 20th Century. More recently, the rapid development and distribution of vaccines against the virus responsible for COVID-19 similarly protected millions of people against serious cases of the disease. Review Boxes 1.2 and 2.1 for more about the benefits of vaccines.

Improvements in water quality. This factor is related to the previous one, as we learned in Chapter 3. An understanding of the ways biological, chemical, and physical pollutants can affect the health of people consuming them (mid and late 1800s) helped usher in better management and protection of drinking water supplies. We will return to the topic of water quality in Chapter 11.

Improvements in nutrition. Agriculture has been around for about 12,000 years (Chapter 9), but increases in the availability of food due to farming and distribution systems only provided much more consistent access to essential nutrients during the past century or so. These changes drove better growth and development, immune function, and prenatal care for many, if not all, peoples.

Improvements in safety. The list of ways human lives have become more safe with time is quite long. Keeping things appropriately brief for our purposes, various laws and trends affecting everything from cars to homes to workplaces to highways have reduced injuries and deaths. Strict regulations on food and drug safety have similarly enhanced both quality and length of lives and contributed to population growth (more about food safety can be found in Chapter 15).

Emigration

As we saw for immigration, this phenomenon varies from country to country and can profoundly affect population growth rate. Among the many factors affecting it are stability of government, civil war, disease, drought and other adverse environmental changes, and poverty. In simple terms, people tend to exit countries undergoing turmoil and enter countries assumed to provide the possibility of a better life.

8.3.3. All inputs and outputs affect local growth rate

Note that equation we saw in section 8.2 can be modified to account for all inputs and outputs in a specific area. Thus,

Growth rate of a region = (births + immigration) – (deaths + emigration)

8.4. DISTRIBUTION OF PEOPLE, GROWTH, AND RESOURCE USE

Human population growth rate and density vary among regions and nations and have also changed with time. Additionally, standards of living, values, and resource use also differ—often by a great deal—from place to place. Here we will briefly examine the major trends and possible explanations for the disparities we observe.

8.4.1. Human population distribution and wealth are heterogeneous[15]

Population size and density vary enormously among (and within) regions, as is well known. More than half of Earth's 8 billion humans live in Asia, and two nations, China and India, each are home to around 1.4 billion people. The United States, with a population of about 340 million, is a somewhat distant third. On the other end of the spectrum are countries such as Saint Barthelemy, Montserrat, and Vatican City with populations of 7,103, 5,414, and 1,000, respectively. Realize that the most populous nations are not necessarily the ones with the highest population density, though. For example, Russia has the 9th largest population (142 million) but one of the lowest population densities, whereas Monaco is one of the most densely populated nations while being among the least populous (about 31,000 people). Standards of living and resource use per person are also uneven; note, though, that there is no

15. All numerical data in this paragraph are from: United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

consistent relationship between the number of people and affluence in an area. Returning to the nations we just compared: Russia, the much more populous of the two, is ranked 72nd in the world in per-capita wealth (as measured by GDP[16] / person, a reasonable reflection of standard of living), whereas tiny Monaco is the second wealthiest per person (the U.S. and China are ranked 17 and 102, respectively, in terms of GDP / person).

8.4.2. Levels of development and wealth often predict population growth

Earth's peoples and nations can be categorized and grouped using several different criteria. One common, if somewhat crude, system is to divide the world into groups based on per capita income and **level of industrialization** or **development**. This strategy is often helpful because it can be used to understand important factors influencing past growth as well as help predict what will occur in the future.

Nations with the highest level of development

These developed countries have relatively high per-capita income, levels of industrialization, infrastructure, life expectancy, education, and overall standard of living (i.e., enjoy relatively high GDP / person). Generally, they tend to be democratic and control a large amount of the world's wealth. Countries such as the United States,

16. Gross Domestic Product. For information about GDP and what it means, consult https://www.bea.gov/data/gdp/gross-domestic-product

Canada, Japan, Australia, Israel, South Africa, most of Europe, and a few other places in Asia, South, and Central America are on this list[17, 18]. Although the developed world is not uniform in this regard, growth rate here is relatively slow, at or below the global average. In fact, in some cases, fertility rates are appreciably lower than replacement; Japan (1.38 child / woman) and Germany (1.57)[19] provide two examples of developed nations with shrinking populations. Moreover, as a group these nations make only small contributions to global population growth—between **10 and 20% of human births occur in the developed world**—and its relative contribution is expected to shrink during the coming decades[20].

Nations with lower levels of development

The list of developing nations is longer than that for the developed world and includes many nations in Africa, portions of South and Central America, some Middle Eastern and Gulf states, China, India, and other parts of Asia[21,22]. These countries are typically grouped together here because, relative to the developed world, they have lower per capita incomes, levels of industrialization, infrastructure, life expectancy, education, and standards of living. They also hold

17. United Nations. 2022. World Economic Situation and Prospects
18. United Nations Conference on Trade and Development. 2022
19. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/
20. United Nations Conference on Trade and Development. 2022
21. United Nations. 2022. World Economic Situation and Prospects
22. United Nations Conference on Trade and Development. 2022

less wealth and tend to be less democratic. This second category is itself rather heterogeneous, though, including a wide range from countries such as Somalia and Niger, arguably on the low end of development, to a place like China, a rapidly changing, advancing nation that, depending on the method used to make the measurement, has either the highest or second-highest GDP on Earth[23]. Fertility rates also vary in the developing world, but they are typically at or higher than the global rates, in some cases by a great deal. For example, Niger and Somalia are two of the fastest growing nations on Earth (fertility rates of 6.8 and 5.2, respectively[24]) while being two of the poorest. On the other hand, Bangladesh, one of the least developed nations, defies the overall trend and has a fertility rate of 2.1[25] (as we will see in more detail shortly, China's is even lower). Despite the less-than-perfect relationship between development and growth, it is fair to say that nations on this list are home to the bulk of global population growth: on the order of **80 to 90% of births occur here,** and the U.N. population division expects the proportion to increase in the future[26].

8.4.3. Level of resource use per person increases

23. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

24. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

25. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

26. United Nations Conference on Trade and Development. 2022

with level of development

It is likely unsurprising to hear that, with increasing per-capita incomes, expenditures on energy, technology, and food go up as well. Put another way, **resource use per person** is generally higher in the developed world than it is in the developing world. Differences in food consumption provide a good example of the way affluence affects access to resources. For example, according to the United Nations, an average person in Europe and North America consumes around 3,600 calories per day whereas the number for Africa is 2,600[27](the amount needed for survival varies but is generally between 1,800 and 1,900 per day[28]). Those regional numbers do not reflect the range, though: people in the United States have access to about 3,770 calories per day (nearly 50% more than is needed for survival) whereas availability in the Central African Republic averages about 1,790 calories per day[29]. Non-food energy usage is similarly unequal. Consider petroleum: although the United States is home to just 5% of the world's population, it accounts for 20% of the use of this important energy source[30] (we will see much more about energy supply and demand in Chapter 10).

27. United Nations. 2022. UN News. Health. Once again, US and Europe way ahead on daily calorie intake

28. Minimum daily requirements for calories, 2020. ourworldindata.org

29. Max Roser, Hannah Ritchie and Pablo Rosado. 2022. Food supply. ourworldindata.org

30. United States Energy Information Administration (EIA). 2022. www.eia.gov

8.4.4. Nations develop and change

The relative contribution a nation makes to global population growth, as well as its share of resource use, can vary substantially with time. After all, the term *developing* implies a dynamic situation: a country may undergo dramatic changes to its infrastructure and standard of living—along with its cultural values and other factors affecting birth rate—such that it might look very different today than it did mere decades previously. As suggested earlier, China provides one particularly vivid example of such rapid evolution. In the 1960s it had a fertility rate higher than 6 and was one of the poorer nations on Earth. A mere 30 years later, its population growth rate had dropped below replacement and its control of the world's wealth began to increase dramatically. See Box 8.4 for more information about fertility, demographics, and development in the country that has the largest population on Earth.

Box 8.4. China: slowing population growth, increasing wealth

As is well known among many people, this population grew rapidly during the second half of the Twentieth Century to become the largest on Earth (just over 1.4 billion people in 2022[31]). Fertility rates were between 5 and 6 each year from the 1950s until the early 1970s,

31. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

between 3 and 4 in the 1970s, between 2 and 3 in the 1980s, and finally dropped below 2 in the early 1990s[32] (in 2022 the number was about 1.45[33]). Although many forces were responsible for this rapid transformation, the so-called "One Child Policy" enacted by the Chinese central government in 1979 had the biggest impact. In short, couples were forbidden by law from producing more than one child. Many exceptions were ultimately allowed, and the law was not enforced uniformly. However, it certainly did bring about the goal, a profound decline in birth rates (the policy was more or less eliminated in 2016). Note that several unexpected and undesirable consequences resulted from the law, including gender inequities (many observers report that some couples wishing to ensure that their one child was male abandoned or killed female infants, actions most will likely view as profoundly objectionable on their own, but ones that also created a much larger pool of young men looking for wives than there were wives to go around in subsequent decades). The population also aged a great deal, meaning that the proportion of older people has grown relative to that of younger

32. Fertility rate, China. The World Bank. 2015. data.worldbank.org/indicator/SP.DYN.TFRT.IN?locations=CN

33. [34]United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

34.

people—among other concerns associated with this change is the fact that an ever-dwindling workforce (as a fraction of the total population) must support an ever-expanding number of elderly people. At the same time this substantial decline in population growth rate occurred, China's economy expanded rapidly, in recent years, far faster than most other nations, and the standards of living of most of its residents improved. The current government's plans include further expansion and encouragement of domestic consumption[35]. So why is China generally still categorized as a developing country? The answer is largely a function of its low per-capita income, ranked 102nd in the world[36], lower than both many nearby and distant countries. In any case, we should expect that, assuming it follows the usual trajectory, continued development will inevitably increase China's resource use and environmental impact.

Acknowledging that not all cultures and peoples have the same goals, values, and aspirations, it is still broadly fair to say that societies tend to progress through similar stages of advancement, albeit at different times, and become more and more technologically advanced, urbanized, and affluent. So, like China,

35. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

36. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

many nations that have historically experienced low levels of resource use and industrialization are currently in a state of transformation, moving them toward those seen in the developed world. What are the consequences of worldwide increases in standards of living and expectations? Put simply: human demands on resources exceed the planet's capacity to supply them. We can visualize the consequences of ever-increasing development among Earth's nations by asking a simple question: how much space would be required if everyone enjoyed the same standard of living as those in developed countries (say, the United States)? Taking into account resources necessary to produce the food, building materials, and energy demanded per person—the so-called **ecological footprint** of an average U.S. citizen—the answer is: the equivalent of just over five planet Earths[37] (see Box 8.5 for a little more about resource use in affluent countries). On the other hand, the same calculation suggests that if everybody lived like citizens of Bangladesh, Uganda, or India, *less* than the equivalent of one Earth would be needed to meet demand for resources. As with any measurement, you should realize that there is a fair amount of uncertainty associated with this one, as well as some disagreement among scientists as to the validity of the assumptions upon which it is based. However, it highlights an important conundrum: people on Earth seem to aspire to standards of living that are simply unsustainable. Creative decision making and solutions, as well as changes to our expectations about how we should live, will be needed to avoid shortages and conflicts in the future (more about this challenge is presented in an upcoming section).

37. Global Footprint Network. 2023. National Footprint and Biocapacity Accounts. www.footprintnetwork.org

Box 8.5. Why the United States?

You might be asking yourself why the United States is singled out in this discussion of consumption and ecological footprint. The short answer is: it is the third most populous nation on Earth and has one of the highest rates of resource use. Additionally, it would be very difficult to argue that it is *not* an influential country—many of its values and habits have been embraced and mimicked for decades. In other words, it often sets the pace for what other nations do. So, it is a very relevant and at the same time instructive example. All that said, you should be aware that consumption rates in some other nations are also quite high. In terms of per-capita ecological footprint, The United States is actually only fifth highest on the list behind countries such as Luxembourg and Qatar (8.2 and 9.0, Earth's required, respectively—see the main text for more about the calculation).

8.5. WHAT DOES THE FUTURE HOLD?

What lies ahead? Will the human population continue to grow exponentially or will its rate of increase slow? Might the total number of people even decline? How will ever-increasing demand for limited resources affect numbers of people and quality of life? Here we will briefly address these important and challenging questions and use what we have learned to predict the future of humans on Earth.

8.5.1. Will our population just keep getting bigger and bigger?

Near the beginning of this chapter we learned how current human population growth is exponential (review Figure 8.1, above). You should realize, however, that this trend cannot continue forever. At some future time, the rate must inevitably slow (perhaps even fall below 0) because the supply of resources will not be able to keep up with the demands for them. If you recall that Earth is closed with respect to materials, you will see how no population can continue to increase in size indefinitely. Fixed amounts of space, nutrients, and water simply limit the absolute number of organisms that can be supported at any one time. Returning to a concept we first encountered in Chapter 1, we would say that a population will continue to expand until it reaches its **carrying capacity**. Importantly, exponential growth can occur only under ideal conditions, principally, so long as necessary resources are available in excess. Plotted on an x-y graph, we see a J-shaped curve (Figure 8.2a). Once a population approaches its carrying capacity, though, growth rate slows, yielding an S-shaped, or **logistic**, curve (Figure 8.2b).

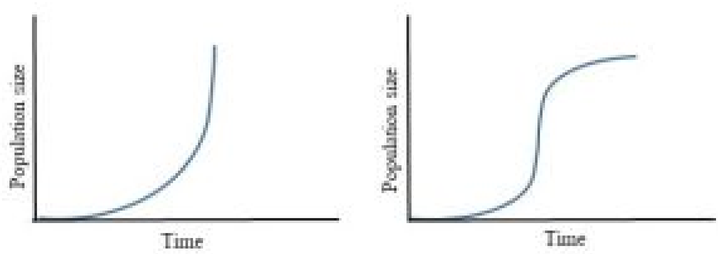


Figure 8.2. *Exponential growth curve, a; logistic growth curve, b. Note that the middle section of b is exponential. Kelsey,* CC BY-NC-SA.

*An **exponential growth curve** represents growth in the absence of limitations, an unrealistic situation in the long run. A **logistic growth** curve describes the likely pattern under real-world conditions of finite resources.*

We know that the human population is ultimately bound by the same rules and limitations as any other on Earth: it cannot increase exponentially forever. Just how big the human population will get and when growth will reach the flat phase of the logistic curve (review Figure 8.2b), though, is unclear and the subject of a fair amount of debate and disagreement. Various researchers have offered estimates of the maximum number of people that can be sustained by this planet. On the low end, the number is probably about 2 billion people, assuming everyone lives like a citizen of a developed country such as the United States and perhaps as high as 40 billion if all were to meet only basic needs for survival (recall the discussion above about ecological footprint and the number of Earth equivalents required under various scenarios). Assuming everybody will aspire to higher standards of living than mere subsistence, the number is almost certainly appreciably lower than our current 8 billion. More about human carrying capacity is presented in Box 8.6.

Box 8.6. Will we exceed carrying capacity? Are we doomed?

What will happen if population keeps expanding at a high rate? Might we reach a point at which disaster will strike? Many have weighed in on these questions, but perhaps the most famous prediction was offered over 200 years ago by Thomas Malthus, an economist and English cleric. In short, Malthus concluded that humans' drive to reproduce will ultimately overwhelm Earth's resources and lead to shortages, war, and mass death and destruction[38]. He has his critics. For example, many demographers and ecologists contend his premise is flawed because humans would never allow their population to grow to intolerably high levels. In other words, growth rate will slow due to negative feedback (Chapter 2). Others hold that our continued growth and existence is evidence his prophecy was false; in fact, there are even a few who think technology and innovation will allow us to grow indefinitely to many tens of billions of people. On the other hand, he has his followers, those who embrace his overall vision. They contend doom is inevitable, but we just have not yet reached the crisis point he imagined.

8.5.2. Can we project what will happen?

According to the United Nations Population Division, growth rate is likely to enter the flat phase of the logistic curve during the coming

38. Malthus T.R. 1798. An Essay on the Principle of Population. Chapter 1, p. 13 in Oxford World's Classics reprint. © 1998, Electronic Scholarly Publishing Project

decades, resulting in a human population of about 10.4 billion people by the year 2100[39] (review Table 8.1). Keep in mind that this projection is based on many difficult-to-predict variables and, therefore, is only a well-informed estimate. For example, the slowing of growth depends on expectations we noted earlier, namely that education, empowerment of women, urbanization, and industrialization will all increase and that fertility rates will decrease. Figure 8.3 illustrates possible patterns of growth that could occur between now and 2100 and the levels of uncertainty associated with this prediction. Note that most reflect slowing growth, however little change from our current trajectory is also possible.

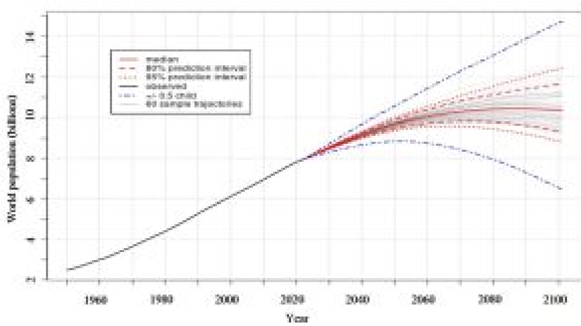


Figure 8.3. *U.N. projections of world population until 2100. United Nations Population Division,*

8.6. CAN WE MINIMIZE THE STRESS FROM HUMAN POPULATION?

Before we end this chapter, we should consider a final question:

39. United Nations Department of Economic and Social Affairs, Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

what can we do to reduce the effects of human population on Earth's environments and organisms? The answer is quite complicated, but we will briefly suggest a few potential solutions and the challenges associated with them.

8.6.1. Make fewer people?

On the surface, this strategy appears straight forward and reasonable: if the human population were smaller, its impact would be, likewise, smaller. We might be tempted to encourage (or even coerce, depending upon how authoritarian we care to be—review Box 8.4) lower birth rates to accomplish our goal. Furthermore, we could also conclude that since 90% of new growth occurs in the developing world, our reduction efforts should focus on those countries with highest fertility rates to relieve the stress brought on by humans. Unfortunately, simply producing fewer babies (no easy task under any circumstances, as we will explore below) would likely not bring about all the benefits we seek because numbers alone do not reliably indicate the adverse impact a country can have on global natural systems. Recalling that per-capita resource use in developing countries is generally far less than it is in developed countries (e.g., the average ecological footprint in, say, Bangladesh vs. that in the United States), we see that lower birth rates alone would not solve the problem. Importantly, a person born and raised in a typical family in the developed world will have greater opportunity to affect global systems than one in the developing world for at least three reasons.

1. Higher level of resource use

We know resource use per person is far from equal on Earth. So, although the fertility rate is over two times higher in a place like the

Democratic Republic of the Congo than it is in the United States, babies born in each place will likely lead very different lives and use vastly different amounts of food, water, building materials, fuel, and space.

2. Greater access to technology

Industrialized nations use more fuel for transportation, computers, cell phones, and the like than do agrarian nations. As we will see in greater detail in Chapters 10 and 14, the production of usable power almost always releases air pollution into Earth's atmosphere, and those products of combustion move great distances from their sources via natural processes. So, a person driving a car in the United States has a greater capacity to remotely and directly affect people on other continents than does a person lacking access to a motor vehicle in a place such as Bangladesh or Haiti.

3, Longer life expectancy

Developed countries generally enjoy lower death rates (review the growth rate equation)—that is, longer lives—than do developing countries. Average citizens in the United States (life expectancy as of 2020 was 77 years), therefore, have even greater opportunities to affect natural systems because they use more food and fuel and so forth for a longer time than do people in the Democratic Republic of the Congo (60 years)[40]. Incidentally, these two countries do not represent the extremes on Earth, which are Monaco (89.5 years) and

40. The World Bank. 2022. Life Expectancy at Birth, Total (years). data.worldbank.org

Chad (53 years)[41]. You will likely find it unsurprising that Monaco has the 2nd highest GDP / person whereas Chad is ranked 220th on a list of 229 reported by the United States Central Intelligence Agency. A few final comments are necessary here. First, average life expectancy for all nations is 72 years[42]. Second, life expectancy is higher for females than males (73 and 68 years, respectively[43]). Third and finally, we saw that life expectancy has generally increased in recently decades, although the COVID-19 pandemic led to a decline between 2019 and 2021[44].

Before moving forward, we should explore a potential downside to reducing the inputs of new humans into Earth's system. Assuming death rates continue to decline as standards of living and medical knowledge increase, people will be living longer and longer lives in the future. At the same time, a reduction in birth rates would lead to fewer and fewer young people as a proportion of a still-growing population. This **demographic shift** is already occurring in many countries around the world in any case (e.g, China, as described in Box 8.4), but a focused effort to bring down birth rates even further would make the change that much more pronounced. Why should we care if elderly people make up more and more of the human population? The answer is best illustrated

41. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

42. The World Bank. 2022. Life Expectancy at Birth, Total (years). data.worldbank.org

43. United States Central Intelligence Agency (CIA). The World Factbook. 2022. www.cia.gov/the-world-factbook/

44. United Nations Department of Economic and Social Affairs, Population Division. 2022. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/No.3

by a look at governmental programs such as Social Security in the United States. Putting it into simple terms, this system was based on the premise that people of working age would contribute part of their wages to a fund used to pay money to retired people. It works well until and unless the number of people paying in dwindles relative to the number of people receiving benefits. As has been reported extensively by various media outlets during recent decades, unless changes are made, the money in the Social Security system will be insufficient to meet demand within 20 or so years (think systems analysis again: inputs cannot keep up with outputs). Reducing the numbers of new babies would certainly bring about some benefits in terms of demand for resources, but it could also change our ability to care for an ever-aging population.

8.6.2. Reduce demand by lowering expectations?

A second approach to minimizing human-caused stress on Earth's systems is to reduce the use of food, energy, and other materials in countries with the highest per-capita resource use. For example, we noted earlier that an average resident of the United States consumes far more calories than are needed. The amount of edible food that goes *uneaten* in that country—wasted on farms, by markets, or by consumers—is also very high, estimated to be 30 – 40 % of all food produced[45]. Similarly, daily water usage per person far exceeds needs in many developed countries (e.g., 590 liters per day in the United States), but is near or below the minimum amount

45. United States Department of Agriculture (USDA). 2022. usda.gov

needed[46] in places such as India (143 L / day) and Mali (11 L / day)[47]. Recall from earlier discussions in this chapter that energy usage follows the same trends. Putting these and other considerations together, we can draw a fundamental conclusion: many more resources are expended by the most affluent nations than are required for survival. Modifications in usage patterns in these places as well as changing expectations worldwide could mitigate impending shortages.

8.6.3. Confront both size and consumption patterns?

It is likely evident by now that alleviating the stress caused by the human population requires a multi-faceted approach; no one strategy alone would be sufficient. Certainly, the increased demand created by more people is a concern, but so too is the high standard of living all the people on Earth expect and achieve. The nature of the problem can be summed up in one more equation:

total effect of the human population =
(effect of each individual) X (total number of individuals).

Notice that an increase in either the effect per individual or the number of individuals will increase the total effect. Given the discussion and data presented throughout this chapter, we can reasonably conclude that *both* of those terms are increasing now and appear to be trending upward into the foreseeable future. The

46. Defined as 50 – 100 L / day by the United Nations

47. United States Centers for Disease Control and Prevention (CDC). 2020. Water Use Around the World. cdc.gov

overall effect of the human population, therefore, is likely to go up at a high rate.

8.6.4. Can we do this? Should we?

Addressing the stresses introduced by the human population is difficult, to say the least—potential solutions are complex, nuanced and, in many cases, unpalatable to people due to the cultural, religious, and social values they hold. Of course, we should also remember what a tough sell this is: reducing consumption when you are accustomed to plenty is simply no fun! Moreover, if we recall the trend in the developing world toward standards established by the United States and other developed countries (Box 8.5), a transformation that many view as a matter of justice and rights, the problem appears even more intractable. The data and evidence presented in this chapter, though, indicate that the time for careful and intensive consideration of the best ways to avoid shortages, conflict, and suffering is upon us.

THE CHAPTER ESSENCE IN BRIEF[48]

The size and behavior of the human population are two fundamental

48. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 8 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

concerns of environmental scientists because they influence resource availability and environmental quality on a planet closed with respect to materials. Chapter 8 begins our study of the many anthropogenic sources of stress affecting natural systems as well as the health and survival of Earth's peoples.

Think about it some more...[49]

How and why has the growth rate of the human population changed during the past century?

How can feedback lead to exponential growth?

What are the effects of increased life expectancy on the growth rate of a population? What other effects could it have?

Which would have the greater adverse effect of Earth's global ecosystem and resource availability: the addition of a baby to the developed world or the addition of that same

49. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

baby to the developing world? You need to consider many variables as you ponder your answer to this question!

9. Agriculture, Soil, and Food

JASON KELSEY

Now we turn our attention to the ways we feed the people we just saw in Chapter 8. The concerns of Chapter 9 are intertwined with many other topics in this book as well, including ecology (Chapter 5), water quality and availability (Chapters 4 and 11), energy use (Chapter 10), air quality (Chapter 14), and environmental toxicology (Chapter 15). Put another way, an understanding of agriculture is critical to the study of environmental science. It also should be of interest to anyone who has ever eaten a meal.

Key concepts

After reading Chapter 9, you should understand the following:

- How and why agriculture is employed to produce human food
- How farms are different than natural ecological systems
- The processes that contribute to soil formation and why soil properties matter to farmers
- How maximum sustainable yield is like carrying capacity (Chapter 1)
- The multiple environmental effects of agriculture

- The potential risks and benefits of the genetic manipulation of agricultural organisms (i.e., the use of GMOs)
- The many different approaches to farming used by modern humans to produce food
- How the issues of human population growth (Chapter 8) are intimately linked to those of agriculture and food

Chapter contents

9.1. AGRICULTURE IS INTEGRAL TO MODERN HUMAN SOCIETIES

Agriculture currently dominates food production for humans, but we have not always been farmers. Much evidence suggests that large-scale cultivation of crops began about 12,000 years ago, meaning earlier peoples fed themselves through hunting and gathering. The development of agriculture had a profound effect on human civilization because, among other things, it allowed people to settle in one area rather than moving from place to place, in a nomadic fashion, searching for food. In addition, it broadly increased food availability. The large-scale farming of the past hundred years or so also has brought on adverse consequences. Do the benefits of agriculture justify its costs? What would we do without it? Might we come up with a better system? You can wrestle with those and other important questions as you proceed through this chapter.

9.2. FARMERS HARNESS ECOSYSTEMS FUNCTIONS

An incomplete, but useful, **definition of agriculture** is: the manipulation and augmentation of natural ecological processes by humans to concentrate and maximize the growth of food in areas that would otherwise be less productive. Employing terminology common in agriculture, farmers use available tools and resources to optimize **crop yields**, the amount of food produced per acre of land. Additionally, large numbers of animals are concentrated and raised so they (or their products) can be consumed by humans. Unlike the hunter-gatherers mentioned above, who had to search—sometimes by traveling great distances—to secure enough to eat, in a sense we have learned to make our food come to us.

9.2.1. How food is produced on farms: a brief overview

Although specific details can differ substantially from farm to farm, an essential set of steps is usually followed when crops and animals are raised by humans. A brief overview of these processes is provided here, but we will explore more details about agriculture and food production later.

Crops

Existing ecological communities such as forests are cleared to make room for crops. The upper portion of soil is typically loosened by plowing or a similar process, and then seeds are planted. If deficiencies are noted, nutrients and water are applied to agricultural fields. Additionally, chemical pesticides may be added or other steps taken to inhibit unwanted organisms that may interfere with productivity. When they are mature, crops are harvested.

Livestock animals

Animals are bred on a farm or at another facility specifically designed for that purpose. Food grown on the farm or elsewhere is given to animals to fatten them up, and certain growth hormones and antibiotics may be added to their diets. Animal waste products must be managed to prevent soil and water pollution, sometimes by using them as fertilizer for crop growth. In the cases where products such as milk and eggs are sought, they are collected and processed. Ultimately, nearly all animals are slaughtered when they reach a target age.

9.2.2. Farms are agroecosystems

Environmental scientists often use the term **agroecosystem** because, although they do take advantage of natural processes, farms differ from naturally occurring, undisturbed, ecosystems in several important ways.

Stunted ecological succession

Crops like corn, wheat, and hay resemble **early stage successional species** more than they do late ones (Chapter 5), so farmers tend to clear land and maintain it in that open condition indefinitely. Since agriculture is most common (and successful) in temperate terrestrial areas, allowing succession to proceed unimpeded would ultimately lead to a climax community dominated by tall, shade-tolerant trees, an environment that discourages crop growth. Unfortunately, artificially preventing succession in this **biome** also provides an inviting environment for unwanted plants—known, of course, as **weeds**—because they too thrive in sunny, barren areas (weeds are typically fast growing and grassy). In short, by creating conditions favorable to the plants we want, we encourage the growth of costly competitors as well.

Monoculture

To maximize efficiency, most farmers employ the practice of **monoculture**: just one crop is sown on any parcel of land. A farm is likely to use several different fields at one time to grow, say, corn, hay, and soybeans, but the species are kept separate from each other. On the other hand, as we saw in Chapter 5, natural ecosystems consist of a community of different organisms (Figure 9.1 compares a corn field to a natural forest).

Figure 9.1. *An agricultural field in which only corn is cultivated, left (a); a natural forest ecosystem containing many different species of plants, right (b). Rob Purvis, CC BY (a); Stefan Wernli, CC BY (b).*

In addition, in an agroecosystem, plants tend to be packed together far more densely than they would be under natural conditions. Two unintended and unwelcome consequences result from the practice of monoculture. First, insects adapted to eat the cultivated crop are provided with a substantial advantage. Since their food source is highly concentrated in a field, nutritional limitations are no longer a problem. They can get on with the energy-intensive and fundamental business of reproduction (Chapter 5). For example, grasshoppers can live on several types of plants, including corn. Imagine yourself to be such an insect and what it must be like to happen upon a 500-acre corn field after struggling your whole life to find food in diverse ecosystems characterized by scarce resources. One of your major worries would be over! Second, a high concentration of one plant species in an area quickly depletes the soil in which it grows of necessary nutrients.

Plowing

This and related practices facilitate seed germination by physically breaking up hard, compacted soil prior to planting. Plowing also enhances water movement through soil and can damage the roots of weeds. Despite the benefits, though, mechanical disruption of soil does have quite a downside: among the adverse consequences of plowing is the way it greatly accelerates the loss of fertile soil from farmlands. Topsoil erosion is a major concern associated with agriculture, and we will devote substantial time and energy to it later in this chapter.

Manipulation of nutrients and water

Recalling that the purpose of agriculture is to produce as much food as possible in a given space, it ought to stand to reason that farmers do everything they can to stimulate growth of crops. For example, nutrients exhausted through excess demand by crops and lost when topsoil is eroded can be artificially replenished through the application of **fertilizers,** mixtures containing biologically available forms of nitrogen, phosphorus, and other elements (Chapter 4), to fields. Natural ecosystems have no such means to quickly replace what is missing. Fertilizer use can provide a lot of benefits, but we will see how it can also lead to pollution of water (later in this chapter and Chapter 11). Water supply on farms can also be supplemented through the addition of water to fields in ways that are not applicable to natural systems. This practice, known as **irrigation**, can be the difference between success and failure of crops, but it requires the costly transport of water to farms from other regions. The environmental, political, and economic consequences of irrigation can be substantial, as described later in this chapter and in Chapter 11.

Inhibition of unwanted organisms

It is important to realize that farmers do not want organisms other than crops (plus any organisms that might enhance crop growth) to succeed in their fields. They can use one or more chemical, biological, and physical strategies to discourage pests that interfere with crop production (more in an upcoming section of this chapter). There are some analogous processes in natural ecosystems, as certain organisms in natural ecosystems possess chemical and other weapons to fight competitors (Chapter 5), but the scale on which they are deployed is much larger on farms.

Minimal recycling of nutrients

On many farms, crops are completely removed at the end of a growing season. In other words, in addition to the target product (ears of corn, the soybeans themselves, etc.), the entire plant is cut down and taken away rather than being left in the field. This practice clears a field for future planting, but it does not take advantage of natural decomposition and the way it helps return nutrients to the soil for next-year's crops. Natural ecosystems are very different because cycling and recycling of nutrients is integral to their functioning and continuation (Chapters 4 and 5). It is worth noting here and will be reiterated below that some farmers do in fact leave plant residues behind in fields after harvest, but that practice is not widespread.

9.3. FOOD PRODUCTION RELIES ON KNOWLEDGE OF SOIL SCIENCE

It tends to be taken for granted by most people, but *soil* is a vital resource that merits special care, preservation, and study. We are interested in it as environmental scientists for at least four reasons. First, with very few exceptions, crops are grown in soil. Our ability to effectively farm is directly related to soil quality: agriculture is successful in soil that is characterized by a certain set of necessary features and properties—put simply, if it is fertile—and fails in the wrong kind of soil. Importantly, though, the very practice of agriculture tends to reduce soil quality, that is, the more it is used for farming, the less able to support crop growth it becomes. Second, the ecology of soil affects all of Earth's systems. Even the tiniest soil organisms are indispensable to the cycling and ultimate availability of nutrients to the rest of the biosphere. Third, pollutants released into natural environments are often modified by processes occurring in soils (we will return to this topic in Chapter 15). Finally, the building of houses, bridges, and other structures is influenced by the chemical and physical properties of soils. Put into stark terms, whether a new bridge stands or falls after it is exposed to its first intense rainstorm is in part a function of the way the soil beneath it behaves. Since knowledge of soil science informs many activities and processes valued by humans, including food production, we will take some time here to study its basic principles.

9.3.1. The definition of soil

Environmental scientists define **soil** as a naturally occurring material that meets the following criteria.

It is unconsolidated

Unlike rocks, the component parts of soil are not cemented together into a solid object. They are loosely packed and can be easily moved, changed, or transported by wind, water, and ice.

It comes from other materials

We will study its formation in more detail shortly, but for now recognize that soil is produced through biological, chemical, and physical modifications of pre-existing organisms and rocks.

It contains solids and spaces

Soil is referred to as a porous medium because it is a matrix of solid materials and pores.

It is layered

As we will see below, to be classed as soil, unconsolidated material must be arranged in characteristic layers which differ from each other in important ways.

It can provide nutrients, habitat, and water for organisms

Soil supports the growth of a diverse array of plants, animals, and microorganisms.

9.3.2. How soil is formed

Soil formation begins when existing organisms and rocks, known as **parent materials**, are first exposed to forces such as **weathering** and **decomposition**. The resulting products undergo additional modifications, and recognizable soil is eventually created. Note that the development of soil often requires decades or centuries and that even after it comes into being, soil continues to be modified by the same forces that initiated its production. The chemical and physical properties of any soil will therefore change as the soil ages. Figure 9.2 provides a simple diagram of soil formation.

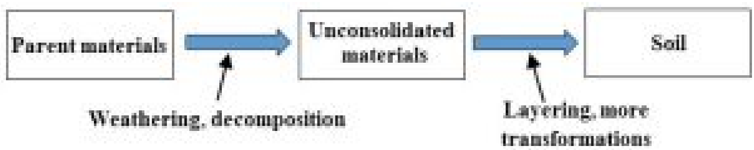


Figure 9.2. *An overview of the soil formation process. Kelsey,* CC BY-NC-SA.

Weathering and decomposition of parent materials

Rocks. Rocks are consolidated objects largely comprised of **inorganic chemical compounds**. They can be broken down into small fragments by both physical disintegration and through a number of chemical reactions (Chapter 3). These products, known as **unconsolidated sediments**, are *not* soil by our definition because they lack the layering referred to in Figure 9.2, above. Given the appropriate conditions, additional modifications yield true soil.

Remains and products of organisms. We have seen elsewhere (Chapters 4 and 5) that decomposers are integral to nutrient cycling because they break down the dead tissues and waste products left by organisms. They are also critical in soil formation because they process and convert these remains into an important material known as **soil organic matter (SOM)**. The chemical and physical properties of SOM are extremely complex and not even completely understood by soil scientists (this is an area of current research). We define SOM by several characteristics.

It originated from organisms

SOM is primarily made up of chemical compounds produced by living things.

It contains materials from multiple organisms

The remains and products of microorganisms, plants, and animals are mixed together and collectively become SOM. Many forms of carbon, nitrogen, phosphorus, and other nutrients are present.

It does not resemble organisms in appearance

Although SOM does come from living members of the biosphere, it does not contain recognizable structures such as leaves or body parts. It is instead comprised of chemical compounds that have been dramatically altered. Some of their inherent identity is intact, but many old chemical bonds are broken and new ones formed as the materials are processed by decomposers. It could be visualized a bit like soup you might start from vegetables, meat, and spices on Monday but then forget about and leave cooking until Sunday. When you get back to it days later, it is likely to be a uniform mush: the distinct ingredients will have been blended together, making it difficult or impossible to recognize them in their original forms. The whole thing is now a new entity with large-scale properties, not a collection of individual pieces of carrots, chicken, and the like floating in water. So, the remains and excrement of fungi, grass, rabbits, bacteria, foxes, and all the rest that live in an area get transformed and combined to form SOM (Figure 9.3).

Figure 9.3. *An identifiable dead rabbit, left (a); SOM contains the altered remains of many organisms mixed together, right (b). Brian Robert Marshall,* CC BY-SA *(a); normanack,* CC BY *(b).*

Development of layers

The relatively uniform accumulation of weathering products from the first step in soil formation can be slowly arranged into layers due to several processes, two of which are briefly described here (see Figure 9.4, an idealized overview of the process).

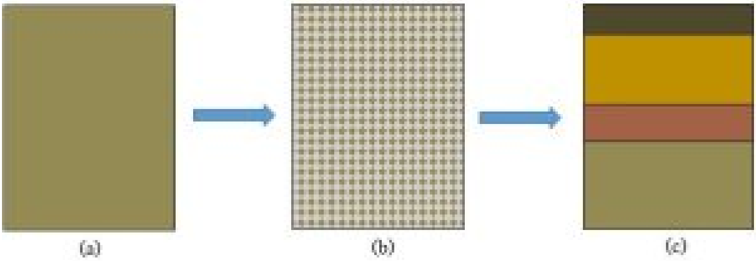


Figure 9.4. *A simplified diagram of the steps in soil formation. Solid rock (a), can be weathered into relatively uniform unconsolidated sediments (b); additional transformations with time can produce distinct soil layers (c). Kelsey,* CC BY-NC-SA.

Removal and re-deposition. Some materials in the upper portion of unconsolidated sediments are easily removed by flowing water and carried vertically downward (a process known as **leaching**—see Chapter 11). Many of these materials leave the water after they have travelled to some depth and end up redeposited into the matrix of sediments. As a result, the properties of both the upper and lower layers are changed.

Incorporation of SOM. Remains and products of organisms accumulating on top of the soil begin to undergo decomposition and transformation to SOM as described above. As time goes by, more and more of this organic material is incorporated into the soil layers near the Earth's surface, combining with the inorganic weathering products that originated in rocks. Microscopic examination of soil reveals patches of SOM adhering to larger surfaces of the inorganic material in soil (Figure. 9.5).

The soil profile and soil horizons

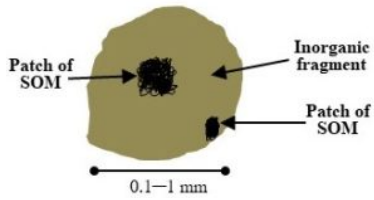


Figure 9.5. *Diagrammatic microscopic view of patches of SOM on surfaces of inorganic fragments. Kelsey,* CC BY-NC-SA.

The **soil profile** is a cross-sectional view of the true soil produced by the differential gains and losses of materials in upper and lower regions, the additional of SOM, and some other chemical changes. It contains several layers, known as **soil horizons**, that are each characterized by distinct chemical, physical, and biological properties. Generally, top horizons contain SOM and are the zone in which plants, including crops, and other organisms can grow. Bacteria are the only members of the biosphere found in the lower layers. Specific properties, including thickness of individual horizons, can vary widely from location to location. Sometimes, certain horizons are completely absent. The major horizons and their properties are summarized in Table 9.1, and Figure 9.6 is a photo of an actual soil profile (note the photo is not as ideal as is suggested by Table 9.1).

Table 9.1. Major soil horizons and their important properties.

Horizon	Important properties	Typical depth (cm)
O	Contains both identifiable and partially decomposed remains of organisms	Surface to a few cm
A	Consists of topsoil; high concentration of SOM; brown / black in color	5 – 30
E	Inorganic materials; lacks SOM; light color	10 – 60
B	Known as subsoil; various colors, can be darker than E; accumulation of materials leached from A and E	30 – 80
C	Partially weathered parent rock	60 – 100
R	Solid bedrock	Below C

Figure 9.6. *A photo of an actual soil profile, showing some visible horizons.* USGS, *Public Domain.*

9.3.3. The composition of soil

Soil is made up of solids and pores, each of which plays important roles for crops and other organisms.

Solids are derived from parents

Mineral fraction. Also known as the **inorganic fraction**, this is the bulk of a soil. It is produced by weathering of rocks, and its composition varies with the chemistry of those parents. It provides the solid support matrix of soils and some inorganic nutrients.

Organic fraction. Soil organic matter generally accounts for 5% or less of the weight of a typical agricultural soil, although it can range from 0 to 90% among soils. Its chemical properties depend on the nature of its parent materials and can vary widely. In soil it provides several important services.

Nutrition

As we know, SOM was produced from organisms so contains a lot of biologically available nutrients.

Structure

SOM provides support to a soil by acting as a kind of glue—it helps hold the mineral fragments together and allows for the formation of large clumps, referred to as **aggregates**, of soil (sizes of 10 or more cm in diameter are possible—See Figure 9.7).

Figure 9.7. *Some soil aggregates.* USDA, *Public Domain.*

Water retention and entry

The structure provided by SOM helps create openings through which water can move into and be held by soil. As we will see, the amount of water present in soil is a critical property affecting crop growth.

Retention of toxins

The chemical properties of SOM make it an ideal place for certain chemical pollutants to bind. See Chapter 15 for more.

Pores are spaces among the solids

Open spaces within the soil matrix can hold water and air. These two vital substances vary in relative abundance as a function of specific properties of a soil as well as climate and weather.

Water. Some fraction of the pore space within a soil will be occupied by water. The amount found here is influenced by the amount of SOM present as well as the sizes of the particles in a soil (described in detail in the section below).

Gases such as air. In most surface soils, atmospheric air is the gas stored in pores. It can provide much-needed oxygen, carbon dioxide, and dinitrogen for various organisms. The amount of water and air in soil are related to each other: as one increases, the other necessarily decreases.

9.3.4. Important soil properties

The chemical and physical properties of any soil are a function of a number of variables, including the chemistry of the parent materials, the climate in which the soil was formed, biological activity, and the age of the soil (more below). In other words, given that each came into being under different conditions, thousands of different soils exist. Several characteristics are used to identify and distinguish soils from one another. These properties are also important in determining the suitability of a particular soil for crop growth. Here we explore a small number of those most relevant for agriculture.

Texture

The solid portion of soil is comprised of particles that are grouped according to their size: clay (diameter below 0.002 mm), silt (0.002 – 0.05 mm) and sand (0.05 – 2.0 mm). The relative proportion of these size fractions, known as **soil texture**, can be determined through laboratory tests and is used to assign soils to different categories. According to the United States Department of Agriculture (USDA), there are twelve major soil classes: those at the three end points (ideally, 100% sand, soil, or clay) and nine intermediates[1]. For example, "loam", consisting of approximately 40% each of sand and silt and 20% clay, refers to a soil commonly used in agriculture. Texture is an important property because it influences the way water flows through soil. Too many very small particles (i.e., clay) will inhibit movement of water, likely flooding plants, whereas too many large particles (i.e., sand) will allow water to drain from soil too quickly for plants to access it.

Soil organic matter content

The percentage of the weight of a soil that is made up of SOM can be readily determined in a laboratory. As noted above, a very wide range of SOM contents is possible. Like most things we have encountered, both too much and too little of this vital constituent can inhibit plant growth and farm productivity. An overabundance of it will hold too much water in soil, and a deficit will deprive organisms of the services provided by SOM that are listed above.

Water-holding capacity

Soils differ in their ability to store water, and this important property is a way to quantify those distinctions. Three points related to this concept merit our brief attention.

Water remains after inputs cease. During a rainfall event water

1. US Department of Agriculture (USDA). Soil classification. https://www.nrcs.usda.gov/resources/guides-and-instructions/soil-classification

is freely available and moves vertically downward through the soil. Once the precipitation stops, water continues to drain for a brief period, leaving behind some amount of water stored inside pores. The amount of water remaining is essentially a soil's water holding capacity.

Water holding capacity affects plants. Water held in pores is very important because it can be used by plants to meet their needs for an extended period, potentially many days or weeks into the future. Under the right circumstances, plants will get what they need from this stored water until it rains again. Plants will eventually wilt and die, however, if dry conditions persist.

Water-holding capacity varies. The capacity to store water depends on soil texture and SOM content, two properties that differ among soils. As we saw, abundant clay or sand will retain either too much or too little water, respectively, for healthy crop growth. Organic matter in soil is also critical: as its content increases, so does water-holding capacity. In short, some soils are simply unsuitable for agriculture because they do not provide the proper amount of water to plants.

Chemistry

Several measurable chemical parameters influence crop growth. We will only consider two of the most important ones, but you should recognize that many others play roles in agriculture.

Nutrients. As we have seen previously, the amount and type of useable nutrients such as nitrogen and phosphorus present are extremely important. Farmers can use various tools to assess nutrient content and make up for deficits they might discover (for example, by using fertilizers).

Acidity. Another critical concern is the pH, or acidity, of a soil. Very briefly, the concentration of hydrogen **ions** (routinely expressed as H^+) present in water can be measured, and that number converted to a unit known as pH. The pH scale is generally understood to run from 0 – 14, with the lower numbers representing the most acidic substances and the higher numbers representing the most basic (or least acidic). A pH of 7 is neutral, neither acidic nor basic. Soils vary in their pH values, so their ability to support crop growth varies as well. Not all plants thrive at the same pH, and growth can be seen at acidic, neutral, and basic conditions. Farmers can determine acidity and then adjust it to some extent if necessary. We will encounter pH again in Chapter 14 when we study air pollution, including the way acidified precipitation can affect soil quality.

Color

This property is a function of soil composition. For example, dark brown or black soils contain more SOM than do lighter soils, red soils are rich in a certain type of iron, grey and yellow soils have different types of iron in them (Figures 9.8.a. and 9.8.b). Color alone cannot provide enough information to completely characterize a soil, but it is used to estimate the likelihood that a soil can be farmed effectively.

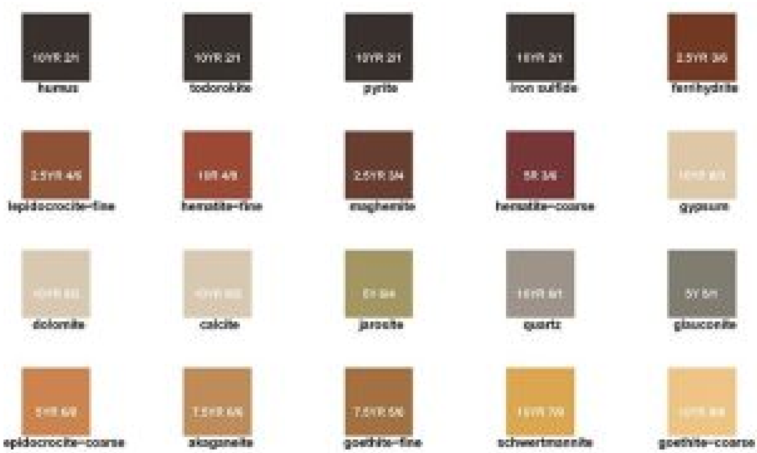


Figure 9.8.a. *Soil color can vary enormously, as suggested by this USDA color chart. The panels represent many possible colors and the main minerals responsible for them. USDA, Public Domain.*

Figure 9.8.b. *Photos of natural soils show just four of the many color possibilities. USDA Public Domain.*

Limiting factor

Of all the requirements for growth, sun, space in soil, water, carbon, nitrogen, phosphorus, and other chemical nutrients, the limiting factor is the **one item least available** relative to demand for it. Imagine a field full of stunted corn, plants that barely produced a few ears. The answer to the question "why did this crop fail to grow as anticipated?" is, by definition, the limiting factor. So, a farmer might investigate what happened and discover everything needed for optimum corn growth is available in sufficient amounts *except* for nitrogen. That nutrient may be added to the soil for the next season (or in the current season, if the problem is detected early enough) and yield should improve. The concept of limiting factor can be applied to a wide range of ecological systems, not just farms, to determine why organisms and communities are not as successful as might be expected. For example, despite abundant sunlight, CO_2, appropriate temperature, and other favorable conditions, the open ocean (i.e., far from a coastline) supports a relatively small number of organisms. What limits growth in this otherwise hospitable environment? Often, it is lack of phosphorus, nitrogen, or iron.

Soil fertility

Soil fertility is defined as the ability of a soil to provide necessary nutrients to support crop growth into the foreseeable future. Realistically, though, since we know that repeated farming tends to reduce soil quality, productivity is likely to deteriorate with time as well. Among the important reasons for this phenomenon is that nutrient availability declines as soils age. In fact, a crucial difference between young and old soils is diminished fertility in the latter relative to that of the former.

Maximum sustainable yield

This final property is really a function of the others on this list and is a critical consideration for farmers. **Maximum sustainable**

yield is the largest amount of a crop that can be produced in a given season without reducing the ability of that field to produce the same amount of that crop in subsequent seasons. Although it might be possible to produce an enormous quantity of food for one or two years, the soil will lose so much of its fertility as a result that crops will be less successful in the future. The goal is to strike a balance that allows for long-term yields to be as even and predictable as possible. This property should seem familiar to you as it is closely related to the concept of **carrying capacity** we encountered in Chapter 1. In this case, though, it is applied specifically to agroecosystems.

9.3.5. Soil properties depend on environmental conditions

It likely makes sense to you that the physical and chemical characteristics of parent materials affect the soil that results from their modification: the nature of the outputs from any process will be closely tied to the nature of the inputs that are available to build them. However, like all reactions, the conditions under which parents are weathered and otherwise transformed also play a critical role in shaping the resulting soil. For example, soils produced in hot, wet environments like the tropics are different than those from cool, dry, polar regions. How much oxygen is present will also influence the forms of iron and other elements in soil. Furthermore, the specific identity of the organisms involved in decomposition is not constant from place to place, so the products of their activities will be similarly different. In short, the properties of parent materials alone cannot be used to predict the nature of the soils produced from them.

9.4. ENVIRONMENTAL EFFECTS OF AGRICULTURE

Employing agriculture to produce food on a scale large enough to meet current demand comes with consequences. As we will see, these adverse effects are relevant at three scales: **local**, those that occur on the site of a given farm, **regional**, those that influence surrounding areas that are not farmed, and **global**, those that affect the entire planet.

9.4.1. Increased rates of soil erosion

Arguably, the rapid loss of fertile soil is the most important effect of agriculture because of the relationship between soil quality and food production noted earlier. It is generally observed on a local scale because it affects fields that are actively farmed; it is very much a problem around the world, though. Put another way, activities in, say, North America do not directly bring about soil erosion in a place like Africa, but loss of soil is still of great concern on both continents.

> *Remember from Chapter 3 that erosion is the removal or transport of lithospheric material.*

Causes: agricultural practices

Land clearing. Earlier in this chapter we saw how existing ecological communities are typically removed prior to farming.

Unfortunately, loss of vegetation leaves an area vulnerable to erosion because plants help to hold soil in place in two important ways. First, living roots create a network that physically stabilizes soil, as in **sand dune succession** (Chapter 5). Second, above-ground structures like leaves and branches can block or deflect incoming rainfall, preventing it from hitting the ground directly. Thus, soil is protected from the full force of the weathering power of water.

Plowing. As we saw earlier, this common practice can aid germination and growth of certain crops. However, plowing breaks large and relatively heavy aggregates into small and light particles, greatly increasing the likelihood that wind and water can easily carry them away (Figure 9.9).

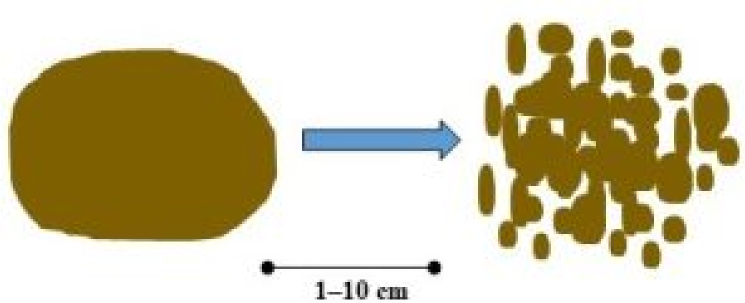


Figure 9.9. *Conceptual diagram of the effect of plowing on soil aggregates. Mechanical disruption can break a relatively large and stable soil aggregate into small, easily eroded particles.* Kelsey, CC BY-NC-SA.

Overgrazing. Farm animals often supplement their diet by eating grasses growing naturally in fields. The trouble starts when their grazing removes the plants faster than they can re-grow, or put another way, when carrying capacity of the field is exceeded. The outcome is essentially the same as that seen after the deliberate clearing of land: cows (or others) rip the grass out of the ground, leaving behind loose soil that can be easily removed during the next rainstorm (Figure 9.10).

Figure 9.10. *Overgrazing by goats can clear land. Lichinga, CC BY.*

Furthermore, overgrazing changes the **structure** of the ecosystem in a field because it favors plants that can quickly grow in response to cleared land. Reduced **biodiversity** is typically a result (Chapter 6), because few organisms possess the traits necessary to thrive under such conditions.

Relevance: loss of topsoil and reduced soil fertility

Topsoil erosion is a natural part of the rock cycle (Chapter 3), and materials weathered, transported, and redeposited can be incorporated into new soils as described in section 9.3. Under natural conditions, these processes balance each other sufficiently to ensure no net change in topsoil availability. Farming, however, has dramatically accelerated the pace of losses, and global rates of erosion are currently many times those of replacement[2]. Drawing upon the symbology and concepts from systems analysis (Chapter 2), we can conclude that agricultural systems are experiencing a net loss of topsoil (Figure 9.11).

This trend is problematic for farmers and limits food production because topsoil is the only layer in the soil profile that contains relevant amounts of soil organic matter. Thus, as topsoil is lost, SOM and the services it provides (described in section 9.3) are lost as well.

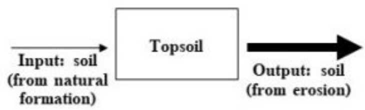


Figure 9.11. *Systems analysis of soil depletion. The rate of removal exceeds the rate of replacement (output arrow is relatively large), so the amount of topsoil in the system is reduced. Kelsey,* CC BY-NC-SA.

Note that (**positive feedback** only makes matters worse (see Chapter 2) because of the ways SOM helps soil resist erosion. In other words, as it is removed, the rates of its future removal will increase. Eventually, soil fertility drops so low that crops will no longer grow.

How to minimize loss and conserve soil

The United States Department of Agriculture, along with analogous state agencies, has developed a set of formal guidelines farmers can follow to help preserve their soil. Many other countries promote similar practices. The strategies described below are a sampling of recommended **soil conservation methods** that could, if widely adopted, dramatically reduce rates of erosion.

Contour farming. Cultivation on the side of a hill can slow the downward movement of water and sediments if crop rows are

2. Pimentel, D, Burgess, M. 2013. Soil erosion threatens food production. Agriculture. 3:443-463

perpendicular to the direction of the slope rather than parallel to it (Figure 9.12).

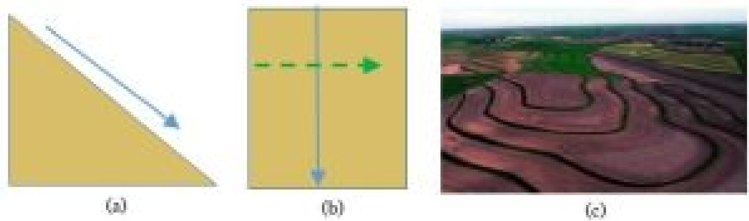


Figure 9.12. *Diagrams of a field on a hillside: cross section (a); map view (b); water moves downhill in the direction of the solid arrows. Rows of crops are perpendicular to direction of water flow (dashed arrow). Aerial photograph of contour plowing in Missouri, U.S.A. (c). Kelsey, CC BY-NC-SA (a and b); USDA, Public Domain (c).*

Cover crops. After a crop such as corn or wheat is harvested, a field may lie empty for many months, leaving its soil vulnerable to erosion. To combat this problem, another plant can be sown in the field after the main crop has been harvested. There are at least two important advantages of this practice. One, the second plant helps hold the soil in place in the offseason. Two, some nutrients necessary for future crop production can be replenished if certain cover crops are used. For example, clover or alfalfa are often chosen because they house **nitrogen-fixing bacteria** inside their tissues (Chapter 4). This **mutualistic association** (Chapter 5) is extremely important because the host plant receives abundant useable nitrogen, and the bacteria are given a suitable habitat and **fixed carbon** (Chapter 4). Since the mature plants are generally plowed into the soil rather than harvested, decomposition will make that fixed nitrogen available for the food crop planted in the next growing season.

Windbreak (or shelterbelt) establishment. Rows of trees can be strategically planted near crop fields to help block wind and reduce the soil erosion it causes (Figure 9.13). The presence of trees can also add to the **biological diversity** of an area, which could help combat insect pests (more below).

Figure 9.13. *A shelterbelt of trees helps protect an adjacent field from wind erosion. MJ Richardson, CC BY.*

No-till farming. So far in this chapter we have assumed agroecosystems utilize mechanical plowing or tilling as a matter of course on all crops. However, an increasing number of farmers (but still less than a majority) choose to sow seeds in fields that have not been disturbed or turned over. Their crops are grown on top of plant residues from the previous year (Figure 9.14 shows a field plowed prior to the sowing of seeds and a field that was not plowed).

Figure 9.14. *Seeds planted in plowed field, left (a); soybeans planted on top of crop residues from the previous year in no-till agriculture, right (b). Volker Prasuhn,* CC BY-SA *(a); USDA,* *Public Domain (b).*

The advantages of this practice are likely obvious by now: soil will be less vulnerable to erosion if its aggregates are not broken down into small fragments. There also are disadvantages. As we saw above, plowing can enhance crop growth. Furthermore, pest management becomes a particularly important concern in no-till agriculture, often necessitating more pesticide use (we will return to this topic soon). Since fertile soil is a limited resource, though, one that is being diminished faster than it is replaced, more and more farmers may elect this approach.

9.4.2. Sediment pollution

Our second adverse consequence is directly linked to the first, soil erosion. What happens to all the soil carried away from farms by wind and water? Surely it does not just vanish, so where does it go? In short, much of that soil ends up moving into bodies of water (and called "sediments"), bringing about substantial changes to those aquatic ecosystems. This effect of agriculture, referred to as **sediment damage**, can be observed on both local and regional scales. Eroded soils pollute the small streams that often flow across farmland and then move hundreds or even thousands of kilometers to rivers, lakes, and oceans.

Causes

The practices that drive the removal of soil from agricultural fields are also responsible for sediment damage. So, land clearing, plowing, and overgrazing all cause an excess of sediments in waterways.

Relevance

Sediment damage reduces water quality. As we have seen repeatedly, organisms in any environment are there because they have adapted to the pressures placed on them. Fish living in a river, for instance, are accustomed to the clarity and temperature of their water as well as the amounts of oxygen and other vital materials in their habitat. Alterations to any of those characteristics will favor different organisms. Two specific concerns are briefly described here.

Increased turbidity. Water becomes more **turbid**, that is, cloudier, when sediments are added to it (Figure 9.15). Many changes can then take place. First, less sunlight penetrates the water, which reduces rates of photosynthesis (Chapter 4). Among other things, less oxygen will be released into water, challenging or killing aerobic organisms. Second, fish sensitive to the stress of more physical objects passing over their gills will begin to die. Third, the added sediments absorb heat, leading to higher water temperature. This last change drives even more oxygen out of the water and it puts pressure on organisms that have become adapted to lower temperatures. Under extreme conditions, water quality becomes so poor that virtually all fish disappear from an aquatic system, leaving behind certain insects, worms, microorganisms, and other organisms adapted to life in murky, polluted water.

Figure 9.15. *These water samples were artificially polluted with sediments. Cloudiness, or turbidity, increases from left to right. USGS, Public Domain.*

Increased amounts of nutrients. Soil contains many nutrients, including biologically available forms of nitrogen and phosphorus. **Eutrophication** may therefore result from sediment damage.

How to minimize sediment damage

Recalling that this effect is brought about by soil erosion, it is fair to say that the soil conservation methods described above would also reduce sediment damage. Prevention, in other words, is an effective way to preempt the problem.

9.4.3. Chemical pollution

Conventional agriculture generally employs chemical substances, collectively referred to as **agrochemicals**, to maximize crop yields with a minimal financial cost. Two basic types of chemicals are used on farms: stimulants and limiters. As we will see, each type of additive can cause local and regional effects. Before we proceed, though, the following clarification might be helpful. The sediment pollution described in the previous section is also termed physical pollution because our primary concern with it is linked to the presence of objects floating in the water. Alternatively, chemical pollution describes compounds that are largely dissolved in water (or other relevant medium) and typically invisible to the naked eye.

Fertilizers stimulate plant growth

Recall that organisms require energy and materials to survive, grow, thrive, and reproduce. Like all plants, crops are **photosynthetic**, so neither fixed carbon nor energy are likely to be limiting. However, other critical compounds, for example those containing nitrogen, phosphorus, and potassium, are often available in only very low quantities. These deficiencies can be addressed by addition of fertilizers, substances rich in nutrients. Here we briefly consider some of the ways this practice is used to stimulate the growth of desired crops.

Types. Conventional farms often use synthetic fertilizers, that is, those that have been manufactured, processed, and packaged in laboratories and factories. In other cases, compost or animal manure can be applied to fields. This second strategy is possible because, as we already learned, waste products released from organisms contain a lot of useable nutrients.

Benefits. Fertilizer can have an enormous impact on crop yield. Without it, growth can be stunted and food production low, but, if added appropriately, growth can be successful and profitable.

Costs. Fertilizers are generally helpful, but their use can also lead to many undesirable outcomes.

Water pollution

As will be described in more detail in Chapter 11, agriculture can contaminate both human drinking water and natural waterways. Although it might seem odd that necessary nutrients like nitrogen can be harmful if they end up in water, you should be aware that the types of nitrogen added to fields are toxic to animals. Moreover, once they enter soil, some of these compounds can be converted into forms that are even more problematic, which then move vertically to the groundwater reservoir with flowing water (Chapter 4). For example, blue baby syndrome can arise from ingestion of a natural product of the nitrogen cycle. This condition can interfere with the body's ability to transport O_2 to the brain, causing death in some cases. Eutrophication is another potential consequence of fertilizer use (Chapter 4).

Stimulation of weeds

As we know, weeds thrive in the same kinds of conditions favoring crops. The addition of fertilizers exacerbates the situation because these unwanted plants also use the nutrients to grow more effectively.

They must be purchased

Any list of the costs of fertilizers would be incomplete without a note about a more familiar type of cost, the financial kind. Fertilizers are not free of charge, so a farmer must determine if increases in crop production justify the investment required to provide supplemental nutrients.

Reducing or eliminating fertilizer use. You should keep in mind that the *right* amount of nutrients needs to be present to support crop growth. Think back to our exploration of biogeochemical cycles in Chapter 4 and recall that there are **critical limits of tolerance** for chemical elements in any system. The same principle applies to agriculture: insufficient amounts of things like available nitrogen will limit food production but too much will lead to water pollution and cost more money than needs to be spent. With that guidance in mind, we can consider ways to minimize the amount of fertilizer that must be applied to fields.

Nitrogen-fixing cover crops

In the previous section about erosion we saw that certain cover crops can help stabilize a field and minimize soil loss. They can also replenish the supply of available nitrogen compounds for next season's crops.

Crop rotation

Earlier we noted how monoculture can both stimulate pests and lead to nutrient depletion. One effective way to reduce the severity of those unwanted consequences is to plant different crops in different fields each season. Monoculture is still employed, but this year's corn field will be used to grow wheat next year, the wheat field will be used to grow corn, and so forth. Ideally, one of the many fields a farmer owns is left fallow, that is, no crop is grown in it at all (except a cover crop that is not harvested) for one season. Crop rotation can reduce the need for fertilizer because, although they all have the same fundamental needs, different species of plants use somewhat different specific amounts of nitrogen, phosphorus and other nutrients. So, if a crop such as corn was grown over and over in the same field for years the nutrients it needs the most would be used up faster than if it and other crops were moved around from year to year (see Figure 9.16 for a hypothetical map of crop rotation for three years).

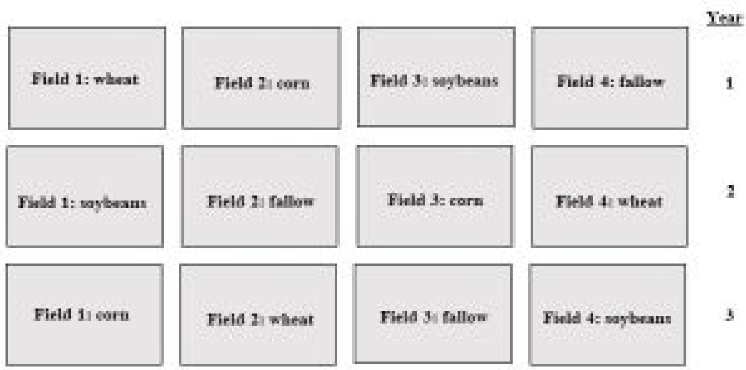


Figure 9.16. *A farmer might use the rotation suggested here. Fallow, no crop is grown. Kelsey,* CC BY-NC-SA.

No-till agriculture

The elimination of plowing and other processes that break up soil is another strategy we encountered during our discussion of erosion, above. In addition to reducing soil loss for the reasons already described, this method can minimize the need for fertilizers because remnants of the previous year's plants stay in the field instead of being completely removed in preparation for upcoming growing seasons. The remains undergo decomposition, and the nutrients stored in their dead tissues are released and made available for the next generation of crops.

Hormones stimulate animal growth

Earlier in this chapter we saw that these chemicals are sometimes given to livestock to promote growth.

Benefits. This practice is used because it can lead to an increase in food production without an increase in the number of animals present. For example, the amount of milk produced per cow goes up if certain hormones are mixed into feed.

Risks. Several adverse consequences are possible, leading some countries, particularly those in Europe, to greatly restrict the use of hormones in livestock; the same regulations do not limit U.S. farmers, though.

Harm to the animals

Data suggest that cows receiving hormones can suffer adverse effects, including abnormal development.

Harm to human consumers

Since hormones play an important role in processes related to growth, development, and sexual function, many people have expressed concern that human diseases could arise from consumption of food containing residues of these chemicals. More research needs to be conducted, but some scientific evidence has linked certain cancers and other disorders, particularly in children, to hormone ingestion (we will see more about human toxicity in Chapter 15).

Harm to natural ecosystems

Waste products from livestock can move into soil and water, carrying hormone residues with them. Both individuals and communities exposed to such chemicals could experience adverse consequences. For example, the presence of hormones in waterways has been shown to disrupt the reproduction of some amphibians (also addressed in Chapter 15).

Reducing hormone use. The financial benefits of hormones make them desirable for some farmers. The alternative is to cease the use of hormones and add more individuals to herds to make up for the difference in production (which would require additional resources, of course). As stated in the previous paragraphs, more scientific research could help guide the development of regulations and practices.

Pesticides inhibit insects and weeds

Unwanted organisms, **pests**, can greatly reduce the production of food on farms. Estimates vary, but on a global basis, the combination of weeds and insects has the potential to reduce yields by somewhere between 20 and 50% (or perhaps more). Consequently, much effort is devoted to controlling these destructive organisms, including the widespread use of pesticides.

What are pesticides? The Greek suffix "cide" refers to killing, and the prefix "pest" refers to something that is bothersome or harmful. The literal translation of this word, then, essentially sums up what these chemicals do: they cause the death of unwanted organisms.

How do pesticides work? Pesticides rely on the basic principle of **selective toxicity** described in Box 5.6 (also Chapter 15): they are poisonous to some, but not all organisms. A well-designed pesticide will not affect humans or other organisms we wish to preserve (referred to as **non-target organisms**) and damage those that are undesirable (called **target organisms**). We will see shortly that there is a limit to just how selectively toxic chemical pesticides can be, though.

Types of pesticides. Pesticides are classified using several criteria.

Specific type of pests they target

For example, insecticides kill insects, herbicides kill weeds, rodenticides kill rodents, and so forth.

Number of different organisms a pesticide is intended to affect

Pesticides with a **broad spectrum** are general in their targets, intentionally killing many organisms, whereas those designed to damage a relatively small number of species have a **narrow spectrum**.

Source or origin

Compounds can be placed into one of two general categories. The first, called **synthetic pesticides**, includes those manufactured by humans. Atrazine, Roundup, Malathion, Sevin, and DDT are but a small sampling from this group. On the other hand, **natural pesticides** (or biopesticides) are obtained from organisms, that is, they are not synthesized by humans. For example, **pyrethrum** is an insecticide produced by plants. Poisons made by fungi and bacteria can be used to kill insects as well. By and large, synthetic pesticides are more toxic to humans and have a broader spectrum than do biopesticides.

Other characteristics

The pesticide story is quite complicated and includes considerations about how to apply them, and when during the season they are most effective, issues that are unnecessary for our discussion.

Benefits. Herbicides kill plants that would otherwise compete with crops, and insecticides kill caterpillars and other herbivores that eat crops. Unrelated to farming, yet important: they also kill insects that carry human diseases (more in Chapter 15).

Risks. Scientists estimate that over 95% of pesticides applied to crops go someplace *other* than their intended target. Unsurprisingly, then, pesticide use can bring about many undesirable consequences.

Human health can be affected

Arguably, ours is the non-target species that gets the most attention when it comes to adverse effects of pesticides. Although we are different than plants and insects and should be protected by the principle of selective toxicity, certain herbicides and insecticides still have the potential to cause human disease. For example, because it disrupts plant hormones, you might expect that the widely used weed killer 2,4-dichlorophenoxyacetic acid (known to most of the world as 2,4-D) would be safe for humans. Some scientific studies suggest otherwise, though: 2,4-D may damage human reproduction by disrupting our own hormones. Other herbicides also seem to pose similar risks. The insecticide DDT (more below and in Chapter 15) is another example of a pesticide that ought to be safe for humans because it specifically targets the nervous systems of insects. It is true that we are not affected the same way target organisms are, but DDT can cause a different set of problems instead. Notably, it has the potential to disrupt reproduction and other processes controlled by hormones in humans. Some people suspect it can also cause cancer, but at this time insufficient convincing evidence supports that hypothesis.

One final note is appropriate here. Humans can contact pesticides through multiple means, including the consumption of contaminated food, but workplace exposure is likely the most important. In other words, the amount of pesticide one will encounter on a piece of fruit from a grocery store is usually very small—so low, in fact, that it will probably have no relevant impact (more about the effects of toxins on human health is presented in Chapter 15). However, a person working in an agricultural field can be exposed to dangerously high levels of pesticides. Consumption of contaminated water is another way these poisons can enter humans, a topic discussed shortly.

Other non-target organisms are affected

Along with killing harmful insects that eat ears of corn or foliage of soybeans, an insecticide can also harm beneficial organisms. Many examples could be cited, but we will limit our discussion to just three. First, malathion is widely used to control insects that eat the leaves of agricultural crops; it also is highly toxic to much-needed pollinators such as honeybees and can affect fish and birds as well. Second, the herbicide atrazine is a very effective way to kill weeds because corn is resistant to its toxic effects. Unfortunately, though, despite what might be our expectations of an herbicide, atrazine can adversely affect animals such as frogs by disrupting their reproduction. Third, the insecticide DDT was widely applied to farms and elsewhere until it was banned in the United States in 1972 (although it is still used to some extent around the world). Its low cost and effectiveness made it a popular choice. Due to a combination of its chemical properties and unexpected behavior in natural systems, however, it was ultimately linked to the deaths of large predatory birds such as bald eagles. More on DDT and the limitations of selective toxicity—including how an insecticide can affect eagles—may be found in Chapter 15.

Water and soil can be polluted

Pesticides that do not remain on their targets move into soil and water. Once there, they can affect soil **communities**, including those made up of microorganisms, and flow to ground and surface waters. The herbicide atrazine we saw in the previous paragraph, for instance, has been found in some drinking water reservoirs, raising concerns about the human health effects it can cause.

Insects can develop resistance to pesticides

A pesticide kills organisms susceptible to its toxic effects. However, it also provides an advantage to those few individuals unaffected by it. As we encountered in our discussions of **antibiotic-resistant bacteria** (e.g., Box 5.6), and **evolution** (Chapter 6), a population can quickly change from one sensitive to a pesticide like DDT to one that can survive in the presence of that same toxin. Since they reproduce much more quickly than humans (but slower than bacteria), pesticide-resistant insects can emerge within a few years of the first use of the compound.

Alternatives to synthetic pesticide use. Acceptable levels of pest control often can be achieved through means other than the use of chemical pesticides. Mostly because of issues related to cost and convenience, though, a large majority of farmers do not rely exclusively on these strategies.

Biological control

In some cases, farmers can take advantage of natural **predator-prey relationships** to reduce the number of insects eating their crops. They can either create and maintain environmental conditions that favor predators of pests or deliberately import such organisms onto their farms. For example, an extremely small wasp known as *Trichogramma* is often used to kill unwanted caterpillars through a form of parasitism. Lady beetles and predatory wasps are also effective biological control agents (Figure 9.17).

Figure 9.17. *A parasitic wasp is attacking an alfalfa-eating aphid in an example of biological control. division*, CSIRO, CC BY.

Crop rotation

This practice was described in the context of fertilizer use (review Figure 9.16), and it is also helpful in pest control. If done carefully, annual shifting of crops among different fields will limit the amount of useable food available to the insect young left behind from a previous growing season.

Figure 9.18. *Different crops are grown in proximity to each other in the practice known as intercropping. manfred.sause@volloeko.de*, CC BY.

Intercropping

Instead of a single plant in one large field, several crops are cultivated in alternating rows within the same field (Figure 9.18, left). As a result, the advantages monoculture provides to insects are reduced.

Removal by hand and physical barriers for weed control

Weeds can be physically pulled out of the soil by farm workers, a practice that persisted for centuries prior to the wide-spread availability of herbicides. This very labor-intensive approach is not practical on large farms, but it can be seen on smaller scales. Alternatively, covering uncultivated soil between plants with tarps and similar objects is a very effective means to prevent weed growth, but it too is most relevant for small farms.

Integrated pest management (IPM)

Rather than relying on a single strategy to control pests, IPM combines many practices to reduce, but not eliminate, the use of chemical pesticides. The relative importance of the different tools—biological control, intercropping, crop rotation, pesticides (both synthetic and biopesticides)—varies from farm to farm and depends on local conditions. Farmers using IPM generally look at whole ecosystems rather than just a single species to manage pest problems.

Antibiotics limit the growth of disease-causing bacteria. The use of antibiotics is the final source of agricultural pollution on our list. Often, particularly on large, commercial farms, animals are fed antibiotics. At first, the strategy might appear to have some merit: money is lost if disease renders livestock (or their products) unfit for sale. In practice, though, widespread use of antibiotics comes with some serious risks. Specifically, and this story should sound familiar by now, bacteria have the ability to rapidly adapt to changing environmental conditions including the addition of poisonous substances. In the case of agriculture, antibiotics moving with animal excrement to soil and water encourage the development of antibiotic-resistant bacteria in those natural systems. Since some affected populations are likely pathogens, people could be sickened by diseases that cannot be easily cured. Livestock animals we wish to protect are similarly at risk.

9.4.4. Desertification

Defined simply, this is the process whereby a desert is formed, through natural or anthropogenic forces, in a temperate or other biome.

Causes: less water is available

Given what we know about environmental conditions in desert biomes (Chapter 5), it should come as no surprise that desertification is triggered by reduced availability of water in an area. But just what is powerful enough to effect such a change in water supply? Although a number of forces and activities contribute, agriculture plays a substantial role in the process.

Reduced precipitation. Perhaps the most obvious reason for desertification is a lack of input of water via important pathways such as rain, snow, and the like. A great deal of geologic evidence supports the notion that climatic conditions, including rainfall patterns, have changed throughout Earth's history. Similar natural forces surely are active today as well, although human activities are also accelerating climate change. As a result, some temperate areas are becoming drier (among other effects). We will encounter this and related topics again in Chapter 14.

Reduced water-holding capacity. Since the amount of organic matter present (SOM) in soil is directly proportional to water-holding capacity, anything that removes SOM will also lower water availability to soil organisms. So, repeated farming of land can speed up topsoil erosion and yield increasingly sandy soil with time. Even if the amount of precipitation falling to the ground does not change, less water will stay around after rainstorms end, depriving plants of a vital resource. Then what happens? Well, as less and less vegetation is present, the soil becomes more and more vulnerable to erosion—fewer roots means less stability, as we know—and even more topsoil is lost. Future generations of crops will be even less healthy, more soil will be lost, and it goes on and on until all that remains is a desert (Figure 9.19).

Figure 9.19. *Loss of SOM due to agriculture has contributed to desertification on this Egyptian farm. Kamal Osama Elgazzar*, CC BY-SA.

Desertification is generally considered to be a regional phenomenon because as a local farm becomes more of a desert, larger-scale precipitation patterns tend to change, causing surrounding non-farm areas to undergo similar transformations from relatively wet to relatively dry. Yet again we see the nature of positive feedback: the system responds to its own output with accelerated rates of change. In this case, as a farm starts to become a desert, the rate at which this shift occurs increases and the size of the area affected grows. If we expand our view further, we can see how erosion and desertification are connected to the growing human population, demand for food, and number of large, commercial farms. In short, this much larger system is likely to be characterized by positive feedback for the foreseeable future.

9.4.5. Redistribution of water

Agriculture accounts for about 40% of domestic freshwater usage in the U.S.[34]. Irrigation, especially in relatively dry climates such as southern California, can withdraw enormous quantities from surface and groundwater reservoirs and lead to shortages and conflicts among various interested parties. We will return to the complex issue of water usage and its consequences in Chapter 11.

9.4.6. Redistribution of carbon

This is the first of two global effects of agriculture on our list. Here, we briefly examine some of the ways farming increases the rate

3. US Geological Survey. 2019. Total Water Use. https://www.usgs.gov/mission-areas/water-resources/science/total-water-use?qt-science_center_objects=0#qt-science_center_objects

4. The term 'sector' is used to indicate a class of water users. For example, the **agricultural** sector includes farms and is the largest user of water. The **power generation** sector can use a lot of water as well, although the amount varies from place to place. The **industrial** and **domestic** sector (homes) tend to use much less water than the others. Consult Chapter 11 for more about sectors and water usage.

at which carbon (as **CO_2** gas) moves from the lithosphere to the atmosphere (see Chapter 14 for more on this topic).

Increased rates of decomposition

Since the environmental conditions inside undisturbed, dense soil clusters are not conducive to decomposition, organic carbon can persist in the lithosphere for centuries. Once a plow breaks soil into small and loose fragments, though, fungi and bacteria can quickly go to work and convert that stored carbon to CO_2 gas.

Use of fossil fuels

Modern agriculture, most notably that carried out on large farms, uses an enormous amount of fuel to power mechanized equipment (tractors, bailers, and others). The organic carbon compounds in those petroleum products (typically gasoline and diesel fuel) are burned, and the resultant carbon dioxide moves into the atmosphere (Chapter 10). Fossil fuels are also used in pesticide synthesis and application, although they are not necessarily immediately converted to inorganic carbon during these processes.

9.4.7. Changes in Earth's albedo

This is the second of our two global concerns. As we noted in Chapter 4 (section 4.3), the exposed, brown soil and small plants typically grown on farms reflect less energy than do the forests that were likely cleared to make room for agriculture. On a small scale, the effect of farming or gardening on albedo is trivial. However, large, continental-scale changes to land usage and surface properties could ultimately increase the amount of energy absorbed, as well as the average temperature of the whole Earth, in appreciable ways. See Chapter 14 for more.

9.5. GENETIC MODIFICATION OF ORGANISMS

We learned in Chapter 6 that **genetic information** is stored in **DNA** and that the specific codes carried in this molecule control the traits of all organisms, including crops and livestock. Agricultural scientists and farmers take advantage of the fact that DNA can be modified to increase our ability to produce food in many ways. Some of these practices are old, time honored, and trusted whereas others are new, relatively untested, and the cause of much debate.

9.5.1. Desirable traits are encouraged

Recognizing that individual organisms vary in their suitability for agriculture, farmers have been optimizing food production by manipulating genetic characteristics in their crops and livestock for centuries. What they do is related to, although not the same as, the process of **natural selection** we encountered in Chapter 6. Recall that individuals that are the most fit for survival are the ones that tend to reproduce, and the characteristics of subsequent generations will reflect the traits that maximize the chances for survival. Agricultural practices artificially manipulate which individuals are allowed to pass on their traits to the next generation. The ability to survive is not necessarily the most important trait to a farmer, rather, other advantages they bring to food production will be favored instead. Practices used to manipulate genetic characteristics and improve agricultural products can be divided into two groups: those that are conventional and widely accepted and those that are newer and controversial.

9.5.2. Approaches to genetic manipulation

Conventional genetic engineering

Selective breeding. A farmer could notice that a few members of a population grow faster, bigger, or otherwise better than the bulk of the others and then intentionally produce offspring from those most desirable individuals. Those offspring would likely have the characteristics of interest and would also be selectively bred until the whole population has been converted.

Cross breeding. Two individuals might each have some, but not all, traits that are desirable—say, one grows unusually large and the second produces uniquely sweet fruit. Offspring possessing all the advantageous characteristics could be produced from crossing these two plants. There are many examples of this practice, only two of which are noted here. First, the many varieties of different apples available for purchase at a supermarket have not always been present on Earth. Instead, they were produced by combining plants with different desired attributes. Other crops like corn and tomatoes have been similarly developed through this practice. Second, Holstein cows, the black and white ones generally used for dairy production in the U.S., are not altogether natural organisms. They are the result of centuries of cross breeding among different types of cattle to produce individuals with optimum traits. The Holstein has been specifically bred for maximum milk production, not necessarily for its meat (although those cows eventually end up slaughtered in any case).

New approaches to genetic engineering using biotechnology

Biotechnology is a general term that refers to various kinds of manipulations of organisms for the purposes of agricultural, industrial, commercial, or medical gains. In the context of agriculture, it can be used to affect specific DNA sequences controlling certain physical traits of crops and animals. Put into commonly used language, biotechnology can help produce **genetically modified organisms (GMOs)**. Unlike cross breeding and other practices which require the relevant organisms to have enough in common that they could plausibly produce offspring without human intervention, these new methods can combine genes from unrelated organisms, those that would not be able to breed under natural circumstances. Genes from many organisms have been combined using these relatively new techniques. For example, plants have been crossed with each other, animals, and even bacteria. The use of biotechnology on farms is a very complex subject featuring both scientific rigor and controversy. Since exhaustive treatment of the topic would require much more space that we have available, we will focus attention on a few major points.

Potential benefits. The reasons GMOs were first developed and continue to be used are numerous. Generally, genetic manipulation that employs biotechnology is desirable because it allows for more targeted and faster modifications than do traditional cross breeding and related practices. A few specific examples of the kinds of advantages that can be realized through this practice are described here.

Crops can fight insect pests

A piece of bacterial DNA can be inserted into the genomes of corn (and others), giving plants the ability to produce insecticides on their own. One application of this idea resulted in a crop commonly referred to as "Bt corn", so called to reflect how it was combined with the prokaryote *Bacillus thuringiensis*. We will return to this case in the section about risks of biotechnology, below, but for now note the benefit of this practice: the need for synthetic pesticides is greatly reduced on such farms.

Crops can survive herbicide exposure

Genes providing resistance to the toxic effects of glyphosate (the active ingredient in Roundup) have been transferred from bacteria to some agricultural crops. Now weed control is relatively simple: crops are the only plants left alive after fields are blanketed with the **herbicide**.

Nutrient content of crops can be enhanced

Genes from daffodils and a certain soil bacterium were inserted into the genome of rice to increase the amount of beta carotene in the crop. The result is a yellow product known as "Golden rice" (Figure 9.20). This modified rice has supplemented the amount of vitamin A available to people in areas that lack the critical nutrient.

Figure 9.20. *GMO golden rice and white rice. International Rice Research Institute, CC BY.*

Crops could withstand stresses of drought and temperature

Research into ways to use biotechnology to increase survival and success of plants under various hostile environmental conditions is ongoing. Since low water availability often limits crop yields, genes from desert plants could be combined with those of corn, rice, or other crops. Similarly, freezing can be a problem with tomatoes and other sensitive plants. Again, genes from plants that have adapted to survive the cold—like those from northern regions—could be inserted into the genomes of crops to produce more robust, resistant crops.

Risks, concerns, unanswered questions. Given the potential benefits described above, along with many others not mentioned, why are some people opposed to the use of GMOs? Several concerns, both scientific and otherwise, are repeatedly cited.

Risks to human health?

Put simply, people worry about the adverse effects that ingestion of GMOs can cause in consumers. Novel DNA sequences lead to the production of new, and potentially toxic, proteins by plants. In principle, negative consequences could range from mild allergic reaction to death. Various governmental agencies do regulate GMO use in human food to protect public health, but not everybody is satisfied. In the United States, genetically modified *crops* can be used in the production of both human food and animal feed, but it is illegal to apply biotechnology in the modification of *animals* used to make human food. Under certain circumstances, these products must be evaluated and shown to be safe, but many observers question the rigor, validity, and objectivity of the tests used (among their concerns: much of the research is conducted or funded by companies that profit from the sale of GMOs). Critics also note that food is not required by law to be labeled to reflect its GMO content, noting that consumers ought to be fully informed as they make nutritional choices. Elsewhere in the world, particularly in Europe, prohibitions against the use of GMOs are much stricter.

Risks to farms and natural environments?

Because of the number of published studies suggesting that GMOs present a small risk to human health, many scientists focus more on the potential threat that GMOs pose to Earth's natural systems. Many concerns and unanswered questions remain.

Transfer of advantageous traits to weeds. Since genes can be passed among organisms through means other than reproduction, the ability to resist herbicides and other stressors might be transferred to unwanted plant species. The resulting organisms, sometimes referred to as "super weeds", would be formidable competitors, potentially taking over farms as well as natural ecosystems. Changes in **ecosystem structure** would surely follow, including lowered **biodiversity** and damage to herbivores that rely on certain plant species for food (see Chapter 5). This phenomenon is poorly understood, and more research needs to be conducted to improve our knowledge about it.

Development of pesticide resistance. In the section on potential benefits of GMOs, above, we saw that Bt crops could produce their own insecticide. However, as with synthetic pesticides, insects have the ability to adapt to it; a population can quickly evolve to one dominated by individuals able to survive the toxin produced by the plant. Weeds can also become resistant. For example, extensive use of Roundup will select for individuals able to survive in its presence, possibly rendering the herbicide ineffective in the future. Agricultural scientists are aware of these problems and have been working on ways to maintain populations that are sensitive to the pesticides used against them, but success in this area is not guaranteed.

Effects on non-target organisms. In addition to its role in the development of pesticide resistance, the use of Bt crops raises another question: does the pesticide produced by plants harm organisms other than pests? As we know from earlier discussions, pesticides are not completely selective and often harm beneficial insects. Bt toxin is the same. Initially, concerns were raised that Monarch butterflies were being killed because corn pollen grains carried on the wind to nearby milkweed plants (a Monarch favorite) contained the toxin in them. Subsequent research demonstrated little evidence for the specific Br corn-milkweed-butterfly connection, but the larger issue about non-target effects remains unresolved.

The effects of sterile seeds

Traditionally, crops were sown and harvested, and seeds were collected from them at the end of the season to be used in the next year. These days, if they wish to use a genetically modified crop, farmers typically must purchase the seeds every year. Why is this? Because GMO crops are manufactured to be sterile, that is, they will not produce useable seeds when they grow (laws prohibiting the unauthorized acquisition and use of GMO seeds are also on the books). Farmers and their advocates cite this as exploitative and inappropriate because large corporations gain too much control over our food supply. The companies that develop the seeds counter that they have a right to dictate what happens to their products and to profit from their work. Furthermore, they say their seeds will perform more predictably and reliably if the randomness that comes from reproduction is not introduced. Like many others related to GMOs, this subject continues to be a source of conflict.

Other, values-based, arguments

For reasons we have seen in detail elsewhere, these are outside the scope of science. People often try to use them to sway their opponents, though—particularly when it comes to GMOs—so you should be cautious to evaluate if what you hear or read is based on measurable, objective data or subjective ideas about what *ought* to be true. Also remember that basing one's actions on opinions, feelings, or convictions is not necessarily wrong, but it most certainly is *not* science.

What happens next? Genetically modified organisms are already being used extensively in agriculture. What the future holds depends on several factors, including our ability to meet demand for food, the work of advocates on both sides, public opinion, and our understanding of the risks and benefits of GMOs. Some people dismiss the idea of ever using genetically modified organisms on principle and will fight against them no matter what—some will even resort to sabotage, deceit, and intimidation. Others are ready to embrace GMOs under any circumstances, despite evidence that might warrant caution, using any means necessary to further their agenda. Scientists try to remain objective and are waiting for more data before making final assessments about the advisability of GMO use.

9.6. APPLYING OUR KNOWLEDGE: APPROACHES TO AGRICULTURE

Now that we have learned about agricultural processes and practices, we can briefly explore and compare some different ways food is produced.

9.6.1. Farming is different today than in the past

Since its beginnings many millennia ago, farming practices have changed a great deal. For example, both the size and scope of farms have grown enormously. People once kept small farms solely to feed their families (i.e., subsistence farms), but modern agriculture has become a large-scale industry employing millions of people to meet ever-growing global demand for food. The past century or so in particular has seen dramatic changes. According to the United States Department of Agriculture (USDA)[5], average farm size increased from about 200 acres in 1900 to about 440 acres in 2012, although American farms these days range in size from fewer than 10 to nearly 1,000,000 acres (globally, the biggest is about 6,000,000 acres). At the same time, fewer farms have become proportionally more important: today, only the largest 4% produce about two thirds of farmed food in the U.S. During the past few decades, though, there has also been an increase in the absolute number of small farms, perhaps indicating a growing interest in a return to local food production.

5. Farm Size and the Organization of U.S. Crop Farming. 2013. USDA Economic Research Service. Economic Research Report Number 152 PLUS 2012 USDA farm census

9.6.2. Some common types of contemporary farms

Farms vary in many ways, including size, type of food produced, strategies used to grow organisms, application of chemical additives, and use of genetically modified organisms. The earlier sections of this chapter present detailed information about the methods described below; consult them as necessary.

Conventional

This is a broad and diverse category of farms that currently produce most of the food in the United States.

Size. These facilities range from small, keeping just enough animals to feed the family living there (known as "rural-residence farms" by the USDA) to intermediate in size, up to a few hundred animals, to commercial farms, those that hold 500 or more cattle (or 1000s of chickens) at any one time (note that the actual maximum numbers in this last category can be very much higher, a hundred thousand or more animals in some cases). Tractors and other machines powered by fossil fuels (e.g., gasoline) are widely used.

Types of food. Plants directly eaten by humans and animals are produced here. Some common crops include corn, soybeans, and wheat. Livestock for human consumption such as beef, swine, and poultry are widely raised on conventional farms as well.

How food is grown. Again, acknowledging that farms differ, we make some broadly applicable observations.

Crops

Land clearing, plowing, crop rotation, and monoculture tend to dominate.

Livestock

Animals tend to be raised in confined spaces (e.g., pens and cages) often referred to as feedlots. Other practices can be used, though, including allowing animals to graze freely or simply have more space to roam (see Figure 9.21).

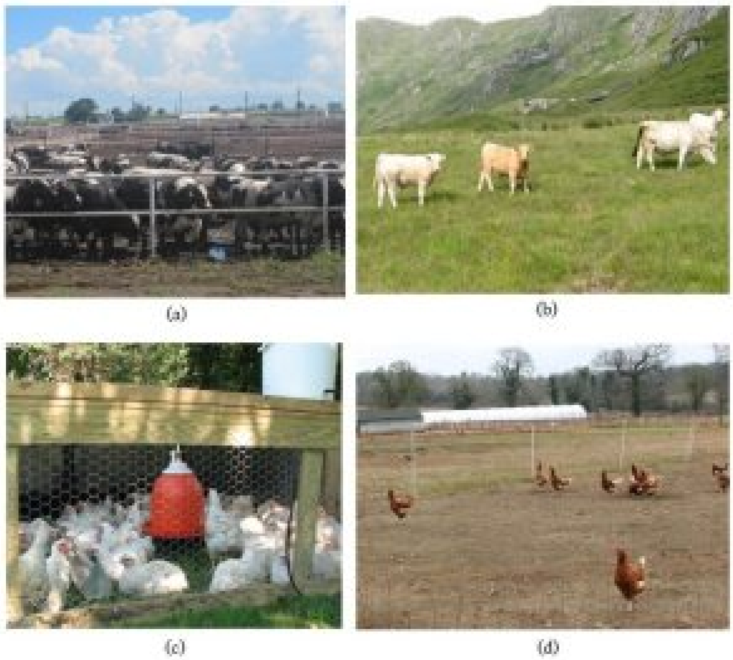


Figure 9.21. *Different ways to keep farm animals. Cattle raised on a feedlot, left (a); free-range cattle, right (b). Chickens can be kept in a variety of cage types, like the one on the left (c), or allowed more space, like the free-range chickens to the right (d). Billy Hathorn, CC BY (a); Jonathan Wilkins, CC BY (b); Chris Packard, CC BY (c); Evelyn Simak, CC BY (d).*

In any case, animals generally are fed large amounts of food and grown as quickly as possible. The waste produced can be used as fertilizer, but it also can move and pollute water. Finally, once they reach maturity (about two years for cattle, two months for chickens), the animals are slaughtered and their meat sold. Some, like dairy cows and chickens, are kept for their products, milk and eggs, respectively. They will likely be killed for their flesh, too.

Use of chemical substances. Nutrient deficiencies in fields are often mitigated through the application of synthetic fertilizers or manure, herbicides and insecticides are widely used to minimize losses, and hormones and antibiotics may be added to the diets of livestock.

Genetically modified organisms. Crops and livestock modified using biotechnology are often used.

Costs and benefits. We have seen the ways the practices used by conventional farmers provide both advantages and risks to human and environmental health in the earlier sections of this chapter. A point worth stressing here, though, is that food produced by conventional agriculture is nearly always less expensive for consumers to buy than that produced by organic means (described immediately below). Those who consume it either have decided its risks are acceptable or simply cannot afford more expensive products.

Organic

A farm can be certified as "Organic" by the USDA if it is operated according to an extensive set of regulations that dictate the conditions under which food must be raised (keep in mind that the conventional farms described above are not bound by these rules). We will refer to some of those items here, but the full list can be accessed on the USDA website[6].

Size. Farms here tend to be smaller—in some cases, by a great deal—than conventional farms. There is no size limit to be certified organic, but the type of work done on such farms is often most realistic at a smaller scale. Tractors and similar equipment are

6. https://www.ams.usda.gov/rules-regulations/organic

used, but more of the work is done by hand than is practical on large, conventional farms.

Types of food. The same kinds of food produced on commercial farms can be grown on organic farms. In addition, many of these small farms raise specialized plants, herbs, spices, and the like, for niche markets.

How food is grown.

Crops

Soil conservation methods tend to be more common on smaller organic farms. No-till agriculture and cover crops, for example, are often seen. Land clearing and plowing are still common, though.

Livestock

Animals are not fed food treated with synthetic chemicals and are also generally given far more space per individual than is seen on conventional farms. Organic farmers often strive to maintain animals in conditions that mimic natural ecosystems—so small cages and large feedlots are rare.

Use of chemical substances. The USDA standards for organic farming contain a list of allowed and prohibited substances, and farmers wishing to receive and maintain their certification must carefully adhere to the rules. Importantly, synthetic fertilizers and pesticides are prohibited, as are antibiotics and hormones.

Fertilizers

The application of animal manure and compost (See Box 9.1 for more on compost) to fields are common strategies used to replenish nutrients lost from soil.

Box 9.1. A few words about composting

Many farmers and gardeners use compost as a source of nutrients for the plants they grow. Just what is it?

Put simply, compost is a nutrient-rich material derived from organic products such as scraps of uneaten vegetables, lawn clippings, leaves, and even certain paper and clothing waste. If these materials are exposed to the proper conditions, the natural process of decomposition will transform them into useable fertilizer. So, for example, the waste products can be put into a container of some kind (structures specifically designed for this purpose can be purchased through commercial vendors) along with some soil and specially adapted earthworms and / or other insects. Once it is closed up, bacteria and fungi quickly go to work and convert the strawberry tops, potato skins, and all the rest into unrecognizable organic residues that combine to create a substance very much like the soil organic matter we encountered earlier. In some cases, the compost is ready to spread on a field within a matter of months. Why is it favored by certain people? First of all, composting can be used to recycle useable organic compounds that would otherwise be thrown away and lost to a **landfill**. Secondly, it is a way to replenish soil nutrients without using expensive and potentially risky synthetic fertilizers. We will encounter composting again in our discussion of waste management (Chapter 13).

Pesticides

Since synthetic chemicals are prohibited on these farms, one or several alternative methods are used to control unwanted organisms.

Weeds. Three practices are often used: removal of weeds by hand, physical barriers such as tarps between rows of crops, and physical disruption of roots by tilling or plowing.

Insects. Biological control, crop rotation, and intercropping are generally employed by organic farmers to keep these pests from destroying their crops.

Genetically modified organisms. Organisms manipulated using biotechnology (i.e., GMOs) are prohibited on organic farms. The use of plants and animals that have been modified and selected through cross breeding and hybridization are allowed, though.

Costs and benefits. One of the goals of organic farming is to minimize the environmental impact of food production. The prohibition against synthetic chemicals and antibiotics certainly reduces pollution of water and the development of antibiotic-resistant bacteria, respectively. It also tends to slow the rate of erosion through its use of soil conservation methods. Both of these outcomes clearly are good for human and environmental health alike. Whether consumption of organic produce is appreciably better for humans than is eating alternatives is a matter of some uncertainty and debate, however. Some have argued that organic produce has a higher concentration of nutrients in it than does conventional, although scientific data have not convincingly demonstrated any such link. Does it taste better? Again, no definitive, objective evidence can be cited to support any claims that it does. The consumption of fewer pesticides, antibiotics, and hormones likely has a positive health impact, but the magnitude of this effect still needs to be rigorously quantified. Finally, given USDA regulations and the increased amount of labor required to grow it, organic food is more expensive to consumers than are produce and meats grown through conventional means. Many people choose to spend more, though, because they feel its benefits are worth the extra costs.

Consider this conundrum: is it better to consume affordable vegetables grown on conventional farms to receive the many benefits they provide than no vegetables at all to avoid the risks they pose? It looks like we need to do a cost-benefit again (consult Chapter 1)!

Blends, intermediates, combinations

Conventional and organic are not the only possible types of farms, but they are often held up as two opposite ends of a spectrum. Whether or not that characterization is completely fair or accurate does not change the fact that farmers can adopt a mix of strategies in response to their philosophies, goals, access to resources, knowledge, and experience. So, for example, the owner of a small farm might elect to use soil conservation methods, only use compost for fertilizer, raise free-range chickens, but also apply small amounts of synthetic pesticides to control insects. Although not certified by the USDA, it would closely resemble an organic farm. On the other hand, large commercial farmers can adopt various practices usually associated with organic agriculture such as no-till cultivation, cover crops, and the use of large spaces for livestock to roam and feed.

9.6.3. Other, less common approaches

Mounting demand for both space and food has driven the development of many creative strategies to grow food. The historic notion of subsistence farming (mentioned in 9.6.1) appeals to certain modern people, even those who live in densely populated areas. Among other approaches used are community gardens (spaces shared by multiple families), greenhouse, rooftop (exactly what the name suggests!) and vertical (various structures are used to minimize the need for extensive horizontal space, see Figure 9.22) farming. Such strategies can certainly yield small amounts of edible crops but are difficult to scale up.

Figure 9.22. *Just one example of vertical agriculture. Note the plants growing on the side of a city building. Bright Agrotech, CC BY-SA.*

9.7. WHAT IS THE FUTURE OF AGRICULTURE?

As we know, Earth is closed with respect to materials like biologically available forms of nutrients, fresh water, fertile soil, and farmable land. How, then, will we be able to keep up with rising demand for food with no increases in the resources necessary to do so? Clearly, several issues must be resolved if we are to be successful.

9.7.1. More efficiency

Production of more food in the space currently available will be necessary. In other words, crop yields will need to increase somehow. Genetic modifications, either using traditional practices or biotechnology, may be part of the solution here. Issues regarding human and environmental health need to be fully researched and addressed, though, before GMOs would likely be widely accepted or even practical (as noted earlier).

9.7.2. Change in consumption habits

As noted in Chapter 8, both the total number of individual humans on Earth and the rate of resource use per person are likely to rise into the foreseeable future. In principle, a reduction in the number of excess calories consumed would reduce some of the stresses of agriculture, but in practice such a goal will likely be elusive.

9.7.3. Soil conservation

Topsoil loss is an enormous problem, one that, arguably, receives far less attention and concern than it merits. As we saw, current rates of erosion outpace rates of soil formation, so changes in strategy will be needed to ensure our ability to farm into the future. If things continue on their current trajectory, Earth's agricultural fields could be effectively barren within a century. Ready to feed that 11 or so billion people with no fertile soil?

9.7.4. Food production without soil?

It is possible to grow plants in water solutions, that is, **hydroponically**. Currently, the amount of food produced in this way is relatively trivial, but research in the area may increase its feasibility. Unless soil loss is slowed considerably, this and other soil-free growing techniques may be required in the future.

THE CHAPTER ESSENCE IN BRIEF [7]

Agroecosystems are used to produce human food. Although specific approaches to farming vary, all harness natural ecological processes and rely on knowledge of soil science. Agriculture tends to cause adverse environmental consequences and is associated with controversial uses of pesticides, fertilizers, and biotechnology.

---

7. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 9 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[8]

What variables influence the suitability of a soil to support crop production?

If nitrogen is so abundant in the atmosphere, how can it ever be a limiting factor for crop growth? Why do farmers often add nitrogen fertilizer to fields?

How could agriculture contribute to mass wasting (Chapter 7)?

Remembering that Earth is closed with respect to materials, what do you think is the best way to meet the ever-increasing demand for food by an ever-growing human population? Are any of the farming practices described in Chapter 9 out of the question for you? If yes, what are the alternatives?

8. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

10. Energy Usage and Environmental Pollution

JASON KELSEY

So far, we have focused on the way **energy** enables processes fundamental to the maintenance of Earth's natural systems, namely, the cycling of materials as well as the growth, survival, and reproduction of organisms. Now we consider how humans utilize energy in non-essential ways: to power the technology of modern societies and support the increasingly high standards of living described in Chapter 8. Given the large scope of this topic, Chapter 10 has been split into four sections. Part I is an overview and introduction, Parts II and III describe non-renewable and renewable sources of energy, respectively, and Part IV summarizes the present and possible future of energy.

Key concepts

After reading Chapter 10, you should understand the following:

- The three fundamental categories of human energy use
- That using energy is associated with unavoidable consequences

- How non-renewable and renewable energy sources are defined and distinguished from each other
- How different fossil fuels are formed by very slow, natural processes
- The environmental consequences of the widespread use of fossil fuels as energy sources
- How environmental scientists can predict the long-term supply of oil, natural gas, and coal
- The challenges, risks, and benefits of nuclear power
- The challenges and benefits of using renewable sources such as direct solar, wind, and moving water to generate power
- How biomass may be used as an energy source
- The distinction between primary and secondary energy sources
- How and why most of the world's energy demands continue to be met by fossil fuels despite the availability of relatively clean renewable energy sources
- How energy supply and usage might look in the future

Chapter contents

Part I: OVERVIEW AND INTRODUCTION

10.1. MODERN SOCIETIES REQUIRE ENERGY

Energy conversions help us meet a variety of demands and perform many types of work (Chapter 4). Here we briefly introduce some important terms and describe different ways we use energy.

10.1.1. Three categories of energy usage

Electricity generation

Any device you plug into a wall outlet—your lamp, air conditioner, computer, mobile phone charger—is directly powered by this important form of **energy**. Although it occurs naturally (e.g., lightning), humans have devised ways to artificially generate and distribute usable quantities of electricity. At the most basic level, the process is straight forward: if magnets are rotated around a copper wire, electrons in the wire flow, that is, electricity is produced. The challenge, then, is to come up with a way to spin something so those magnets move as needed. As you might imagine, a wide array of possible solutions is available. You could employ your hamster's wheel or your bicycle, for example, converting the energy in consumed food to electricity. Unfortunately, such small devices would yield limited amounts of usable **power**. Plus, of course, you and your pet rodent would tire after a while, halting the flow of electrons. These problems of scale and continuity of supply were overcome through the invention and development of the generator and steam turbine during the early 1800s and middle 1900s. Today, large power plants use the 19th-Century concepts of **thermal generation** to meet nearly all human demand for electricity. In simplified terms, these facilities convert energy stored in a fuel such as coal (there are others, as we will see) to heat, and that heat is used to boil very large quantities of water. Resulting high-pressure steam is directed at **turbines** and the magnets to which they are connected (Figures 10.1.a and 10.1.b).

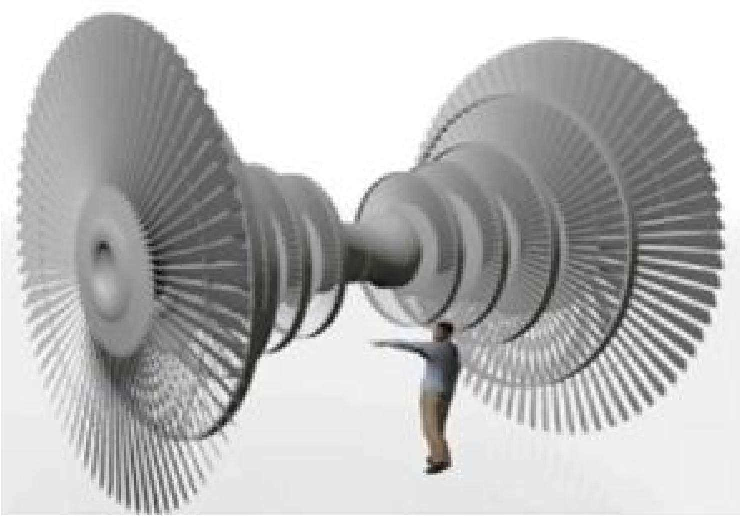


Figure 10.1.a. *Steam spins a turbine spins at 3000 rpm—note the person for scale. US DOE, Public Domain.*

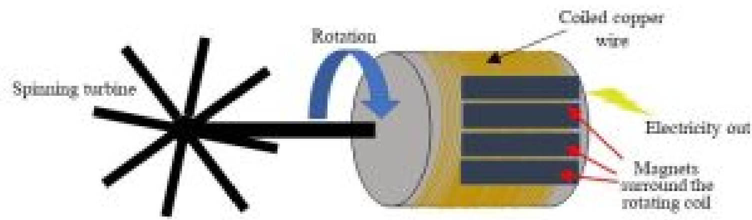


Figure 10.1.b. *A turbine attached to a generator produces electricity. Not to scale. Kelsey,* CC BY-NC-SA.

Steam is cooled, condensed, and then returned to the boiler to be turned into steam again. Given the large amounts of water used in the condensation process—it serves to lower the temperature of the steam—electricity generation plants are often built near natural reservoirs of the hydrological cycle (Figure 10.2).

Figure 10.2. *Power plants are often located near natural rivers, streams, or lakes. Pibwl,* CC BY-SA.

Electricity is then sent along wires into a larger network known as the **power grid**, or just "the grid" and distributed to homes, factories, and the like. The details of this last step are complex and not necessary for us to consider here; suffice it to say that electricity usage is far from uniform, changing with both location and time, and successfully meeting such heterogeneous demand requires extensive planning, monitoring, communication, and maintenance.

Note that the amount of **power** generated varies among facilities. Large coal-fired plants continuously meet the demands of hundreds of thousands of homes so long as the water is boiling and the turbines are spinning. Other types of plants can supply less or more power as a function of many variables we will explore later. Finally, recognize that the laws of thermodynamics (Chapter 4) govern and limit the efficiency of all energy conversions, including those that produce and distribute electricity. Among other important concerns, the amount of available electricity diminishes with distance from its point of generation—that is, more and more usable energy is lost as it moves along power lines to its point of use. So, electricity generated in, say, Arizona (USA), could not plausibly meet demands a few thousands kilometers away in the state of Vermont.

Temperature control

We generally strive to regulate the temperature of water and air in our homes, offices, and other interior spaces. When too warm, we turn to an electrical device such as an air conditioner or freezer to chill them. If cooler than we would like, though, we use hot water heaters or furnaces to convert chemical energy to heat and raise their temperatures. The process resembles that of electricity generation in that fuel is burned, often natural gas or oil, but here we do not go to the next step of spinning a turbine. Instead, hot water and / or air is distributed to a building's climate control and plumbing systems, often with the aid of an electricity-powered blower or analogous machine, as desired by human users.

Transportation

The ability to reliably and regularly move people, animals, food, and other objects both short and long distances has become central to the functioning of modern societies. As is well known, coal and petroleum products have served as portable fuels to power nearly all this work—transportation—since the beginning of the Industrial Revolution and appear poised to continue in that role for some time (much more about these energy sources is described in Part II, below). For example, when gasoline is burned in an internal combustion engine, chemical energy is released and converted into kinetic energy: pistons move up and down, wheels spin, and a motor vehicle is propelled along the street.

10.1.2. Total energy vs. electricity

A word of caution is in order at this point. You should note the potentially confusing distinction between **total energy**, which includes electricity, heating, and transportation, and **electricity** alone. Be on the lookout for the two terms as we proceed, and remember that they do not refer to the same thing.

10.2. CONSEQUENCES OF ENERGY USAGE

Energy conversions to meet demands for electricity, temperature control, or transportation initiate processes that lead to two consequences, one of which is highly likely, and the other of which is unavoidable.

10.2.1. Environmental degradation

Although the nature and severity of it varies, this consequence is virtually certain. For example, in addition to the large amounts of useable energy they release, both coal combustion and nuclear fission produce harmful products. Even the use of energy sources generally regarded as clean, such as solar and wind, has some environmental costs. Details about these and other adverse consequences linked to power generation will be presented in Parts II and III of this chapter.

10.2.2. Loss of usable energy

Substantial amounts of energy must be invested—never to be used directly by us—in the interest of power generation. Two phenomena are responsible for this second, inevitable, consequence.

Technology used to obtain energy needs power to operate

Speaking rather informally, it takes energy to get energy. Keep in mind that the machinery necessary to acquire, process, and transport our various energy sources must be powered somehow. For example, drills, pumps, and marine tankers, just some of the technology used in the exploitation of fossil fuels, generally rely on the combustion of fossil fuels themselves. Although our goal is to capture as much *net* useable energy as possible, losses are unavoidable. Importantly, the worth of a particular resource is often judged by the relationship of the amount of energy needed to obtain it to that provided by it. No fuel will produce more energy than is invested in it, but efficiencies do vary substantially.

Energy conversions are inefficient

Recall from our consideration of the **second law of thermodynamics** in Chapter 4 that every energy transformation leads to a reduction in the amount of usable energy in the universe. So, when the **potential energy** stored in a fuel such as gasoline is converted to the **kinetic energy** of a moving car, some amount of low grade heat is always released (review Box 4.6). Powering technology, therefore, will necessarily reduce the amount of high-grade energy available in the future. The practical relevance of this reduction varies considerably among energy sources as some are so abundant that usage has a trivial and irrelevant effect on their availability whereas others are appreciably diminished when they are used to do work. We will hear more about the importance of this distinction below.

10.3. Renewability: an Introduction

We group energy sources according to the effect current usage has on future supply (as noted just above), or on what is commonly known as their **renewability**.

Non-renewable sources. The supply of these sources is substantially reduced when they are used to do work. Once again, systems analysis is useful here: since the rate of usage, or output, far exceeds that of input, or formation, the quantity of non-renewables in Earth's systems decreases (Figure 10.3).

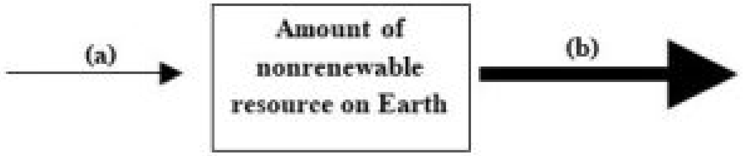


***Figure* 10.3.** *The rate of formation (arrow a) is lower than that of use (arrow b), so the amount of a non-renewable resource declines, or is depleted (review Chapter 2). Kelsey,* CC BY-NC-SA.

Importantly, although it is true that materials such as coal and petroleum *are* replenished by natural processes, their rate of replacement is far too slow to keep up with demand. We will focus a great deal of attention on these and other non-renewable energy sources in Part II of this chapter.

Renewable sources. A commonly used way to assess renewability is to determine the length of time required to replenish a resource after the high-grade energy in it has been degraded. If it can be replaced in a period that is short enough to be useful and relevant for humans, then it is counted in this group. Note that renewability does *not* imply the energy is somehow recycled or reused (in what would be a violation of the laws of thermodynamics), rather, high-grade energy degraded to low-grade energy is replaced by an additional quantity of high-grade energy. It is also appropriate to classify a source as renewable if its supply is sufficiently large that humans could not plausibly deplete it. Consider the sun, which easily meets this second standard of supply vs. usage. It will continue to release energy so far into the future, for several billion more years, that it might as well be infinite and inexhaustible as far as we are concerned. Solar, wind, and hydrothermal energy are among the renewables we will study in some detail in Part III.

Part II: NON-RENEWABLE ENERGY SOURCES

10.4. NON-RENEWABLES DOMINATE

These have provided the bulk of our energy since the beginning of the Industrial Revolution of the 1800s. In this section we will explore advantages and disadvantages of each, paying close attention to three considerations: how it is formed and / or used to provide power, the environmental costs associated with it, and how much demand it can meet.

10.4.1. Fossil fuels

These have well-known economic and cultural importance. They are energy rich, portable, and versatile enough to provide power in different ways. Note they are naturally occurring, having been derived from organisms that died millions of years ago (Chapter 3). You should also realize that their ultimate source of energy was the sun because the relevant organisms lived in **ecosystems** powered by **photoautotrophs** (Chapter 5). The fossil fuel group includes crude oil, natural gas, and coal.

Crude oil and natural gas

These first two are fluids with similar chemistry but different physical properties. Each is composed primarily of carbon and hydrogen atoms (i.e., they are called **hydrocarbons**—see Figure 10.4), however, crude oil is a dense, thick, tarry liquid, whereas natural gas is a light, gaseous substance.

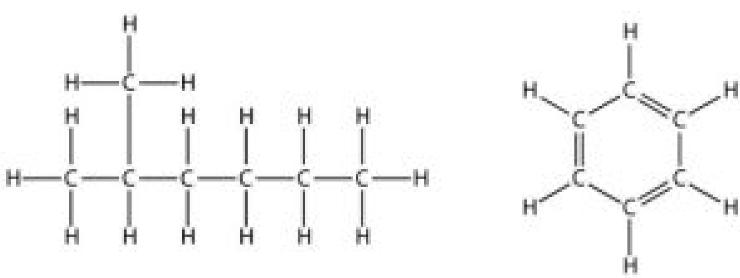


Figure 10.4. *Hydrogen and carbon atoms can be connected in many ways, including chains and rings, to form different kinds of hydrocarbons. Pictured here are just two possibilities. Jessie A. Key, CC BY-NC-SA.*

The term **petroleum** is often used as a synonym for oil, although some classify natural gas as a petroleum product as well. They are paired in our discussion here, with "petroleum" used to refer to both, because they generally form under the same extremely specialized and unusual conditions.

Formation. The story of oil and gas begins with tiny **plankton**—primitive plant- and animal-like organisms—living in surface waters of ancient oceans. After they died, decomposition converted all but a small fraction of their remains to inorganic products like **carbon dioxide**. The **organic molecules** that persisted sank and mixed with **inorganic materials** (e.g., weathering products like sediments) in low-oxygen environments on the seafloor. Note that this lack of O_2 gas limited **aerobic** organisms that would otherwise quickly break down the accumulating cells. Burial under additional sediments pushed the materials deeper into the **lithosphere**, subjecting them to elevated temperatures (as they moved toward Earth's hot interior) and pressures (squeezed together by overlying layers of sediments). Metabolism by **anaerobic** bacteria (Chapter 1) likely acted upon them as well. At a few hundred meters depth, the biological materials became a waxy intermediate product called **kerogen**. Beginning around 2 km, the kerogen underwent conversions to crude oil. Increasing temperatures and pressures with continued burial yielded **natural gas**. Realize that at this stage the materials were not concentrated in some sort of underground pools just waiting for humans to find them, instead, they were dispersed in the pores of sedimentary rocks (Chapter 3) located several kilometers below the surface. It took millions of additional years for the fuels to move upward from what are known as their **source rocks** to shallower locations where they collected in **reservoir rocks**. Why did they move against gravity? Most petroleum is less dense than solid rocks and will rise through cracks and pores, assuming such openings are present, toward Earth's surface. In fact, oil and gas continue to move up until and unless they encounter a barrier, or **oil trap**, an impermeable geologic structure through which fluids cannot readily pass. For example, a fault (Chapter 7) can trap large amounts of petroleum and lead to the formation of an economically important oil field. When all was said in done, on the order of 200 million years elapsed between the death of those ancient plankton and the appearance of useable quantities of petroleum. Figure 10.5 presents an overview of the formation of oil and natural gas reservoirs, including an example trap, and Box 10.1. provides a bit more information about the biological origins of oil.

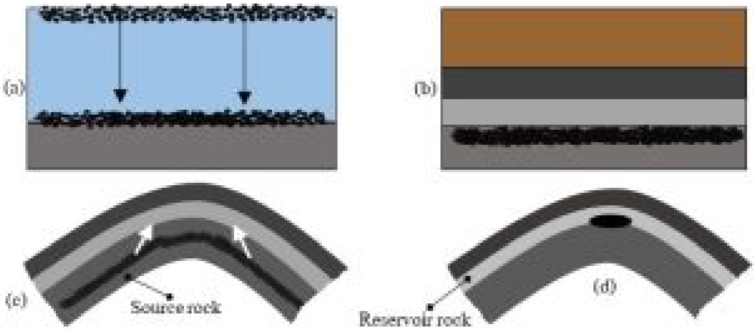


***Figure* 10.5.** *Formation of petroleum reservoirs. Remains of organisms settle to the ocean floor (a). The remains from the first step get mixed with inorganic materials, modified by reactions, and buried (b). Sediments are solidified into rocks and folded into a structure such as the one shown; the organic material undergoes more transformations and then migrates upward from source rock until it reaches an impermeable boundary (c). After millions of years, the petroleum dispersed in the lower layer has been concentrated in reservoir rock, under an oil trap (d). If discovered, the trapped oil may then be pumped to the surface. In cross section, not to scale. Kelsey,* CC BY-NC-SA.

Box 10.1. Not dinosaur fossils!

Oil is often mistakenly thought to be the remains of dinosaurs, but alas, the real story of its origins is quite different. Most of the oil we see these days got its start hundreds of millions of years ago, before those famous extinct animals even roamed the Earth. Chemical analysis of petroleum in the lithosphere confirms that it is made up of the remains of tiny plankton that once lived in ancient oceans. Perhaps this is a less glamorous story than we would like, but it is one supported by data.

Note that in some cases oil does not encounter an effective trap during its upward journey but slowly bubbles out, or **seeps**, into ocean waters from the seafloor. During millions of years, certain marine bacteria have adapted to eat this oil that enters their environments! Scientists take advantage of these microorganisms to clean up spills, as described in more detail in Box 10.2 (upcoming).

Environmental costs. Each of the many steps required to bring oil and gas from their natural reservoirs to the places they are used by humans causes degradation of natural environments.

Exploration

Oil and gas are found in places subjected to the conditions that led to the formation and accumulation of these fuels hundreds of millions of years ago. The challenge, then, is to find those reservoirs from which useable quantities can be extracted (next step, below). Several tools are used to explore. First, a study of the geology of an area is essential. Since we broadly understand how oil is produced, we can examine data to narrow down locations that are likely candidates. For example, places such as present-day Saudi Arabia and Texas were once environments conducive to petroleum formation. Second, satellite images and maps can be used to search for surface features associated with reservoirs. Finally, specialized instruments allow a direct study of rock layers below Earth's surface in areas thought to harbor oil: the behavior of artificially generated vibrations sent underground can indicate whether the drilling of exploratory wells is warranted. If oil is discovered, an above-ground ecosystem can be profoundly altered as land is cleared and roads, buildings, and more wells are constructed. When reservoirs are beneath the ocean, analogous changes will occur as oil platforms are built.

Extraction

Removal of petroleum from reservoirs—known as petroleum recovery or extraction—is accomplished through several different methods, and each can cause appreciable environmental degradation.

Conventional oil drilling. This is currently the most widely used technique. Here, large metal drills are driven into the Earth, producing a vertical hole that leads to underground reservoirs. The opening is then subsequently lined with pipe and other materials. Such wells can be built in marine or terrestrial systems, using oil rigs or oil platforms, respectively (Figures 10.6.a and b).

Figure 10.6.a. *Oil recovery via a marine oil platform.*

Figure 10.6.b. *Oil recovery via a terrestrial oil rig.*

The process is complicated for many reasons, including the fact that the methods used to obtain oil from any given well generally must change with time. When a drill first penetrates a new reservoir, a tremendous amount of pressure can be released, sending oil gushing up the pipe toward the surface. During this early stage, oil is obtained relatively easily because it is not tightly bound in reservoir rocks. However, only about 10% of the fluid in a typical well can be recovered through simple pumping. Large amounts of energy (and money) are required to remove oil that is locked up in pores and other places. In the end, it is not unusual for over half of the oil in a reservoir to be left behind because it is too costly to obtain it.

Extraction from either land- or sea-based wells is linked to several adverse environmental consequences. First, oil can spill into natural systems due to damaged equipment or errors. These events can be particularly problematic when pipes or pumps lie deep below the surface of the ocean. The case of the Deepwater Horizon accident of 2010, which occurred under about 1500 meters (~1 mile) of water in the Gulf of Mexico, provides an extreme example of the potential dangers of oil drilling. An explosion killed eleven people, sank the platform, and ultimately generated a large leak on the seafloor—some 15 million liters (~4 million gallons) of oil escaped into the water during the nearly three months that passed before the gushing well was capped[1]. Spills of this kind can cause substantial amounts of damage to environmental and human systems (see Box 10.2).

Box 10.2. Marine oil spills: consequences and clean up

1. Adverse consequences.

(a) Petroleum can kill fish, mammals, seabirds, and

1. National Commission on the BP Deepwater Horizon oil spill and offshore drilling, 2011, Deep Water: the gulf oil disaster and the future of offshore drilling. Public domain. https://www.gpo.gov/fdsys/pkg/GPO-OILCOMMISSION/pdf/GPO-OILCOMMISSION.pdf

other organisms living in and near the water. It is poisonous and can also reduce the insulating capacity of feathers, causing coated waterfowl to freeze to death (Figure 10.7, left)

(b) Spills can create public-relations nightmares for responsible parties, as when they leave beaches coated with oil (Figure 10.7, right).

Figure 10.7. *Oil coats both animals left (a) and beaches right (b). Marine Photobank, CC BY (a); Exxon Valdez Oil Spill Trustee Council, Public Domain (b).*

(c) Oil spills cause human suffering. Substantial economic losses can affect individuals and communities that rely on ocean resources for their livelihoods. The spill associated with the Deepwater Horizon explosion (described in the main text), damaged fisheries, including those that harvest oysters.

2. Can they be cleaned up? As noted in the main text, petroleum hydrocarbons can be **decomposed** by marine

microorganisms adapted to use them as nutrient sources. Some fraction of spilled oil is, therefore, broken down by natural processes. However, since the amount of oil released at one time by a large accident can overwhelm ocean systems, human intervention is generally needed to reduce adverse consequences. Time is often of the essence, because we want to keep spilled oil from spreading over the water's surface and moving onto beaches where its removal is even more difficult. Several **remediation** (i.e., clean-up) strategies may be used.

(a) Containment. As soon as is possible, efforts are made to slow the spread of oil from its source (e.g., a damaged well or tanker). Floating barriers known as booms (Figure 10.8) are often set up at the periphery of a spill to reduce damage to both water and beach ecosystems.

Figure 10.8. *A floating boom deployed to contain a marine oil spill. US Coast Guard, Public Domain.*

Although containment is clearly just a short-term solution, it can buy clean-up crews some time to deploy other strategies to eliminate oil.

(b) Physical removal. In effect, floating oil is scooped off the surface of polluted water using devices called skimmers. Skimming is a slow process but can be effective in reducing the amount of oil that ultimately moves onto beaches. High-pressure washers have been used to blast oil from rocks and sand, but their hot water tends to kill many of the bacteria that would eventually be able to remove much of the oil if allowed time to do so.

(c) Breakdown. Naturally occurring microorganisms with the appropriate adaptations can be harnessed to clean up oil. The extent to which this strategy is effective depends on many difficult-to-control variables. If the ocean is very cold (as in the arctic), nutrient poor, or otherwise hostile, bacteria able to break down oil in idealized laboratory experiments might not do so in a real-world situation. Scientists continue to study ways to enhance the survival and activity of microorganisms to help remediate ocean oil spills. We will see more about the use of microorganisms to breakdown environmental pollutants in Chapter 15.

Second, the drills, pumps, and other machinery required to recover oil release air pollutants because they are powered by fossil fuel combustion. Third, existing ecosystems and landforms are disrupted as space is made for oil rigs, roads, pipelines, and the rest of the infrastructure associated with oil drilling. Land subsidence can even occur in areas that lie above wells, because, in some cases, critical support for Earth's surface is lost when large amounts of oil are extracted (Figure 10.9). Buildings can be damaged by shifting ground, and inundation by water can cause widespread flooding.

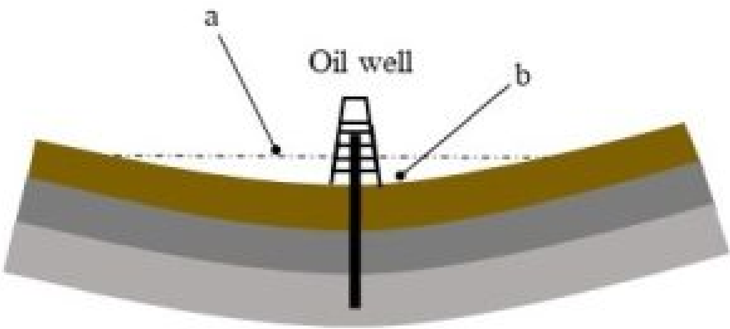


Figure 10.9. *Land can subside as a result of oil extraction. Height 'a' is the elevation of the land before the well was constructed, and 'b' is the elevation of the land after the well started pumping. In cross section, not to scale. Kelsey,* CC BY-NC-SA.

Conventional natural gas drilling. Oil and natural gas are often found near each other and are recovered from porous rocks through similar methods. They are also associated with many of the same environmental concerns: air pollution, ecosystem disruption, and changes to landscapes. Since it is gaseous, though, accidental releases of natural gas cause air pollution to a far greater extent than do oil spills. As we will see in detail in Chapter 14, methane (an important component of natural gas) plays an important role in raising Earth's average temperature and changing its climate.

Unconventional methods. For various reasons, much of Earth's petroleum does not lend itself to recovery through conventional wells. In such cases, one of several alternative extraction techniques can be employed. Although these methods have the potential to greatly increase the amount of oil available to us, they generally are more expensive than conventional ones and present their own environmental challenges. Commonly used unconventional methods include mining from oil sands or oil shales and hydraulic fracturing. We will look only at the last technique on this list.

Hydraulic fracturing, commonly referred to as **fracking**, is used to free up so-called **tight** petroleum that cannot be extracted by conventional drilling methods. In these deposits, oil and gas—in some cases, enormous quantities of them—are trapped in individual, isolated pores within shale (a type of **sedimentary rock**). Importantly, these openings are not connected to each other, meaning that fluids do not flow through the rocks efficiently enough to justify the expense associated with conventional pumping (Figure 10.10 distinguishes porosity from permeability, two properties that influence storage and extraction of oil and gas within rocks).

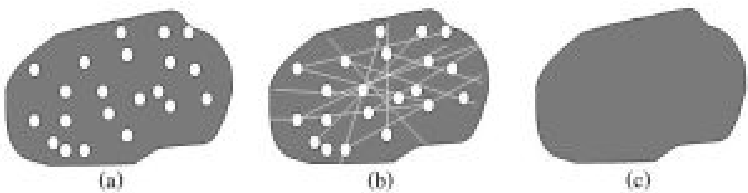


Figure 10.10. *Porosity vs. permeability. a: isolated pores in a rock with low permeability. Petroleum stored in these pores would not be recoverable with conventional methods and would require fracking to allow fluid to flow through the rock; b: pores are connected in a rock with both high porosity and permeability, and petroleum would be relatively easy to recover; c: a rock with low porosity and low permeability would not store petroleum. In cross section, not to scale. Kelsey, CC BY-NC-SA.*

Fracking was developed to increase permeability in such deposits. In short, fluids consisting of water, chemical compounds, and sand-like materials are pumped into reservoir rocks at such high pressure that new cracks and channels form. After the rocks have been fractured, the contents of those many pores flow toward each other and become concentrated in relatively small areas; rigs can then pump the petroleum to the surface. In recent years, extraction efficiency has been further improved through **horizontal drilling**. Here, a vertical well pipe is curved so it runs in a direction roughly parallel to the Earth's surface through reservoir rocks (Figure 10.11).

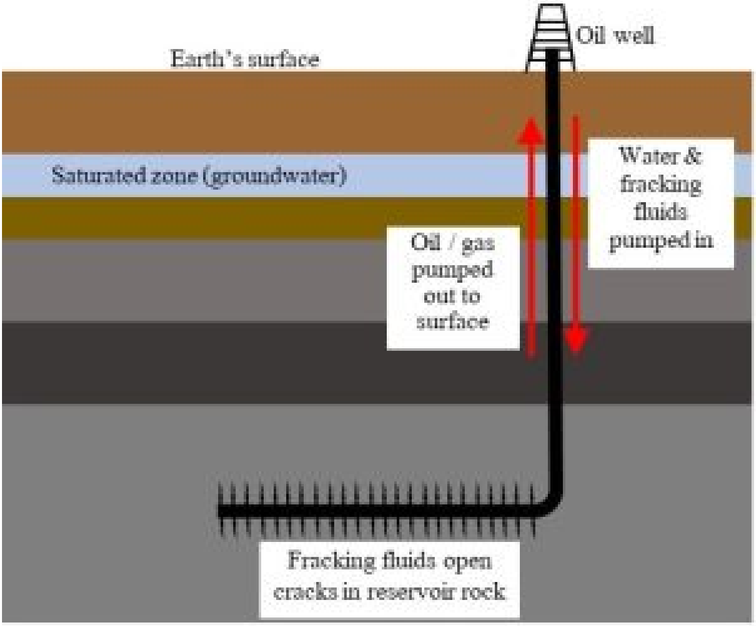


Figure 10.11. *Hydraulic fracturing. See main text for details. In cross section, not to scale. Kelsey, CC BY-NC-SA.*

Several environmental concerns unique to fracking have made it a controversial technique. First, large volumes of water are used to break up rocks and increase permeability. In places like Pennsylvania, where water is relatively abundant, high usage is generally not viewed as limiting or even much of a concern. However, fracking in Texas and similarly dry areas could further strain already limited water supplies. Second, the fluids injected into rocks to increase permeability could spill onto the surface or leak from broken pipes and move into groundwater (Chapter 4 and Chapter 11), substantially degrading an important source of drinking water for many people. Finally, there have been suggestions that methane gas released by fracking can make its way into the water supplies of nearby houses. A quick internet search can even reveal sensationalized video images of homeowners igniting their sink faucets! Many questions, both scientific and legal, have arisen about the possible link between fracking and what appears to be combustible water, but definitive answers have yet to be found.

Transport

Transport is needed at two different stages of the petroleum cycle: to move materials pumped from wells to refineries (below) for processing and conversion into various fuels, and to move the products of refining to the places at which they will be used. It is important to recognize that, when initially extracted from the Earth, petroleum is generally not in a form that is directly useable by humans. For example, **crude oil** is the material obtained through conventional oil rigs. This viscous, tarry substance can rarely function as an energy source for electricity generation, transportation, or heating. Crude must first be transported from wells to facilities at which it is converted into separate fuels (more on these conversions is presented below). For both crude oil and refined fuels, one of four transport methods are common.

Pipelines. Essentially large metal tubes that can extend for thousands of kilometers across and under Earth's surface, pipelines are commonly used to transport petroleum (Figure 10.12).

Figure 10.12. *Pipelines such as this one in Alaska, U.S.A., can carry petroleum from wells to refineries and refined products to places they will be used.* Luca Galuzzi, CC BY-SA.

Two important environmental concerns arise from pipelines. First, land is cleared and ecosystems disrupted when they are constructed. Among other problems, movement of migratory land-based organisms can be interrupted by the presence of these large barriers (recall **habitat fragmentation** in Chapter 6). Second, ruptures in oil pipelines can cause the same kinds of consequences seen above in the discussion of oil extraction, namely pollution of water and soil.

Pipelines are the dominant means by which natural gas is transported from wells to processing plants and users (as we will see, other methods also play important roles in the transport of oil). Large-diameter tubes move it long distances, and then smaller ones carry it into homes and businesses. The need for pipelines has placed limitations on how much gas can be moved among continents (i.e., they are not conducive to transporting material across oceans), meaning that it cannot be as easily distributed from producers and consumers as can oil. However, when it is cooled to very low temperatures, it becomes a liquid that can be shipped via vehicles (like trucks, trains, and marine tankers). Increasing interest in this liquefied natural gas (LPG) could change global markets, a point we will briefly revisit again later.

Rail cars and on-road trucks. These means of transport carry less oil than do pipelines and have considerably less capacity than do marine tankers (next), but they still are used in many cases. Trucks are valued for their flexibility—they obviously can deliver directly to many more places than can pipelines. The risks they pose are like those we have seen already: air pollution from fossil fuel combustion and human and environmental damage due to accidental releases.

Marine oil tankers. When over-land methods are impractical, large ships are used to move oil across oceans (Figure 10.13).

Figure 10.13. *Marine oil tanker docked at a platform. US Navy, Public Domain.*

Just one of these tankers can hold up to 2 million barrels[2] of liquid (note that one barrel contains about 169 liters, 42 gallons, of petroleum), so, if all goes as planned, this means of transport is relatively cost efficient. Unfortunately, though, tankers have been known to crack and release their cargo into seawater. Most of the time, such accidents involve small quantities of petroleum, but since 1969 there have been 44 events releasing 10,000 or more barrels each (about 1.6 million liters) worldwide[3]. The consequences of spills depend on the location in which the oil is released, but damage to both natural coastal ecosystems and human welfare is common (review Box 10.2 for more).

Refining

2. U.S. Energy Information Administration, 2022

3. States National Oceanographic and Aeronautics Administration, Office of Response and Restoration, 2017. http://response.restoration.noaa.gov/oil-and-chemical-spills/oil-spills/largest-oil-spills-affecting-us-waters-1969.html

Crude oil extracted from the lithosphere is taken to large facilities known as **refineries** where it is treated in a multi-step process to produce different fuels. Refining starts when crude is heated to yield several separate fractions (i.e., individual fuels). Relatively light gasoline, kerosene, and some diesel fuels are among the outputs collected after this stage. Next, the heavier fractions left from step one are subjected to high pressure, further heating, and various other processing to generate additional gasoline and other valued commodities. Finally, treatment to remove impurities and otherwise prepare the refined products for sale and use is carried out as needed.

In the end, for every barrel of crude that enters an oil refinery, about 20, 12, and 4 gallons of gasoline, diesel fuels, and jet fuel, respectively, come out (the remainder becomes asphalt and other petroleum products)[4]. Figure 10.14 summarizes refining and its products.

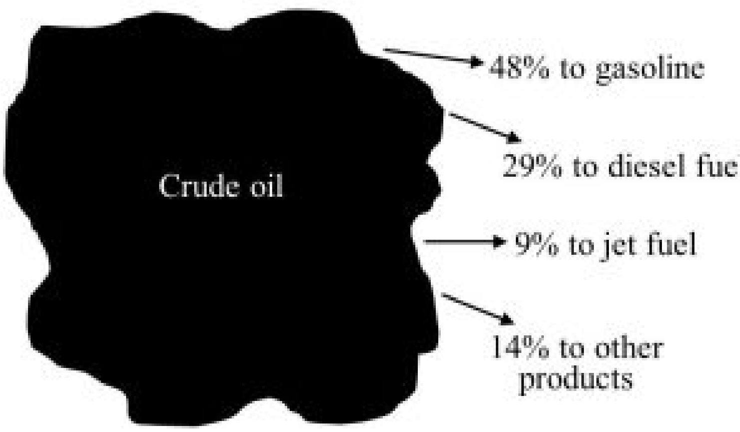


Figure 10.14. *Refining yields many useable petroleum products from crude oil. Kelsey*, CC BY-NC-SA.

4. U.S. Energy Information Administration, 2022

Refining causes some important adverse environmental effects. Many air pollutants are routinely emitted, including heavy metals, dust particles, carbon dioxide, sulfur dioxide, nitrogen oxides, and others (we will see more about sources and effects of air pollution in Chapter 14). Hazardous waste regularly enters waterways adjacent to these facilities as well. Moreover, refining is very energy intensive, and a lot of fossil fuels are burned to carry out the various stages of it. Accidental releases from refineries can also damage natural systems, and air and water pollution are ongoing concerns (Figure 10.15). Issues related to transport are applicable here as well because the finished products must be moved to their points of sale and use.

Figure 10.15. *Refineries like this one in Louisiana, U.S.A. are visible sources of air pollution. Wclarke, CC BY-SA.*

You should be aware that natural gas extracted from wells does not undergo the same kind of refining as crude, although it still must be processed to some extent before it is usable. Various hydrocarbons, water, and other impurities are removed at both the site of a well and remote treatment facilities to yield a product that is nearly 100% methane.

Use

The outputs from refineries are transported to places like gasoline stations and power plants. A great deal of energy is released when fossil fuels are then burned and converted to products—after all, that is why these materials are so prized—but, their use generates several air pollutants as well.

Carbon dioxide. Upon combustion, nearly all the organic carbon in petroleum ends up bonded to oxygen atoms to form **carbon dioxide** gas. On the order of 8,900 grams, or 19 pounds, of **CO_2** is released for each 3.8 liters (1 gallon) of gasoline burned (and about 511 liters—135 billion gallons—were burned in the United States during 2020)[5]. Note that, when compared on the basis of equivalent amounts of energy yielded when combusted, diesel fuel releases about 5% more and natural gas releases about 30% less CO_2 than does typical gasoline[6]. This gaseous product presents environmental concerns not because it is directly toxic to humans (although in very high concentrations it can push oxygen away and cause asphyxiation), but because it is one of several important greenhouse gases that contribute to global warming and climate change (Chapter 14).

Other carbon products. The remaining C atoms get incorporated into carbon monoxide molecules, methane, and various hydrocarbons. The first on that list of three, CO, is poisonous to humans and other organisms and can undergo chemical transformation to CO_2. Methane is of concern because, like CO_2, it plays an important role in the process of global warming that we will explore in Chapter 14. Recall that "hydrocarbon" is a general chemical term that refers to any compound made up of C and H atoms. Small amounts of many different hydrocarbons are produced when fossil fuels are combusted, and they contribute to a phenomenon known as smog (also coming up in Chapter 14).

Nitrogen and sulfur gases. Nitrogen oxides, often represented by "NOx" to suggest that the number of nitrogen and oxygen atoms

5. U.S. Energy Information Administration, 2022

6. U.S. Energy Information Administration, 2022

can differ among compounds, and sulfur dioxide, SO_2, also can exit the tailpipes of motor vehicles (combustion of diesel fuel is a particularly important source of sulfur gases). Many of these gases may be transformed into acid precipitation and are important components of smog (again, more about both consequences will be presented in Chapter 14).

Figure 10.16. *Particulates from diesel combustion. US EPA, Public Domain.*

Particulates. These dust-like physical products are often visible as dark-colored clouds in the exhaust of diesel vehicles (Figure 10.16). Concerns about particulate matter include respiratory and other health effects among exposed humans, reduced visibility in affected air, and pollution of water and soils into which the solids ultimately settle.

Mercury. Since petroleum often contains mercury, combustion releases this toxic metal into the atmosphere. Through various mechanisms it is eventually deposited into soil and water where it can damage individuals and ecosystems. In a related note, lead was once added to gasoline and emitted like mercury when fossil fuels were burned, but, starting in the 1970s, it was phased out of fuels used by passenger vehicles in the United States.

Meeting present and future demand. It would not be an exaggeration to say the industrialized world is dependent on petroleum. In 2021, oil and natural gas met approximately 36% and 32%, respectively, of the total demand for energy in the United

States (worldwide, the numbers are similar)[7]. Since humans consume them faster than they can be naturally replaced, these levels of use are not sustainable, and, if nothing changes, the planet will be effectively emptied of oil and natural gas at some point in the foreseeable future. A universally agreed upon answer to the question "just when will we run out of petroleum?" has proven to be elusive, however, with estimates ranging from as low as a few decades to as high as a few centuries. These disparities arise from several sources, including differences in assumptions about the size of conventional and unconventional reservoirs, future demand for fuels, and the future costs and availability of alternative energy sources (more about these later, in Part III). Here, we briefly explore some important facts about production and use of petroleum and then return to the question of how long our supply might last.

Calculating supply: reserves vs. resources

Not all petroleum on Earth is equally accessible. A portion of it can be readily extracted using conventional methods, but much of it is not so easily obtained. In the latter case, one of the unconventional methods noted above may be tried, however, some material is so tightly held within the lithosphere that it cannot be removed with currently available technology. Furthermore, most scientists assume some modest quantity of petroleum is yet to be discovered, and the search for new oil fields continues. Since exploration is generally informed by knowledge of geology and predictions about probable locations of untapped reservoirs, estimates about how much more we can expect to find are often made. Combining considerations about relative accessibility with informed guesses about future discoveries, we have established different ways to quantify and report supply. Quantities of petroleum (and for that matter, any natural material of interest) that have been located and can be extracted in an economically

7. U.S. Energy Information Administration, 2022

viable way are classified as **proven reserves**. On the other hand, the term **resources** refers to known reserves *plus* material that is more difficult to recover for one or more reasons. For example, oil that has been discovered but is stored such that it is either currently impossible (due to technological limitations) or too expensive (i.e., it costs more money to produce than it will return when sold) to extract would be counted here. Oil that is thought to be present in an area without confirmation that it really exists is also included with resources. The word "resources" clearly represents a far more optimistic appraisal of supply than does "reserves", so we should interpret and use the two terms with great care.

Petroleum distribution is heterogeneous

Oil. Due to past differences in environmental conditions around Earth, the current availability of oil and natural gas is unequal. It would be fair to say, in fact, that petroleum is concentrated in a relatively small number of nations. An estimated 80% of conventional oil reservoirs (i.e., proven reserves), about 1.2 trillion barrels, are controlled by the Organization of the Petroleum Exporting Countries (OPEC)[8]. This group is made up of about fourteen nations (the number can change from time to time) from the Persian Gulf region, northern Africa, Latin America, and Asia. Saudi Arabia and Venezuela, two members, together hold over half of OPEC's oil. On the other hand, most of the world's unconventional supply is in the United States, Canada, Venezuela, and Russia. Estimates vary, but another 2 trillion barrels (possibly much more, according to some rather sanguine calculations) could be held in these difficult-to-exploit reservoirs. Recent studies even indicate that, when all resources are included, more oil is present in the United States than in any other country[9].

Conventional reservoirs dominated world supply for most of the

8. opec.org

9. U.S. Energy Information Administration, 2022

past century, and combustion of them has eliminated nearly 1 trillion barrels of Earth's supply (you should know there is some debate about the size of that figure). However, with the development of new technology, oil from unconventional reservoirs, even some that has historically been inaccessible, is now extractable and relevant. In the U.S. alone, the percentage of crude oil production from hydraulic fracturing and related methods rose from essentially 0 in the year 2005 to 48 in 2017, and the amount of oil coming from those unconventional sources is predicted to increase by a factor of two or three in the next thirty years[10].

Natural gas. As with oil, reservoirs of this important energy source are unequally distributed: about two thirds of the 200 trillion cubic meters (tcm) of proven natural gas reserves are controlled by just five countries: Russia, Iran, Qatar, the United States, and Saudi Arabia[11]. Relatively new technology allowing extraction of **tight gas** has greatly increased the estimated size of Earth's natural gas resources by about three fold (i.e., including all recoverable reserves and resources). The continued high cost of exploiting these unconventional reservoirs, though, is an important factor limiting long-term supply.

Demand

Oil. Demand for oil has risen during the past several decades and is likely to continue that trajectory for the next several. Humans use around 100 million barrels per day (37 billion per year) and that number is projected to increase in coming years, especially as usage in China, India, and other nations in the **developing world** goes up[12].

Natural gas. Several factors, among them worries about crude

10. U.S. Energy Information Administration, 2022

11. U.S. Energy Information Administration, 2022

12. U.S. Energy Information Administration, 2022

oil supply, reduced air pollution when it is burned instead of oil, and the emergence of new approaches to transport that allow intercontinental distribution of natural gas to regions that lack significant local suppliers (using LPG), have stimulated a doubling in global consumption since 1980[13] (current demand is about 4 tcm and, despite current short-term declines in demand, likely to rise in coming years[14]). Recent discoveries of vast reservoirs of tight gas, including those in the Marcellus Shale deposits of New York, Pennsylvania, West Virginia and a few other eastern U.S. states, along with the Barnett in Texas and elsewhere, have also stoked usage. This fuel has become increasingly important in electricity generation, heating, and transportation, and that trend is likely to continue. In just one example, power plants that rely on coal as an energy source are steadily being replaced by those that use natural gas because the former emit more air pollution than the latter (will we see more about coal below).

How long can demand be met?

Oil. We now attempt to address the seemingly straight-forward, yet difficult-to-answer, question we posed at the beginning of this section. How much oil remains on Earth is hotly contested and depends on both who you ask and how they measure. The most conservative, if widely accepted, assessments count only the so-called proven reserves of oil, or about 1.6 trillion barrels[15]. Now, if we divide the amount of oil left by current demand (1.6 trillion or 1,600 billion barrels / 37 billion barrels per year) we can conclude it will run out in **43 years**. There are two caveats to keep in mind here: 37 billion barrels a year surely underestimates future demand,

13. U.S. Energy Information Administration, 2022
14. Gas Market Report, Q3-2022. International Energy Agency. www.iea.org/reports/gas-market-report-q3-2022
15. for example, see opec.org

and total supply is likely higher than 1.6 trillion barrels (remember, some additional untapped reserves probably will be discovered and we can count unconventional reservoirs). Thus, this number may represent the low end of future availability. Less conservative estimates add oil believed to be tied up in difficult-to-access reservoirs, as well as undiscovered crude, that is, include all oil classified as **proved, probable,** and **possible** (three categories ranging from most certain and accessible to least). Taking those less-reliable pools into account, the total amount of oil left could range from 2 – 4 trillion barrels, or last up to about **130 years**. Finally, some extreme estimates envision 10 trillion additional barrels of unconventional oil (yielding a theoretical sum of 11.6 trillion), meaning that present-day demand could be met for well over **300 more years**. Consult Box 10.3 for more about predicting the future of oil and the concept known as **peak oil**.

Box 10.3. Peak oil?

In 1956, an American geologist named M. King Hubbert made an evidence-based prediction about future supply and demand for oil. He observed that oil will progress through stages of increasing discovery, production, and use that inevitably leads to depletion and ever-declining availability. His model, represented in what is now known as a **Hubbert Curve** (Figure 10.17), is based on several assumptions. First, in the early years of the industry, discoveries of petroleum reservoirs outpace the use of oil. Then, demand increases as supply increases until production reaches it maximum, or **peak level**. After the peak, supply cannot keep up with demand and prices increase irrevocably.

Eventually, the amount of oil on Earth diminishes to levels so low as to be irrelevant.

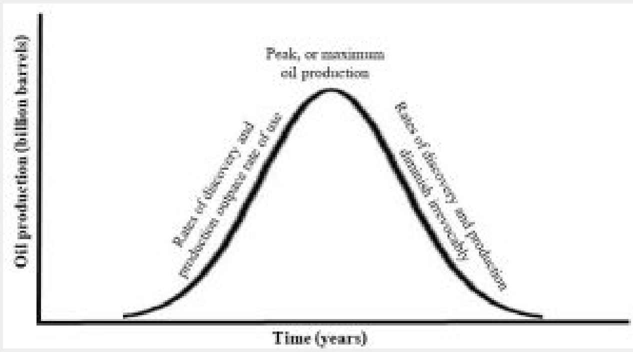


Figure 10.17. *An idealized model of the life of a finite resource (like oil). Absolute dates are not known so are left vague. Kelsey,*

Hubbert was partially correct when he predicted that petroleum would peak in the 1970s in that conventional oil production in the 48 contiguous U.S. states did reach its maximum then. But he did not account for oil from places such as Alaska, OPEC, and Russia, nor did he anticipate unconventional oil. Many have discounted his model completely because it failed to predict the correct year of peak oil. Others take it more seriously, noting how it reasonably describes the fate of oil and any finite resource. In other words, although the absolute timing of his prediction was clearly off, the overall trend of Hubbert's curve fits our basic understanding of what will happen if we continue to consume oil far faster than it is replaced by natural processes.

The story gets quite a bit more complicated when we consider how price affects supply. For instance, increases in the cost of petroleum—vividly manifested by soaring prices at gasoline stations—during the 1970s led to economic turmoil for nations that imported much of their oil. By and large, short-term demand during these periods of rapid change remained constant; people continued to buy more expensive petroleum products because they simply needed them to drive. In the long term, though, cars got smaller and fuel efficiency standards rose during the 1980s (i.e., less fuel was used per kilometer or mile travelled). Expensive unconventional methods to obtain petroleum as well as renewable sources of energy also began to attract attention and investment dollars. Unsurprisingly, oil supply increased, substantially exceeding demand at times (during periods known as **oil gluts**). Resulting declines in petroleum prices then cut into the profits of oil producers and necessitated the closing of many costly wells. Economic, political, and military conflicts during recent years and decades have meant that price and supply remain volatile. These days, OPEC—the group that controls most of Earth's conventional oil—is pitted in a complex battle against producers in Canada, the United States, Russia, and other countries that rely on unconventional methods. The former would like to ensure elevated and stable prices whereas the latter rely on relatively high prices to enable their continued existence even as they drive down prices by increasing supply. Put bluntly, some in OPEC have sought to flood the market with oil to lower prices and put competitors out of business. Interestingly, though, expensive unconventional wells have continued to produce petroleum when prices have decreased, in part because the industry desires to remain relevant in the long run. Of course, all types of oil producers are threatened by renewable sources of energy we will describe below.

Two final points about oil deserve some attention. First, in projecting the remaining life of oil, we should assume that its price will go up as it becomes increasingly scarce. In other words, oil will probably be around far longer than 43 – 130 years (currently, the most plausible estimates) because it will be too expensive for most people to use in any relevant quantity. Of course, in such a case it would no longer provide as much energy as it currently does, presumably due to the rising viability of other non-petroleum options, and remain on Earth indefinitely. Second and related, before we embrace the notion that 2, 4, or even 12 trillion barrels are left, a thorough examination of the consequences of extracting, refining, transporting, and burning all the remaining oil we can ultimately access is warranted (we will return to this point in Chapter 14). Informed by the data and scientific ideas presented here, we could undertake serious discussions about the future role of petroleum in meeting human demands for energy. Put into other words, although we might be able to eventually develop ways to recover every drop of oil remaining in the lithosphere, we need to properly weigh all the risks before we set out to do so.

Natural gas. The long-term viability of this fuel is difficult to predict because it depends on technology and costs of natural gas and other energy sources. With global demand at roughly 4 tcm / year and the amount in proven reserves, 200 tcm, we could conclude that the answer is **50 years**[16]. The story is more nuanced, however, than this simple calculation suggests. Importantly, more than 200 tcm is likely present—possibly by a great deal—although we need to keep in mind that much of the gas assumed to be available is in reservoirs only accessible through fracking.

16. U.S. Energy Information Administration, 2022

Coal

Coal has been prized as an energy source for centuries. It was first used to produce electricity during the mid 1880s and has been an important fuel ever since. It is made of hydrocarbons so is chemically like our other fossil fuels. As a solid, though, it is physically different and not appropriate for the same applications (e.g., coal is not as portable as oil and gas, limiting its usefulness for transportation).

Formation. Like petroleum, coal originated from organisms that lived hundreds of millions of years ago. However, in this case, the starting point was dead terrestrial plants like ferns that settled into shallow and stagnant swamps, not marine plankton sinking to the ocean floor. The remains escaped decomposition in low-oxygen environments and were transformed into an organic-rich and unconsolidated material called **peat**. High temperature and pressure from burial induced a reduction in moisture content as water got squeezed out, an increase in density, and a higher relative amount of organic carbon in the deposits. At depths of a few hundred meters, and after many millions of years had passed, layers (ranging from a few centimeters to several meters thick) of a rock formally recognized as **coal** developed (Figure 10.18).

Figure 10.18. *Stages of coal formation. Upper left (a), remains of swamp plants escape decomposition; upper right (b), peat bog; lower left (c), coal seam; lower right (d), single coal rock (the one shown is approximately 8 cm across). David Monniaux, CC BY-SA (a); Miraculix HB, CC BY-SA (b); Peabody Energy, CC BY (c); USGS and Mineral Information Institute, Public Domain (d).*

It is important to note that the term “coal” represents several subtypes, known as **ranks** of coal. These ranks may be viewed on a continuum, representing increasing degrees of burial, age, and energy content. Figure 10.19 summarizes some important properties of different coal ranks.

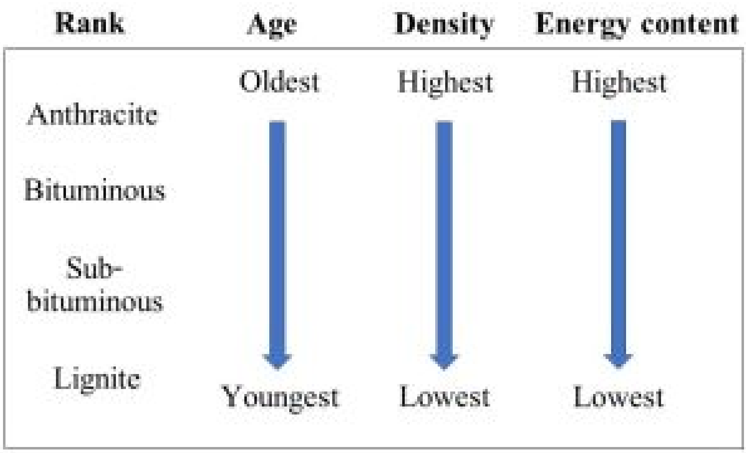


Figure 10.19. *Ranks of coal and some of their important properties. Kelsey,* CC BY-NC-SA.

Note that energy content is directly proportional to heating value. Coal **grade** is another way to categorize coal, and it is a function of how much sulfur it contains and some other properties related to its usefulness for various tasks. When we consider the environmental impact of coal in an upcoming section of this chapter we will see that low-sulfur grades are preferable, if not always available.

Environmental costs. As is the case for all fossil fuels, each stage in the life cycle of coal causes adverse environmental consequences.

Extraction

Some coal is found on Earth's surface, but much of it can only be obtained through the mining of buried material. Extraction techniques are divided into two broad categories: surface and underground mining.

Surface. When layers of coal—known as a **coal seam**—lie no more than about 60 meters beneath soil and rock, this approach can be used. It dominates the U.S. industry, responsible for approximately two thirds of coal extraction[17]. In short, machines remove the rocks and soils lying on top of the seam, known as **overburden**, and the coal is dug out and hauled away. Several specific approaches are grouped together here, including strip mining and mountaintop removal (MTR).

- Strip mining. This technique is used when coal is very close to the surface. Strip mining literally skims off long, thin swaths of overburden—strips of land—to reveal the underlying seam.
- Mountaintop removal. As the name suggests, MTR works on a larger scale than does strip mining. Here, entire mountain peaks are cleared, using explosives and related methods, and the resulting materials are deposited in nearby valleys (Figure 10.20.).

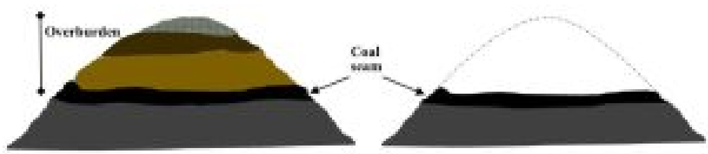


Figure 10.20. *Diagram of a mountain before, left, and after mountaintop removal, right. The dotted line on the right shows the former surface of the mountain. In cross section, not to scale. Kelsey, CC BY-NC-SA.*

Surface mining can cause many adverse consequences in affected areas, but just three important examples are described here. First, as land is cleared, existing ecosystems are destroyed. Consequently, the services they deliver, such as providing habitat for organisms, stabilizing **topsoil**, **fixation** and storage of carbon, and water retention, are lost as well (Chapter 5). Second, the removed overburden tends to accumulate in valleys, changing habitats and

17. U.S. Energy Information Administration, 2022

water **runoff** as well as spreading toxins that are the products of mining. Such areas can remain hostile to **ecological succession** ([Chapter 5]) and the return of healthy ecosystems for indefinite periods of time after mining ceases. Third, the digging of rocks and soils that have been buried for thousands or millions of years leads to a phenomenon known as **acid mine drainage**. Briefly, when certain sulfur-containing minerals such as pyrite are unearthed during mining, they undergo a chemical reaction that produces a very corrosive and concentrated substance known as sulfuric acid. Runoff and infiltration flowing from mines carry this acid and other toxic substances to nearby reservoirs (Figures 10.21 and 10.22).

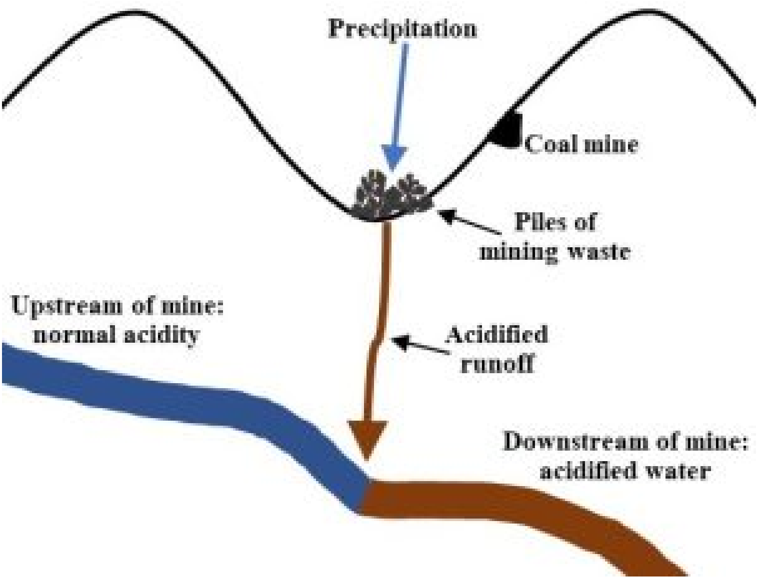


Figure 10.21. *Diagram of acid mine drainage. Water flows through the piles of mining waste and runoff carries the acids to a nearby stream. In cross section, not to scale. Kelsey, CC BY-NC-SA.*

Figure 10.22. *The visible effects of acid mine drainage on a stream. US EPA, Public Domain.*

Surface mining is quite controversial, and arguments about the relationship between its costs and benefits rage. Opponents note the extensive damage caused by the practice, especially the widespread changes brought about by mountaintop removal; images of portions of states like West Virginia and Kentucky show some dramatic effects of strip mining and MTR (Figures 10.23.a. and 10.23.b.).

Figure 10.23. *Surface mining in Germany, left (a), and Kentucky, right (b), cause large-scale changes to Earth's surface. Ekem, CC BY-SA (a); iLoveMountains.org, CC BY (b).*

Anti-MTR people also point out the long-term loss of habitat caused by the accumulation of toxic substances in mined regions. On the other hand, proponents of the practice, including mining companies, see the situation differently. In fact, to reflect their more positive outlook, they use a different name: MT**M**, in which the second M refers to "mining". They cite how surface mining allows for far more recovery of coal from a seam than does underground mining (described below) and is the less expensive of the two approaches. Risks to human life are also reduced when miners do not enter deep and dangerous shafts. Furthermore, supporters point out how federal laws require responsible parties to return former mines to their pre-mining condition. You should realize, though, that **remediation** of areas subjected to strip mining is extremely difficult and yields mixed results.

Underground. If coal is too deep for surface extraction to be practical, this second technique is used. In the U.S., only about one third of mining is done this way, although that number is not consistent from region to region: for example, more than three quarters of the coal taken from the Appalachian Mountains comes from underground mines[18]. These sites consist of long shafts dug from the surface down to the appropriate depth—often around 300 meters, although they can be 500 or more meters deep—and chambers that spread out laterally to follow a seam. Miners use various tools and strategies to remove coal from caverns. Rail cars and conveyor belts move the extracted rocks to the surface (Figures 10.24 and 10.25).

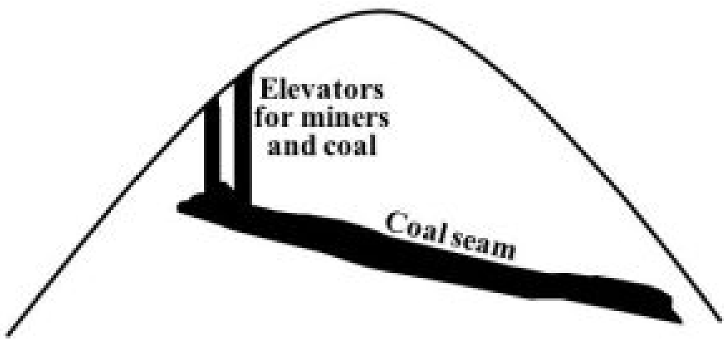


Figure 10.24. *Diagram of an underground mine. In cross section, not to scale. Kelsey*, CC BY-NC-SA.

18. U.S. Energy Information Administration, 2022

Figure 10.25. *Coal is transported from mines to the surface by rail. U.S. National Archives and Records Administration, Public Domain.*

Underground differs from surface mining in several important ways. Most obviously, people must descend the hundreds of meters necessary to reach a seam and subject themselves to great danger. The risk of death from asphyxiation, exposure to poisonous gases, explosions, and mine collapses continues to be quite real to this day. Additionally, underground mining is less efficient than is surface. Since underground coal provides structural support for the mines themselves, only a fraction of it can be extracted before conditions become too hazardous to continue (the amount varies among sites).

Like surface mining, underground mining causes environmental degradation in the form of acid mine drainage (review Figures 10.21 and 10.22, above) as well as damage to existing ecosystems. The latter issue is less pronounced here because overburden is not removed, however, the roads, rails, and other infrastructure which are built to support mining are disruptive. Finally, mine fires are a serious, albeit somewhat quirky and infrequent, consequence of coal mining. Since buried coal is in a largely oxygen-free environment, it is not likely to burn while under the ground (oxygen is needed for combustion). However, once mining commences, the situation changes because shafts leading from the surface allow O_2 gas into the coal seam. An explosion or other ignition source can start a fire that burns for years or decades, possibly continuing until all the coal has been consumed. Very large coal mine fires are currently burning in India, China, and Pennsylvania, U.S.A., riddling these places with sinkholes and smoking vents, and generally rendering them uninhabitable. In the last case, the town of Centralia has been essentially eliminated due to a fire burning since 1962—little more than roads and warning signs are all that remain (Figure 10.26).

Figure 10.26. *A sign posted in Centralia, PA, warning of the hazards of a long-term problem. Lyndi and Jason,* CC BY-SA

Transport

Mined coal is moved to processing plants and the places at which it is burned. As always, such transport requires energy from the combustion of fossil fuels for coal to be available to users. Of course, unlike oil and gas, it is solid rocks—so rail cars, barges, and trucks rather than pipelines are employed.

Processing

Refining of coal is not necessary to produce a useable fuel—it can be burned as soon as it is extracted. Coal is often subjected to some processing, though, to improve its quality: rocks, chemical impurities, and other materials that reduce its efficiency as an energy source can be removed before it is shipped to users. Also, it generally is pulverized before it is burned. As with all the other steps, this one is powered by the burning of fossil fuels so is associated with air pollution.

Use

The role played by coal has changed during the past century or so. Historically, energy released during its combustion was widely used to heat buildings, power the movement of trains and ships, and starting in the late 19th Century, generate electricity. During the early and middle parts of the 20th Century, coal became less and less important in transportation and the direct heating of spaces because it was replaced by various petroleum products in cars, trains, and homes. The reasons for this change include the more portable nature of liquids like gasoline and diesel as well as the relatively large amount of air pollution released when coal is burned. It continues to be heavily used in electricity generation, though (as well as in steel and cement making). Notably, it is still responsible for nearly 40% of worldwide electricity production, although its importance varies by country[19].

The combustion of coal produces the pollutants we saw for other fossil fuels. Three warrant some additional attention here, and we will return to them again in Chapter 14. First, coal burning releases more carbon dioxide per unit of energy than oil, exceeding that from gasoline by 30 – 40% (depending on the rank of coal used—recall Figure 10.19). Second, the sulfur and nitrogen oxides emitted from coal-fired power generation facilities can cause a great deal of environmental damage, particularly when high-sulfur grades are used: smog and acid precipitation are two consequences we will explore in some detail. Third, coal burning produces fly ash (light enough to rise into the atmosphere) and bottom ash (adheres to the floor and walls of a furnace), two solids *not* released by oil and gas combustion. These last pollutants are generally recycled and used as additives in the manufacturing of materials like cement or buried in landfills (Chapter 13).

19. International Energy Agency. iea.org

Supply and demand. As we saw with oil and natural gas, humans have come to depend on coal. For example, in 2021, it accounted for approximately 11% of total energy supply and 22% of the electricity generation in the United States[20]. We know natural processes form this non-renewable fossil fuel far more slowly than humans use it, so we are keenly interested in the many variables influencing coal's availability, use, and long-term prospects.

Most of the world's coal is concentrated in just five countries

China and the United States together control almost 50% of the world's reserves of coal, and about 30% is found in Russia, Australia, and India[21]. The rest is in South America, Asia, Africa, and Europe.

The value of coal varies

It is important to realize that coals differ in many ways and that they are not all equally valued. As we saw above, rank is one critical consideration (review Figure 10.19). Anthracite is the densest and hardest, and it is prized because of the large amount of energy released when it is burned—that is, less of it is required to power the same amount of work compared to, say, low-carbon and soft lignite. Unfortunately, anthracite is by far the least abundant (and, therefore, costliest) coal, making up only about 1% of reserves. Bituminous and sub-bituminous coals are softer coals than anthracite and have lower energy content per unit mass, however, since they are most abundant, they are most often used. A second important property is sulfur content. It is fair to say that low-sulfur coals are most sought after because they burn relatively cleanly, but as with anthracite, their supply cannot keep up with demand in many places.

The U.S. has abundant coal

The United States possesses large reserves, as well as unproven resources, of bituminous and sub-bituminous coal. It is extracted

20. U.S. Energy Information Administration, 2022

21. U.S. Energy Information Administration, 2022

by miners across the country and has been a dominant source of energy for well over a century. More than 40% of that coal is produced in Wyoming, and most of the rest (about 30%) comes from states within the Appalachian Mountains and Illinois[22]. Worth a final note: coal from western states such as Wyoming has less sulfur in it than does that from the east.

How long will coal last?

The amount of proved worldwide reserves is 1100 billion tons of coal, and current annual demand is around 8 billion tons[23]. So, we can use the same type of simple calculation used for oil and gas to estimate that coal would last for approximately **140 years** if demand remains constant. Some people are more optimistic, though (including the U.S. Energy Information Administration, U.S. EIA) projecting that supply will last another **several hundred more years**[24]. As with oil, different assumptions about probable supply and future demand yield vastly different predictions.

Will it be used extensively in the future?

Coal could be available for human use for many decades, if not centuries. Whether it will, or even should, be a primary energy source in the future depends on several considerations, though. First, as noted above, extraction of coal causes extensive environmental damage—the effects of surface mining, for example, are undeniable and very difficult to reverse. Second, burning of coal generates chemical and physical products that can substantially degrade the quality of air. In fact, its use as a fuel to heat American homes declined precipitously during the middle of the 20th Century in part because of the smog and other air pollution problems it

22. U.S. Energy Information Administration, 2022

23. World Coal Association. 2022. What is coal and where is it found? worldcoal.org

24. U.S. Energy Information Administration, 2022

caused. Since that time, laws have been enacted and measures taken to reduce the release of many of the most troubling products of coal combustion, like sulfur and fly ash, but such clean-up is quite costly. Still unsolved, however, is the problem of carbon dioxide, a major contributor to global climate change. In the U.S. alone, coal-powered electricity plants release approximately 900 metric tons of CO_2 annually (or 60% of the 1600 metric tons released from all electricity generation)[25]. How to capture that gas before it enters the atmosphere is a current area of research, but practical and cost-effective solutions appear to be years away. We will revisit coal's role in the generation of air pollution, as well as strategies used to reduce its negative impacts, in Chapter 14. Finally, the availability of other sources of energy, particularly renewables that bring about far less environmental disruption, will influence discussions of whether the costs of coal are justified by the benefits provided by its continued use.

10.4.2. Nuclear power

Now we turn from fossil fuels to explore a very different strategy to generate power. In this case, the energy of certain types of atoms is converted into electricity through means other than combustion. We begin with a brief description of the process and then will consider the environmental consequences and prospects of this non-renewable energy source. Note that not all nuclear power plants are the same, and the information provided here is intended to provide a generalized overview.

25. U.S. Energy Information Administration, 2022

Nuclear fission releases large amounts of energy

Recall from Chapter 4 that atoms are the minuscule building blocks of everything in the Universe and that each of them is comprised of even smaller particles known as protons, electrons, and neutrons. Since these subatomic particles are held together with very strong bonds, it is quite difficult to break atoms apart. When they *are* split, though, during a process known as **nuclear fission**, an enormous amount of energy is released. Under the right circumstances, some of that energy can be captured and converted into electricity.

Unstable atoms can be split. For several reasons, most notably the fact that they are relatively easy to split, some uranium atoms (U) are well suited for nuclear power generation. In fact, certain U **isotopes**, referred to as U^{235}, are so unstable that they can spontaneously break apart (the superscript indicates the number of neutrons in the atoms, which can differ for atoms of the same element). Since natural U^{235} is found in very low concentrations in rocks and soil, mined uranium must be extensively processed before it is fit to serve as a fuel. Several steps and specialized equipment—along with a great deal of energy—are necessary to produce a usable fuel. This product, known as **enriched uranium,** consists of a mixture of uranium isotopes: 2 – 4% is U^{235} and the rest is U^{238} (we will see more about this second isotope shortly). Fission (splitting) releases several products: smaller atoms (known as **fragments**), neutrons, ionizing radiation, and heat.

Fragments

The original, or **parent** atom, is converted into smaller, **daughter** atoms. Since the initial daughters produced are also unstable, they too break into even smaller fragments; this process, termed **radioactive decay**, continues through several steps until stable atoms form. Note that unstable atoms vary widely in the time required for them to become stable, ranging from fractions of seconds to billions of years. This period is quantified with a unit known as the **half-life**, or the time required for half of the parent atoms present in a space to undergo conversion to daughters.

Neutrons

These **subatomic particles** are important products of fission because after they are released they can strike other unstable atoms and initiate additional fission. Nuclear reactors rely on interactions among neutrons and fissionable atoms to operate. We will return to the important role of neutrons in power generation below.

Ionizing radiation

This is a type of invisible energy that can travel through air (or even a vacuum) and induce profound chemical changes in the objects it encounters. For example, by transforming important structures inside of cells, radiation can damage or kill organisms. Gamma radiation—a very dangerous type that can penetrate deep into both living and non-living materials—is released in nuclear fission. Bodies that emit gamma or one of several other types of radiation are said to be **radioactive**.

Heat

Fission of U^{235} also releases enormous amounts of heat, meaning that a relatively small mass of material can be used to generate a great deal of electricity (more below). Heat from fission reactions is used to produce steam.

Nuclear fission can generate power. Simply put, a nuclear reactor is a reinforced chamber into which fuel (usually the mix of U isotopes described above) is submerged in water and allowed to undergo fission. The process involves several important steps.

A neutron gets it started

Uranium atoms are placed close together, and a neutron released when one undergoes fission collides with an adjacent one. That second atom then splits and releases one or more neutrons, leading to the fission of additional atoms.

A chain reaction ensues

At first, a small number of atoms are transformed into products. As more and more are split, the number of neutrons in the reactor increases, causing more and more atoms to undergo fission (Figure 10.27). Such a process is known as a **chain reaction** and ought to look somewhat familiar because it resembles the many examples of positive feedback we have seen throughout this book (first encountered in Chapter 2). As the reaction grows, so does the amount of heat released into the water. Eventually, the proper temperature is reached, and steam is continuously sent from the reactor to spin a turbine.

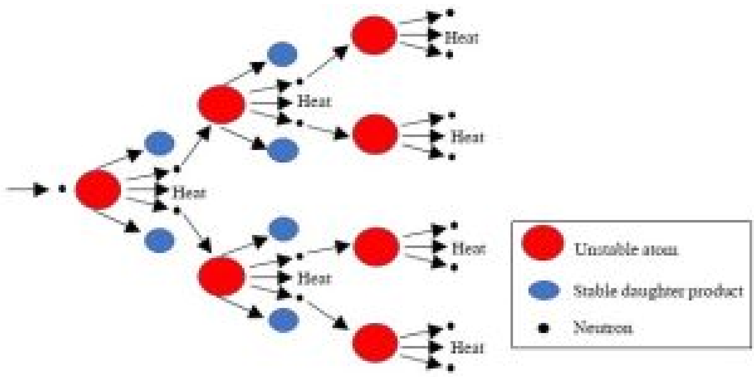


Figure 10.27. *A diagram of a chain reaction. The heat released at each step is used to boil water; the neutrons released at each step split additional atoms. Kelsey*, CC BY-NC-SA.

Plutonium is produced

While the U^{235} undergoes fission, another important process occurs inside the reactor: some of the U^{238} atoms present (the bulk of the material) are converted into plutonium (Pt) atoms. Atoms of one of the plutonium isotopes, Pt^{239}, also decay and release a great deal of additional heat (which contributes to steam generation).

The chain reaction proceeds in the reactor core. A nuclear reactor consists of several essential components (Figure 10.28).

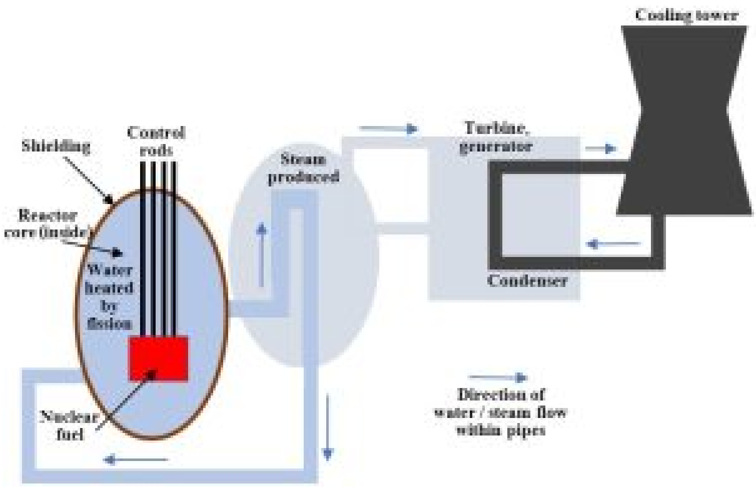


Figure 10.28. *Simplified nuclear reactor and associated structures. Details about the components shown are provided in the main text, below. Not to scale. Kelsey*, CC BY-NC-SA.

Nuclear fuel (fuel rods)

Uranium is packed into long, thin metal tubes. Several of these fuel rods are then assembled and can be lowered into a reactor.

Water

Although the details of this step vary among power plants, the heat released during fission ultimately boils water and generates steam that is directed at turbines to generate electricity. The water is maintained in a closed loop and used repeatedly, so it must be cooled and condensed after it has boiled and then cycled back into the reactor. The critical cooling step is generally accomplished using hyperboloid or hourglass-shaped cooling towers that are so often associated with nuclear power (these are large, well over 100 meters tall; Figure 10.28 shows their basic shape). A second batch of cold water taken from a nearby surface reservoir is placed close to pipes carrying hot steam from inside the reactor. Re-condensed water is then ready for another round of boiling and electricity generation.

Control rods

These metal cylinders are constructed of materials that absorb neutrons without undergoing fission and perform two important functions. First, they help keep a reaction running at a constant rate and make electricity generation as efficient as possible. Briefly, they are lowered into the reactor core once the water temperature has reached the target level and are subsequently retracted and inserted to maintain it. Second, they provide a means of reducing the risk of an accident. Without control rods, a chain reaction could proceed rapidly enough that the chamber in which it starts can no longer contain it. In other words, many atoms would simultaneously release energy and cause an explosion. For reasons that are likely obvious, operators of nuclear power plants would prefer to avoid exploding reactors. Many precautions are taken to minimize the likelihood of a runaway chain reaction, but control rods are among the most important.

Shielding

Nuclear reactors are built to prevent radiation from escaping into natural environments. Gamma radiation is the biggest concern as it can pass through solid objects (as noted above), so the core is typically constructed of thick steel and concrete. Water serves as a shield as well.

Waste is removed and stored. Two types of radioactive waste must be managed.

Low-level waste

This emits relatively little radiation. Some of the materials used to generate electricity at nuclear power plants fall into this category. Various storage strategies are used to protect organisms, soil, and water while its radioactivity is lost due to natural decay, but it generally does not pose a serious threat.

High-level waste

Although it makes up less of the total volume of waste produced at a nuclear power plant, it presents by far the bigger risk of the two types. Simply put, it consists of uranium, plutonium, and other fission products. As it contains materials that are quite radioactive (as well as very hot) this **spent fuel** must be handled carefully to ensure that humans and natural systems do not contact it. The first step in its management is to place fuel rods in large tanks of water for several years to allow them to cool and provide time for decay to remove some of the remaining radioactivity[26]. Then, the material is moved to dry storage. These two stages generally occur at a power plant and are intended to serve as a temporary holding strategy. Some of the waste products have very long half-lives, so they will continue to release radiation for thousands of years. Still unresolved is what to do with these dangerous materials because no permanent storage solution currently exists for the U.S. (the country that generates the most electricity through nuclear fission). We will return to this point below.

Nuclear power provides benefits and poses risks

Nuclear power has both advantages and disadvantages. Like fossil fuels and the other energy sources described in this chapter, we can use science to help us decide if the benefits of using it are worth the risks.

Advantages. There are reasons nuclear power is favored by many people as an energy source.

It releases no carbon emissions

Advocates point out that nuclear fission releases a great deal of energy without emitting carbon dioxide. Relative to fossil fuel combustion, it does not increase the concentration of CO_2 in the atmosphere or directly contribute to global warming (Chapter 14).

It releases no radiation

26. You should realize that the only known remedy for radiation is time

Normally, radiation emitted by unstable atoms is contained inside the reactor by various structures.

It uses relatively small amounts of material to generate large amounts of electricity

Far more energy is concentrated in U^{235} than is found in coal, the largest global contributor to electricity production: just one ton of uranium can generate the same amount of electricity as 16,000 tons of coal[27].

It generates relatively little physical waste

About of 2000 tons of spent fuel is produced annually in the United States[28]. For comparison, coal combustion produces about 130,000,000 tons of ash each year[29].

Disadvantages. There are costs and risks of nuclear power and reasons to be cautious about it.

The risk of catastrophic accident is low but not zero

The many redundancies and safety features built into nuclear power plants make nuclear accidents unlikely, and it is fair to say that a great deal of power is generated without radiation leaks or other incidents. However, the potential for destructive events certainly exists, as shown by three serious accidents that have taken

27. Argonne National Laboratory, U.S. Department of Energy Office of Environmental Management, 2017, Depleted UF6 guide. http://web.ead.anl.gov/uranium/guide/facts/

28. Argonne National Laboratory, U.S. Department of Energy Office of Environmental Management, 2017, Depleted UF6 guide. http://web.ead.anl.gov/uranium/guide/facts/

29. U.S. Environmental Protection Agency, 2017, Coal ash basics. https://www.epa.gov/coalash/coal-ash-basics

place in the past four decades[30]. The first was in 1979 at the Three Mile Island power plant near Harrisburg, PA (U.S.A.). In this case, the outcome was relatively minor, but some amount of radiation escaped the confines of the facility. Due to multiple errors and technical issues, one of the reactor cores underwent a partial **meltdown**, which is more or less what the name suggests: cooling and monitoring systems failed, allowing the core to get so hot that parts of it melted. It could have been far worse, for example, an explosion that would have exposed millions of people to high levels of radiation. As it was, public health concerns were heightened but widespread effects on the nearby population were never confirmed. Very importantly, public opinion turned decidedly against the nuclear industry, and no new facilities were opened in the U.S. during the decades following the accident (until 2023, as noted below). Furthermore, many new regulations were introduced to improve the safety of existing plants. The second occurred in 1986 in at the Chernobyl nuclear plant in Ukraine (then in the U.S.S.R.). It was a devastating accident, as an explosion destroyed parts of the power plant and allowed high levels of radiation to escape and travel thousands of kilometers throughout Europe and Asia. Around 30 people are thought to have died in the initial accident, but thousands more have been affected by chronic conditions like cancer in the years since. Today, even though the damaged reactor is encased in a containment structure, the zone within a 30-km radius of the plant is all but off limits to humans. Thirdly, the Dai-ichi Fukushima nuclear power plant of Japan was damaged by an earthquake and subsequent tsunami in 2011, leading to meltdowns of multiple reactors and the release of radiation and radioactive material into air and water. Tens of thousands of people were evacuated from affected areas, and due to continuing high radiation in many of those places, few have been allowed to return.

30. U.S. Nuclear Regulatory Commission, 2017. https://www.nrc.gov/

Long-term health effects are still to be determined, but this massive accident again raised worldwide doubts about the safety of nuclear power. Some countries even vowed to abandon their usage of fission in response to the event, citing what they saw as unacceptably high risk.

We have no long-term plan for spent fuel

Arguably, this issue is the biggest problem with nuclear power and the one that is currently the most intractable. Half-lives of the different radioactive atoms in the waste range from around 30 to 24,000 years (plutonium is on that high end), meaning that the mixture of these life-threatening materials must be securely isolated for tens of thousands of years. In the United States, the need for a designated storage facility was identified decades ago, but one has not yet been built. A place in Nevada known as Yucca Mountain was extensively studied and nearly chosen as the site, but due to many scientific, technical, security, and political issues the plan has been scrapped. Other countries face similar problems. The search for a suitable solution, at a national or perhaps even international level, continues. For now, nuclear power plants produce radioactive waste and hold it in the structures we noted were only designed to be temporary solutions. Unfortunately, some of these containers are disintegrating, raising concerns about potential environmental contamination. The amount of high-level waste needing long-term storage *could* be reduced if spent fuel was treated to extract some of the usable isotopes and used to generate more electricity. In other words, instead of storing waste for thousands of years, the useable fraction would be **reprocessed** into new fuel. The U.S. has prohibited the reprocessing and reuse of plutonium from spent fuel to reduce nuclear proliferation, however (next item). Some countries using nuclear power do not adhere to such rules.

It could contribute to the proliferation of nuclear weapons

Nuclear power is sometimes referred to as a *peaceful* use of fission. Unfortunately, some of the same equipment used to enrich uranium for power generation can produce fuel concentrated enough for a nuclear weapon. In addition, high-level radioactive waste contains some chemicals that could be incorporated into weapons. For example, plutonium, a product of uranium decay, could be isolated from spent fuel and built into a bomb.

Cost of plant construction and decommissioning are high

Nuclear power plants are very expensive to build, operate, and ultimately shut down when they have been taken off line. The construction costs are quite difficult to estimate because they vary so much by location and often go over budget in any case; plus, since only one new plant has been opened in the U.S. (2023, in the state of Georgia) in the past few decades, the real cost of a new one is particularly hard to calculate with much certainty. A figure in the range of at least $10 to $15 billion is likely. Often overlooked by most people is the big expense that arises at the other end, that is, when a plant is shut down after its useful life has been expended (more below). Among other tasks, all the radioactive material must be handled, and the facility decontaminated. Again, it is hard to come up with a definitive figure here, but certainly hundreds of millions of dollars at a minimum is reasonable. Running a plant (labor costs, maintenance, etc.) costs money too, of course, but relatively speaking, this is a small figure.

Mining of uranium has environmental costs

As noted earlier in this section on nuclear power, uranium is obtained through mining. Although it might go without saying, a brief reminder is in order: mining is responsible for a great deal of environmental degradation. Many of the consequences of coal mining such as land clearing, ecosystem damage, soil erosion, toxic runoff, and air pollution, for example, are seen here as well.

What is the future of nuclear power?

We will end this section with some thoughts about the role nuclear power will likely play in the future and some factors affecting its long-term viability as an energy source.

Usage[31]. Nuclear power has met approximately 20% of electricity demand (9% of total energy) in the United States for most of the past 30 years and is projected to remain near those levels for the next 30. Thirty other nations rely on fission to generate electricity to different extents, such as France (where it accounts for nearly 80% of demand) and China (about 2%). Again, the U.S. generates more electricity from nuclear than any other nation—nearly double that of the second-biggest user—but nuclear accounts for a smaller percentage of its total than, say, that of France because usage in the former country is so much higher than that in the latter.

Uranium supply. Various estimates suggest the fuel supply is likely to last for another century or so, assuming neither the technology of nor demand for nuclear fission changes during that time. Now, it would be possible, under a set of unlikely circumstances, for the availability of useable nuclear fuel to be extended by many centuries if a different plant design, known as a **breeder reactor**, were widely used. In short, reprocessing of nuclear waste, including the plutonium we discussed above, would increase fuel supply and decrease the need for extensive management of high-level radioactive waste. Many obstacles, including very high costs to build hundreds (or more) necessary new facilities, security concerns, safety issues, and considerable political opposition to nuclear power, make such a system extremely difficult to bring to fruition. A second point about supply merits a brief comment here: although uranium mines are widely distributed, Kazakhstan, Canada, and Australia control about 70% of the world's reserves. In other words, most countries with nuclear programs import the bulk of what they use. In the U.S., for example, domestic mining only meets about 10% of uranium demand[32].

Meeting demand as old plants are decommissioned. Nuclear

31. U.S. Energy Information Administration, 2022

32. U.S. Energy Information Administration, 2022

plants are not designed to operate indefinitely—eventually, they must be shut down and replaced with new ones if the share of electricity generated by nuclear power is to remain constant. In the U.S., such facilities are licensed by the Nuclear Regulatory Commission (NRC) to operate for 40 years, and the current average age of the 61 facilities (99 total reactors) is 36 years[33,34]. Now, it is common for operators to apply for (and receive) a 20-year extension to their license, and four new reactors are currently under construction. What will happen to nuclear power during the coming decades, however, is by no means certain.

Management of nuclear waste. Many people argue that nuclear fission is generally a safe, efficient, and effective way to generate electricity, but the problem of what to do with high-level radioactive waste remains unsolved. Without a permanent solution in place, nuclear power presents a continuing threat to environmental and human health. How much, and even whether, it can continue to serve as an energy source will depend in large part on how much risk people are willing to tolerate as well as how the U.S. and other countries choose to manage their waste in the future. We will return to this and related questions when we explore the science of risk assessment in Chapter 15.

A role in efforts to combat climate change? As noted here already, under normal circumstances, nuclear power plants emit no carbon emissions. We will learn a lot more about the effect of CO_2 on climate in Chapter 14, but for now let us just say that meeting growing worldwide demand for electricity through fossil fuel combustion will be highly problematic. Some argue that nuclear fission provides an environmentally friendly way to provide

33. U.S. Energy Information Administration, 2022

34. Other countries have their own strategies for oversight and management, and the International Atomic Energy Agency works to promote the safe and peaceful use of nuclear power worldwide.

power without air pollution and should, therefore, be expanded. Work to design safer and cleaner reactors that generate less-risky waste is ongoing, but whether it produces viable alternatives to other sources of energy remains to be seen.

Part III: RENEWABLE ENERGY SOURCES

10.5. RENEWABLES OFFER ALTERNATIVES

Renewable energy sources—sometimes called **alternative energy sources**—have become increasingly relevant during the past few decades for two reasons. First, on-going volatility in the price of petroleum and worries about the long-term supplies of fossil fuels have driven many to explore options that are, at least in principle, free and inexhaustible. Second, extracting, processing, and using non-renewables lead to substantive environmental degradation, including climate change (Chapter 14). Although renewables are by no means without environmental cost, it is fair to say they are generally more benign than non-renewable sources. Here, we will examine several renewable energy sources, briefly considering how they are harnessed to generate power, their advantages and challenges, and their potential to supply energy in the future.

10.5.1. Direct solar

The sun provides power for nearly all the processes active on Earth, including **photosynthesis**, wind, and the cycling of water. These and other *indirect* manifestations of solar energy are clearly vital to human survival, as described in previous chapters. An understanding that the sun also can serve as a *direct* energy source developed relatively recently, although it is hardly a new idea: people have exploited the sun's heat to raise the temperature of homes, dry objects, and even cook food, for millennia. During modern times, we have developed additional uses for direct solar, namely, to heat water and interior spaces as well as produce electricity.

Ways to use solar energy directly

Heating. The sun's energy can be harnessed to raise the temperature of water or air through one of two mechanisms. In the first, **passive heating**, objects are exposed to sun and their temperature is increased without the aid of any moving parts or distribution systems. For example, windows can be installed in the south-facing side of a building to maximize natural heating of an enclosed space (Figure 10.29).

Figure 10.29. *Passive solar heating. Use of sun-facing glass maximizes energy absorption and heating of this building. Marion Schneider & Christoph Aistleitner, Public Domain.*

The second is known as **active heating**. Here, a substance with a high capacity to absorb and hold heat (a fluid of some type) is exposed to sunlight, and a mechanical system then distributes the heat. The pumping of sun-heated water throughout pipes in the walls of a home is just one application of this approach. Solar-energy-driven domestic hot water heaters are also a very common use of active heating.

Electricity generation. Electricity can be produced from solar energy using two distinct approaches.

Photovoltaic cells (PV cells)

These are made of specialized materials that absorb solar energy in such a way that electrons are induced to move (i.e., electricity is generated). Individual cells are generally small and flat, a few centimeters square and less than one millimeter thick, but many of them can be arranged together on the surfaces of solar panels (Figure 10.30).

The size of panels depends on how much electricity is needed, so just two or so centimeters will suffice to run a calculator, but several panels that each are a meter or more are required to provide enough power for an entire house. Hundreds of panels arranged on a rooftop or in an open field could generate larger amounts of electricity, enough to supply many homes, businesses, or communities with needed energy.

Figure 10.30. *Many small solar cells can be grouped to form larger panels like the ones shown here. David Monniaux, CC BY-SA.*

Thermal power generation facilities

The idea is deceptively simple: heat from the sun is used to generate steam which then spins turbines. The large scale of these operations presents design challenges that have been approached in different ways. In any case, a great deal of sunlight must be directed at a vessel of water for this to be an effective system. For example, moving mirrors can concentrate the sun onto a relatively small area, raising the temperature of a fluid to nearly 500 °C (around 1400 °F).

Figure 10.31. *Thermal solar power generation with a solar power tower. Afloresm, CC BY.*

A small number of thermal power generation plants are operational, and research continues into ways to improve them. Figure 10.31 shows a solar power tower, one such facility.

Advantages and challenges of solar

Interest in solar energy has increased steadily during the past few decades for a simple reason: it holds great potential as an energy source. However, some challenges continue to limit its widespread use and ability to replace non-renewables such as coal. Here we take a look at both the ups and downs of solar power.

Advantages. The list of reasons to embrace solar power consists of economic, environmental, cultural, and security-related arguments.

Its supply far exceeds demand

The enormous amount of energy moving from the sun to the Earth more than meets our needs. One useful way to visualize the quantity is suggested by the U.S. Department of Energy (DOE): enough energy hits the Earth in 90 minutes to meet human demands for a year[35].

The energy itself is free

Solar energy is readily available and not controlled by any one country or group. To be sure, some places on Earth receive more than others, but it is accessible to anyone with the proper equipment to harness it.

Its environmental impact is low

As we will see below, solar energy is associated with some amount of risk to natural systems. Relative to fossil fuels and nuclear power, however, the effects of solar are small. Importantly, solar power is classified as a non-carbon energy source, and, as such, it does not release carbon dioxide (or any other air pollutants, for that matter) into the atmosphere.

It is renewable

Solar power is listed in this section because it fits our definition of renewable energy. It is inexhaustible in any practical way, potentially providing energy to humans in perpetuity (or as long as we are likely to need it, i.e., it will not run out for about 5 billion years).

35. U.S. Department of Energy, 2022. energy.gov

Challenges. If solar is so great, why doesn't it make up the bulk of our energy supply? Some of the answers can be found on the following list.

It is dispersed

This problem is linked to the first item on the list of advantages. Yes, supply exceeds demand, but the energy is spread out over the entire planet. In other words, each square meter of the surface receives a small percentage of all incoming sunlight. Panels must therefore be large if they are to collect enough solar energy to generate relevant amounts of electricity.

It is variable

As is quite familiar, the amount of sun striking a portion of the Earth is not constant. For example, the sun sets in the evening, cloud cover changes, days are shorter in winter, and equatorial regions receive more direct sunlight than do those at high latitudes. Put another way, solar energy is intermittent and variable. These and other factors mean that the amount of electricity generated from solar power will differ in both space (i.e., location), and time (i.e., season, month, and hour). It is simply a reality that some areas are better situated for solar power than are others. For example, it is more efficient and realistic in southwestern portions of the United States than it is in northern Canada.

There are technological obstacles

The first two problems on this list, the dispersed and variable nature of solar energy, are made even more challenging by the state of the currently available technology. Even though we have clearly learned to effectively capture energy from the sun and convert it into heat and electricity, some improvements would allow solar power to make a more substantial contribution to our energy supply. For example, photovoltaic cells are inefficient, converting only about 20% of received solar energy into electricity (worth noting, though, is that efficiency has been steadily improving during recent years). Furthermore, better batteries or other storage solutions are needed to increase the viability and relevance of solar power.

It is costly to harness

We noted above that the energy itself is free and readily available. The technology required to convert sunlight into electricity requires a financial investment, however. Yes, it has gotten steadily cheaper during the past few decades—and is competitive with or even less expensive than conventional energy sources in a few places—until recently the cost of solar has been quite a bit higher than that of fossil fuels and only affordable by a small number of people. Centralization is an important key to the viability of solar: large power generation plants that use photovoltaic (PV) systems can provide the same quantity of electricity for far less money than can an equivalent combination of small PV systems installed on individual homes.

Solar brings its own environmental disruptions

Solar power has a relatively small effect on natural systems, but two important concerns are worth noting. First, the chemicals in batteries and photovoltaic cells are toxic and therefore pose a risk to humans and other organisms if not properly managed.

Second, given the need for large panels noted above, solar arrays cover a lot of surface area. Land clearing can affect ecosystems and increase rates of erosion (we have seen as much in Chapter 9 and earlier in this chapter) but is not inevitable if panels are placed on existing buildings and other structures instead of on open ground (Figure 10.32).

Figure 10.32. *Rooftop solar array on an existing structure. hjl, CC BY (modified by Kelsey).*

How much energy could we get from the sun?

Stooping to the use of a somewhat painful pun, the future of solar could be very bright—*if* the challenges enumerated above are met. As long as coal and natural gas remain cheap compared to solar, though, the former sources will continue to thrive. Certainly, concerns about air pollution and other environmental degradation associated with fossil fuels will drive some shift toward solar, but cost is likely to exert more influence over energy choices made by individuals and governments. What if we ignore costs for the moment? How much of our demand for energy could realistically be met by the sun? If you consider that only about 3.7% of worldwide electricity currently comes from the sun[36], you might not be surprised to read that a lot of untapped potential is still out there. According to estimates described by the United States Department of Energy (DOE), if fully exploited, rooftop PV alone could meet 39% of U.S. energy demand[37]. The addition of large PV power generation plants would increase that number considerably, optimistically up to 90% under, admittedly, extraordinary circumstances. Many governmental and private organizations have weighed in on the question, and, depending on the assumptions they make, the amount of electricity likely to come from the sun will range between 3 and 30% by the year 2040. Put another way, the gap between potential and actual amount of energy supplied by

36. Piotr Bojek. 2022. Solar PV. The International Energy Administration. https://www.iea.org/reports/solar-pv

37. Gagnon, P., Margolis, R., Melius, J., Phillips, C., Elmore, R., 2016, Rooftop solar photovoltaic technical potential in the United States: a detailed assessment. National Renewable Energy Laboratory, U.S. Department of Energy, Technical Report NREL/TP-6A20-65298

solar power will probably continue to be large, but just how big the difference will be is very hard to predict with certainty.

10.5.2. Wind

Due to many factors, including differential heating by the sun, the temperature of Earth's atmosphere is not uniform. The resulting pockets of relatively warm and cool air move, and wind is generated. Humans learned thousands of years ago that wind can be harnessed to do work, including powering sailing ships and pumping water. Wind was first exploited to generate electricity on a small scale in the late 1800s, became more efficient and viable in the 1970s, and expanded in its importance during the past three decades. Today it is widespread, relatively mainstream, and likely to supply an increasing amount of electricity in the future.

Converting wind to electricity

Turbines spin. We saw earlier that electricity can be produced when steam turbines are connected to a generator. Wind energy is similar: structures known as **wind turbines** are turned by moving air, and that **kinetic energy** is converted into electricity (Figure 10.33). There are some differences, such as wind turbines do not spin at a constant rate, but many of the basic principles are the same.

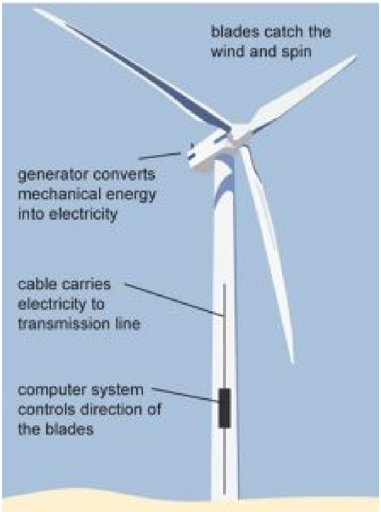


Figure 10.33. *Basic turbine design. US National energy education development program, Public Domain.*

Scale varies. The size and number of wind turbines constructed in an area affects the amount of energy captured. For a single house, one small turbine is generally sufficient, but larger numbers of turbines, each of which could be 60 meters tall and possess 30-meter-long blades, can be clustered together in a centralized **wind farm** (Figure 10.34) to produce enough electricity to meet demand for communities and cities.

Figure 10.34. *An off-shore wind farm featuring multiple turbines in the ocean, left (a); wind farm on land with turbines of varying heights to maximize energy capture (note each turbine is over 100 meters tall), right (b). Mariusz Paździora CC BY-SA (a); James McCauley, CC BY (b).*

There are many wind farms in operation. For example, the Horse Hollow Wind Energy Center in Texas (U.S.A.), has over 400 turbines on 47,000 acres of land and can supply enough electricity for about 300,000 homes[38]. Of course the amount of electricity produced depends on how much wind is blowing at a given moment, a concern we will revisit shortly.

Advantages and challenges of wind energy

We have improved our ability to harness the wind, and it has become a practical and affordable source of energy. At the same time, though, some obstacles to wider usage persist.

Advantages. Many of the items on the list for solar are here as well: the energy is free and accessible, produces no chemical pollution and otherwise has a low environmental impact. Furthermore, it is abundant, although its supply does not exceed our demand for energy (more below) and, of course, it is all but inexhaustible.

Challenges. Wind has some problems in common with solar.

Relatively high costs historically

The required technology was expensive for decades, although costs are now about the same as those for coal. Past high prices delayed its growth, and only in recent decades has it become attractive as an alternative to fossil fuels.

It is variable and intermittent

Some areas simply receive more wind than others, and it is not constant with time, weather, or season.

Some environmental degradation

A great deal of space is needed for a wind farm. As we saw with solar, though, turbines could be placed in areas that already have been cleared (e.g., agricultural fields). In addition, spinning turbine blades in terrestrial farms could harm birds and bats, and marine systems can be disrupted by off-shore farms. Audible warning

38. U.S. Energy Information Administration, 2022

signals and appropriate location choices can minimize these problems.

Aesthetic degradation by windmills

The aesthetic degradation brought about by wind farms has raised concerns about and opposition to the construction of electricity plants powered by wind. On the other hand, some point out that a coal-powered plant, with its smokestacks and other features, is at least as objectionable.

How much energy could we get from wind? How much *will* we get?

The amount of energy wind contributes to the U.S. electricity supply has increased from 1 to over 9% since 1990, and it currently meets about 6.5% of total worldwide demand[39]. China, the European Union, the United States, and Germany have the most installed wind power, but more than 100 countries use this renewable source to varying degrees[40] Clearly, it is still a minor piece of the overall energy picture, but interest has grown as prices of wind turbines have fallen and concerns about fossil fuels have increased. Wind's role in electricity generation may increase, but there is not agreement on just how important it will become. For instance, the International Energy Agency suggests wind could generate nearly all of our electricity if offshore farms (i.e., turbines

39. Hannah Ritchie, Max Roser and Pablo Rosado. 2022. "Energy". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/energy'
40. U.S. Energy Information Administration, 2022

are located in the ocean, not on the mainland) were fully utilized[41], whereas a report by the U.S. EIA is more modest in its prediction that it will account for about 31% of our supply by the year 2050[42]. As with solar, the role of wind during the next several decades will depend on the cost and reliability of fossil fuels, public support, as well as trends in air pollution expectations and laws (more below).

10.5.3. Hydro power

Put simply, **hydro power** refers to various ways to capture the energy associated with flowing or falling water, such as that in a river channel, waterfall, or an ocean. The use of the energy of moving water is not a new idea; like wind, humans have been using hydro power for millennia in the operation of mills and other machinery. Hydro power was first used to produce electricity in the late 1880s and was the dominant renewable energy source for decades. In fact, until quite recently, it contributed more energy than solar, wind, and the others on this list combined. It was overtaken by wind recently and still lags behind fossil fuels and nuclear, though, despite its potential.

How the energy of moving water becomes electricity

The story should be familiar by now: turbines are spun, this time by water moving from high to low elevation (or by waves or changing tides), and electricity is generated as we described earlier in this

41. Wind. 2022. International Energy Agency. www.iea.org/fuels-and-technologies/wind

42. U.S. Energy Information Administration, 2022

chapter. The details of how the process is accomplished vary from place to place, with one of two general strategies used most often in **hydroelectric power plants**. In the first, turbines are placed within an unaltered body of moving water and spun by the natural current. In the second, a dam is built across the path of a river. Here, a reservoir of water is formed upstream and water is released through relatively narrow pipes in which turbines are installed. This strategy allows for the flow rate—and resultant electricity production—to be controlled and constant (Figure 10.35).

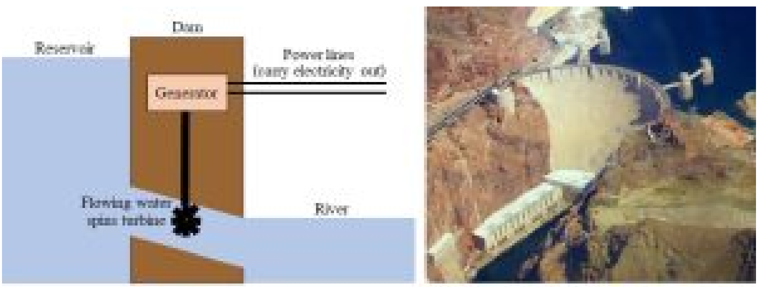


***Figure* 10.35.** *Capturing the energy of flowing water. Left (a): diagram of a reservoir created by a dam (in cross section, not to scale); water is directed through a channel to spin a turbine as it flows downstream. Right (b): photo from the downstream side of the Hoover Dam, Nevada, U.S.A.; note the reservoir at upper right. Kelsey, CC BY-NC-SA (a); Licko, Public Domain (b).*

Advantages and challenges of hydro power

Advantages. As with solar and wind, there are many compelling reasons to use the energy of water to generate electricity.

The energy is free and accessible

Water flows naturally in response to gravity, and its energy can be harnessed without the need for mining or other complex extraction techniques.

There is no chemical pollution of water

Hydro power uses moving water and does not introduce any toxic substances into reservoirs.

It has a low environmental impact

Compared to the many consequences of fossil fuels and nuclear, hydro power is relatively clean, requiring no extraction, transport, refining, combustion, or extensive waste management.

It uses proven and cost-effective technology

We have been generating electricity with hydropower for over a century and have developed practical and efficient ways to do it. In fact, when total costs are considered (construction, maintenance, fuel, and labor), hydro power is considerably cheaper than nuclear and about equal to fossil fuels.

It is consistent

Unlike solar and wind, hydro power can be captured and released at a constant rate. Among other benefits of this trait is the way it can be used in conjunction with intermittent renewable sources like solar and wind to provide uninterrupted power.

It is renewable

Although not as abundant as solar, its supply is inexhaustible under any likely scenarios.

Challenges. As you likely expect, there are some down sides of hydro power.

An initial financial investment is necessary

This problem certainly is not unique to hydro power—construction costs for a nuclear plant, for example, are far higher—but this challenge cannot be ignored for any renewable on our list.

Its feasibility varies

Simply put, hydro power is not a good option everywhere because it depends on the presence of appropriate topography (remember, downhill flow is an inherent part of the process), precipitation, and space. In the case of ocean-related power, proximity to a coastline is clearly required.

It can disrupt existing ecosystems

Dam construction brings big changes, including the flooding of large **terrestrial** areas (review Box 4.4, including Figure 4.17, for many important consequences of dams). Often overlooked is how the creation of a large, relatively stagnant reservoir on the upstream side of a dam can increase **anaerobic** decomposition and the release of methane. This is no trivial matter: levels of methane in the atmosphere are directly linked to Earth's temperature and affect climate (Chapter 14).

Dams can displace people

Flooding pushes many terrestrial organisms out of an area and fundamentally changes the character of existing ecosystems. It can also cause a great deal of human suffering as people are forced to evacuate their homes before they are permanently inundated with water. For example, the Three Gorges Dam in China (built during the 1990s and early 2000s), created the largest hydro power facility in the world by flooding about 1000 square kilometers of land[43]. Although the project provides vast amounts of electricity, it is controversial due to the ecological damage it caused and because it forced over one million people to relocate.

What is the future of hydro power?

Both the present and past suggest a long-term role for hydro power going forward. Currently, it contributes about 6% of the electricity used in the United States, second among renewables to wind (which, as we learned, is 9.2%), but it accounted for nearly all of *renewable* electricity from 1950 to the mid 1990s[44]. Some countries generate a much larger fraction of their electricity than the U.S., if a lower total amount overall, with hydro power (e.g., Norway meets

43. https://www.usgs.gov/special-topics/water-science-school/science/three-gorges-dam-worlds-largest-hydroelectric-plant

44. U.S. Energy Information Administration, 2022

over half of its demand this way). As of 2020, it was responsible for about 17% of worldwide electricity production[45].

The extent to which hydro power will meet energy demand in coming years depends on how much untapped capacity can be exploited by both improving existing facilities and constructing new ones. The story is an interesting and complicated one, in part because many dams were built for purposes other than power generation (flood control, etc.). In other words, the environmental and financial costs associated with them have already been borne, yet they produce no electricity. Although it is difficult to predict an exact amount with certainty, hydro power has the potential to contribute substantially more electricity than it does currently. Conservative interpretations of available data suggest 15% of U.S. electricity could come from this renewable resource, with more optimistic estimates ranging up to about 30%.

10.5.4. Biomass

The general term **biomass** refers to living material in plants, animals, and microorganisms and is often used by ecologists when describing **structure** and **function** in **ecosystems** (Chapter 5). It has a related meaning in our current discussion of energy: biomass fuels, often called **biofuels**, are produced or derived from present-day organisms (as opposed to ancient ones in the fossil fuels seen earlier). Of the many products that could be included in this discussion, we will take a brief look at just three: liquid ethanol, wood, and biogas. Before we proceed, a small caveat is in order. Many people categorize these energy sources as renewable because

45. International Energy Agency. 2021. Hydropower Special Market Report. www.iea.org/reports/hydropower-special-market-report

they are ultimately powered by the sun; indeed, they are also grouped like that by the author of your textbook (if reluctantly!). Arguably, though, because they are not necessarily replenished as readily and completely as are solar, wind, and hydro power, and depend on things like fertile soil for their production, this classification may not be appropriate. We will see more about this issue below.

Ethanol (or ethyl alcohol)

Ethanol is produced by fermentation, a process whereby crops such as sugar, wheat, or corn are metabolized by a type of yeast, one of the microorganisms we encountered in Chapter 3. It is familiar to many people as "alcohol", the intoxicating ingredient in beer, wine, and liquor. In the current context, though, we are interested in this substance as an energy source. To summarize its value, ethanol is a relatively clean-burning liquid added to fuel to reduce air pollution and extend our petroleum supply.

Ethanol use is higher in some countries than others, but it still only makes up a small fraction of the global energy supply. In the early 2000s, it appeared poised to grow substantially and become far more important, perhaps even replacing gasoline completely. So, what happened? The answer is complex, consisting of many parts. One of many issues is related to reduced fuel economy: since its combustion releases less energy than does that of gasoline, engine efficiency declines as ethanol content in fuel rises. A more fundamental problem is the fact that ethanol is an agricultural product. Extensive production of corn and other plants causes the many adverse consequences of farming we explored in Chapter 9 (some, like corn, are more damaging than sugar and others). Because it depends on conventional agricultural practices that are unsustainable (loss of fertile soil being among the more notable), it is arguably unfair to call ethanol renewable. Furthermore, some people hold that **photosynthesis** completely offsets the **carbon**

dioxide ethanol combustion adds to the atmosphere, but such an assertion ignores the many ways farming in general also leads to the production of this important greenhouse gas (e.g., fossil fuel use and stimulation of **aerobic decomposition**—again, review the link to Chapter 9). Finally, the opportunity-cost associated with using agricultural fields and resources to grow fuel instead of crops deserves some attention. Since Earth is closed with respect to materials, including arable land, whether we can expand biofuel production and still meet ever-increasing demand for food is not altogether clear.

Wood

If not strictly a biofuel like ethanol, wood certainly is an energy source derived from biomass. People have burned wood for heat for many thousands of years, and it continues to be the primary fuel source throughout the developing world. Even today, a fraction of home heating in the developed world is accomplished with various types of wood stoves. Trees present many advantages, including their accessibility and versatility Moreover, relative to fossil fuels, wood is readily replenished (requiring decades rather than millions of years). However, the harvesting and combustion of wood brings with them several adverse consequences (briefly presented here—we will return to forestry and deforestation in Chapter 12). First, the growth of relevant numbers of trees requires hundreds of thousands of acres of space, water, healthy soil, and nutrients in natural ecosystems or tree farms. Clearly, land and other resources dedicated to tree farming cannot be used for other purposes such as housing, industry, or food production. Second, widespread removal of trees can cause substantial damage to forest ecosystems (Chapter 12), increase rates of soil erosion (Chapter 9), and affect the hydrologic cycle by changing **runoff** and **infiltration**. Third, wood burning converts organic carbon stored in trees to inorganic carbon dioxide. In other words, instead of acting as a reservoir for the storage of C, forests become a source of CO_2 released into the atmosphere (we will see more about this point in Chapter 14). Keep in mind that many other air pollutants are released by combustion, including gases that are toxic to humans and other animals. Finally, although wood can, in principle, be grown sustainably, trees are often removed faster than they are replaced. Yes, its energy source is the sun, but whether in practice wood is truly renewable is debatable.

Biogas[46]

This is a general term that includes methane-containing products from anaerobic decomposition (Chapter 4) of organic carbon in animal waste (Chapter 9), sewage (Chapter 11), and landfills (Chapter 13). In these and other cases, methane must be separated and concentrated from a gas mixture. It can then be transported and burned to generate electricity just like the methane described in the discussion of petroleum, above. It currently makes up a small fraction of supply, but it has the potential to grow in coming years. Again, we can raise the question: is biogas really renewable? When we consider all that is required to produce it, including farming, we should be prepared to provide a nuanced answer!

46. For further reading on biogas, consult www.iea.org/reports/outlook-for-biogas-and-biomethane-prospects-for-organic-growth/an-introduction-to-biogas-and-biomethane

10.6. PRIMARY VS. SECONDARY ENERGY SOURCES

Petroleum, radioactive materials, sunlight, wind, hydro power, and biofuels are all known as **primary energy sources** because they can be processed through combustion or other means to release energy for transportation, electricity production, or temperature control. A **secondary energy source**, on the other hand, can be viewed as a sort of go-between that carries energy from a primary source to a place in which it can be used. Electricity is a good example of a secondary energy source because it is derived from something like the combustion of coal or the spinning of a turbine by moving water. The electricity itself is not a *source* of energy, rather it is a readily used form of energy that can be distributed via power lines to homes and other buildings.

Another important secondary energy source is hydrogen. This gas can store large amounts of energy, is quite portable, and can provide extremely clean power to cars and other vehicles. It also lends itself to easy transport and distribution via existing infrastructure such as pipelines. However, it has some serious limitations. Most importantly, H_2 gas is difficult to obtain in the quantities needed. Certainly a vast number of hydrogen atoms are present on Earth, but nearly all of them are bound up in either H_2O or CH_4 and must be separated from these larger molecules before they can be used. Several strategies are available, but the most widely used and feasible ones require a large energy input—that is, a great deal of primary fuel must be expended to produce this secondary fuel. For example, if an electric current is passed through water, the bonds between the hydrogens and oxygens will break; this is done easily enough, except you should note it requires electricity. As we know, renewable or non-renewable energy sources which bring on varying amounts of environmental degradation can be used, but most of our electricity comes from fossil fuels. In short, the feasibility and wisdom of using hydrogen will depend on the primary sources used to produce it.

Part IV: CLOSING WORDS AND FUTURE TRENDS

10.7. THE STATUS QUO

10.7.1. Many sources combine to make up energy supply

No single primary source currently meets all our energy demand. In fact, it is reasonable to visualize our supply as if it were a puzzle composed of several pieces. Because some of the fuels play larger roles than others, though, the sizes of those pieces are unequal. If we combine oil, natural gas, and coal into one group and renewables into another, fossil fuels clearly contribute the vast amount of energy to the U.S. and most of the industrialized world, accounting for about 80% of the total energy supply for most of the past few decades (Figure 10.36).

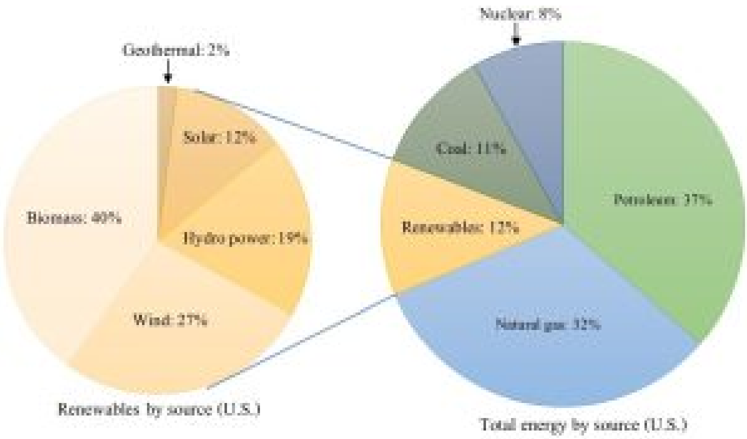


Figure 10.36. *Total energy supply by source in the U.S. (for 2021). Note the pie chart on the right contains all sources, and the chart on left shows the many renewable sources that make up the 12% wedge in the right-hand chart.* *Kelsey*, CC BY-NC-SA

10.7.2. Renewables play a relatively small role

As we have learned, solar, wind, and hydro power offer many advantages—including the fact that they will not run out in any relevant way—yet they continue to meet less of our demand than their potential suggests. You might wonder: why is this transition taking so long? The short answer is that several factors, some of which are described below, have put up various amounts of resistance through the years.

Existing infrastructure

The Industrial Revolution was powered by, and modern societies still largely run on, fossil fuels. These non-renewable fuels are simply an integral part of day-to-day existence: transportation, electricity, heating, food production, manufacturing of plastics, cosmetics, drugs, and building materials, and many other processes are enabled by oil, coal, and natural gas. To reduce our reliance on fossil fuels, many modifications to the way we live clearly would be required. Think, for example, what a transition away from gasoline-powered cars would involve. We would need new mechanisms for fuel transport, battery charging, refueling, and service. Mass transit would also likely need to take on an expanded role. A related issue: people often find it difficult to accept any kind of changes, particularly those that threaten long-standing and entrenched conventions.

Lack of government support

Historically, the U.S. government provided far more financial support to the fossil fuel industry than to various entities working on renewable energy. Thus, private petroleum companies, which are usually very profitable in any case, have enjoyed tax breaks for oil exploration and other activities whereas research that could lower costs and improve performance of photovoltaic cells and batteries received comparatively little funding. Recent developments may lead to changes in the balance, however. The Inflation Reduction Act, passed in the U.S. 2022, includes large increases in funding for solar, wind, and other technologies defined as "clean energy", and China, the European Union, and countries in other regions have made similar investments[47].

Perceived economic risks

Opponents of an intentional transition away from non-renewable

47. The International Energy Administration. 2022. Global Energy Review. iea.org

fuels to renewable ones often cite a fear of economic hardship as a concern. Although changes certainly would occur, evidence suggests that the relative growth of solar, wind, and related energy sources would not necessarily damage our economy, rather, it would likely lead to new opportunities, industries, long-term jobs, and resilience against inflation.

10.8. AN ALTERATIVE FUTURE?

Where do we go from here? Non-renewables will not be able to meet demand indefinitely, and their continued use threatens natural ecosystems along with human well being. Could renewables take on a larger share? Various models predict that the contributions they make to the world's total energy supply will rise from their current level of 12% to around 27% by 2050 (for electricity alone the number is expected to increase from 28% to 56% during the same period[48]). Could we do better? Scrutinizing the potential of solar, wind, and hydro power to provide power (data we explored in our earlier discussions) we see that far more than half of our *total energy* could come from renewable energy sources. That is, 27% could be the worst case, not the best.

Among the many benefits of a future powered by renewables: the multiple costs of dependence on fossil fuels would be substantially diminished. The proposed shift would not be simple, though. Numerous changes to infrastructure and expectations, including the three described below, would be required.

48. U.S. Energy Information Administration, 2022

10.8.1. Increased production of secondary sources from renewable primary sources

Petroleum is a highly practical fuel for transportation because it is energy rich and portable, and any replacement would need to possess the same characteristics. Both hydrogen and electricity—two commonly used secondary sources noted in section 10.6—could fill this role if they were generated in centralized locations with abundant sun, wind, or flowing water and then distributed from there to users. See Box 10.4 for a related discussion.

Box 10.4. More electric cars?

Electric cars have become more and more common, but they still make up a small proportion of vehicles on the road. Can we increase their use? The answer depends on cost and convenience, as is so often the case. Prices of many electric vehicles have started to approach those powered by gasoline, but the limitations of batteries persist to some degree. Frankly, people are hesitant to invest in a car that has an insufficient range and is difficult to recharge. The first issue is linked to our discussion of the need for better battery technology, and the situation has slowly improved during the past several years. Re-powering electric cars remains a hurdle, particularly in the United States, but the number of public recharging stations is on the rise. Unfortunately, the time required to get back on the road during a lengthy trip can be prohibitively long, though.

At this point, a second, very important, question about electric cars must be asked: *should* we increase their use? Here, the answer depends in large part on the energy source for electricity generation. Fossil fuels could certainly do the trick, but using them would hardly alleviate the problems associated with these non-renewable energy sources. Urban smog would diminish due to less petroleum combustion in cities, but coal or natural-gas burning would increase elsewhere as more electricity is needed. It is true that the amount of air pollution, notably CO_2, is lowered slightly overall by a conversion away from gasoline-powered vehicles, but it would likely worsen in the long run as demand for electric cars rises worldwide. In other words, if you plan to stick with fossil fuels as your primary energy source, electric vehicles provide minimal benefits. If, however, solar, wind, and hydro power are the sources used to generate electricity, a transition to electric cars would solve several problems.

10.8.2. Improvements and expansion

Technological improvements

We have already discussed many of the obstacles limiting the feasibility of renewables, but two merit a brief reiteration here. The feasibility of renewable sources to meet more of our energy demands would be substantially enhanced with increased efficiency and affordability. Research has led to improvements in recent years, and the work continues. In addition, variable sources like solar would be more attractive and portable with enhanced battery technology, a need we noted earlier.

Biofuel production

Production of fuels from biomass *could* be an important piece of the energy puzzle, but as we know, it is limited in large part by land use and other issues related to soil-based agriculture. Ethanol, for example, is not particularly practical if it is derived from a resource-intensive crop such as corn. However, when it is produced from sugar, other plants, or even portions of plants that are otherwise of no use as food, it could make a substantive contribution to our energy supply. Research into biofuels is ongoing, and includes work to produce fuels from giant kelp, large **algae** that do not require soil because they are grown in the ocean.

10.8.3. Living off the grid

Most of the developed world is designed around centralized power-generation plants that supply electricity to communities of people. The network of generators and power lines is assembled into the so-called grid we mentioned at the beginning of this chapter and is managed on a large scale to maximize efficiency and electricity availability for all. Individual users pay separate bills to the power company according to how much electricity they have used. In recent years, renewable energy sources have enabled more and more people to use solar or wind to generate electricity on a small scale—on their own property or rooftop, for example. Some structures are not connected to the community grid at all whereas others use electricity from a power plant as a back up only. This emerging type of infrastructure is creating an important new trend because of the freedom it gives people to choose how they obtain power. Moreover, it allows regions within the developing world that would otherwise be remote from a large power plant and its grid to obtain electricity through other means. Local access to power enables improvements in standards of living without the cost and environmental degradation associated with the construction and use of coal or nuclear plants.

THE CHAPTER ESSENCE IN BRIEF[49]

Demand for energy to power technology is both high and increasing. Non-renewable fossil fuels have dominated supply since the start of the Industrial Revolution and continue to do so despite the substantial environmental degradation they cause and the availability of many relatively clean, renewable energy sources. A move away from the status quo has started, yet change has been slow for multiple reasons.

---

49. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 10 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[50]

Besides meeting our basic biological needs, what do humans do with energy?

Why is oil classified as a non-renewable source of energy if it is formed by natural processes? Why is the sun classified as a renewable source of energy if it is not being replaced?

Imagine we discover a previously unknown oil reserve, one that could easily meet our demand for energy for 500 years. Would you embrace it? Would our problems be solved?

What variables would you consider in a risk-benefit analysis (Chapter 1) of nuclear fission and coal combustion as energy sources? Which do you think would come out on top: risk or benefit?

Some people consider ethanol to be a renewable source of energy. Given what is required to produce it, do you agree with this characterization?

50. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

Why do fossil fuels continue to meet a large majority of our energy demands when alternatives such as solar, wind, and hydropower could (mostly) take over?

11. Human Water Usage and Pollution

JASON KELSEY

In Chapter 4 we learned about the cycling of many of Earth's materials. Here we re-visit water and consider the effects of human activity on its storage and movement. Importantly, we will see how humans can reduce the availability of freshwater by using it faster than natural processes replace it and by adding pollutants to it. A review of the hydrologic cycle might be helpful as you begin to study Chapter 11.

Key concepts

After reading Chapter 11, you should understand the following:

- The different ways humans use water
- How tapping into reservoirs of the hydrologic cycle can lead to water shortages and conflicts
- Why people intentionally pollute water
- The many sources and types of water pollutants
- The strategies humans use to prevent and clean up water pollution
- How water conservation can be encouraged

Chapter contents

11.1 WATER USAGE BY HUMANS

11.1.1. Categories by use type

As is well known, we need to consume water regularly to stay alive. However, the list of ways we utilize water goes beyond just basic survival. To begin our discussion, we will group these various applications into two categories: in-stream and off-stream use.

In stream

The use of water in place, that is, inside the reservoir in which it is naturally found, is known as in-stream water use. In these cases, water is not removed from oceans, rivers, or lakes. Instead, people travel to those reservoirs and carry out their activities within them. Hydroelectric power generation (Chapter 10), boating, swimming, transportation of goods, and fishing, are good examples of this type of usage (Figure 11.1).

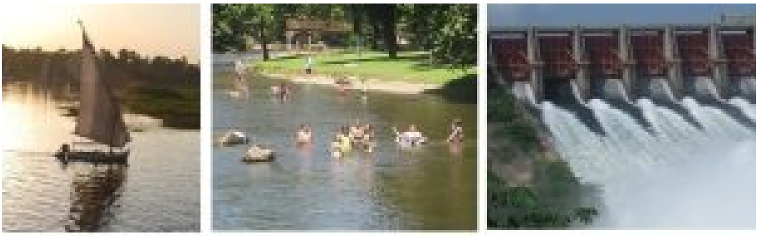


Figure 11.1. *Three examples of in-stream water use: boating, left (a); power generation, middle (b); swimming, right (c). GayleKaren, CC BY (a); Billy Hathorn, CC BY (b); ZSM, CC BY (c)*

Off stream

When water is removed from a reservoir and transported elsewhere, the activity is categorized as off-stream water use. We can further divide this group into **consumptive** and **non-consumptive.**

Consumptive. After it has been used, this water is not returned immediately to the reservoir from which it was taken. It is said to be consumed, although clearly that term is not accurate when we recall how water moves and is recycled among Earth's systems. However, since the water will potentially travel along multiple pathways, including evaporation, freezing, and other processes described previously, before returning to a specific reservoir, for all practical purposes it is gone and unavailable for some period of time. For example, people drink water from a faucet and later flush some of that water down a toilet. It is likely that the water will then begin a long journey through facilities designed to clean and ultimately release it back into natural systems. Other applications appropriately categorized here include bathing, clothes and car washing, lawn watering, irrigation of agricultural fields (Chapter 9), and some industrial processing. See Figure 11.2 for two examples.

Figure 11.2. *Two examples of consumptive use of water: irrigation, i.e., agricultural use, left (a); water in a faucet, i.e., domestic use, right (b). Jeroen Komen, CC BY (a); Nicole-Koehler, CC BY.*

Non-consumptive. At times, off-stream use is followed by a near-immediate return of the water to its source instead of a tortuous trip along the many pathways of the hydrologic cycle. For example, an industrial site might be located near a river to take advantage of the cooling power of water. Water can be withdrawn from the stream and circulated in pipes around a piece of machinery to help maintain it at a low temperature. Once the water has absorbed some heat, it is very often released back to a downstream location on the same river (Figure 11.3). The use of water in power generation (Chapter 10) is another important example of this type of water use.

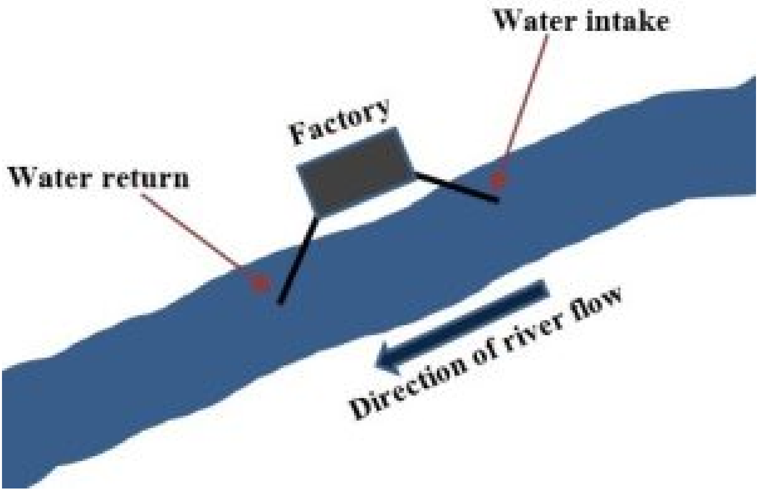


Figure 11.3. Diagram of non-consumptive use of river water in an industrial setting. Water is taken in to the plant to cool a generator or similar machine and then returned downstream. Map view. Kelsey, CC BY-NC-SA.

Power generation plants use water in this way as well (see Chapter 10 for more on power generation). It is worth noting that chemical pollution is generally not associated with the use of water to cool. However, the water can be returned to a river at a higher temperature than it was when withdrawn. See Box 11.1 to learn more about the phenomenon known as **thermal pollution**, the introduction of heat into a natural waterway.

Box 11.1. Thermal pollution

Water used to cool an industrial process (as in Figure 11.3 in the main text above) is itself heated before it is returned to its reservoir. In some cases, the change in temperature between withdrawal and release can be dramatic, with increases of 10 °C (18 °F) or more possible. Should we care about a little heated stream water? Well, even a few degrees Celsius can have a big impact on aquatic **ecosystems**. First, warmer water cannot hold as much O_2 gas as can cooler water, so **aerobic** organisms will suffer and sometimes die as a result of thermal pollution. The accumulation of organic waste from the die-off can contribute to further changes, including **eutrophication**. Second, toxic materials might behave differently at higher temperatures, endangering the health of organisms. Third, some organisms can only succeed under narrowly defined temperature conditions, so a small increase could permanently change the **structure** of downstream ecosystems.

11.1.2. Water demand outpaces supply

*Water demand varies by **sector**. **Agriculture** and **power generation** together account for about 80% of water use in the United States. **Industrial**, **domestic** (i.e., household), and other types of usage make up the rest*[1].

1. US Geological Survey. 2019. Total Water Use. https://www.usgs.gov/mission-areas/water-resources/science/

Competing demands for water

Given freshwater's scarcity and value, it is probably not difficult to imagine how a single reservoir could be coveted by people with different priorities. A body of water that provides water for domestic use, for example, can also be in demand for activities such as irrigation, boating, fishing, and swimming. Conflicts over rights to water can take many forms, ranging from hard feelings to political / economic rivalry to war. See Box 11.2 for three stories about struggles brought about by water shortages.

Box 11.2. How should we divvy up the water?

The problem of insufficient water supply is faced again and again around the world. And things will likely only get worse as populations grow. Stories of scarcity, unfulfilled demand, unpopular decision making by politicians, and unequal distribution are numerous enough to fill a large book, but we will briefly consider only three here.

1. Southern and Central California, USA: can agriculture, lawns, and drought all coexist?

The climate in these regions has been naturally dry for at least centuries, but matters have gotten progressively worse as a severe drought has persisted for most of the time between 1990 and the present day.

total-water-use?qt-science_center_objects=0#qt-science_center_objects

The past decade or so has seen especially severe depletion of already limited water reservoirs, and no end appears to be in sight. It is worth noting that a relative lack of water only became particularly relevant after the number of people living in California grew dramatically to approximately 40 million during the past century. Prior to that time, the organisms living in the area were well adapted to the prevailing environmental conditions, and water-loving plants and animals were largely absent. Currently, local water supply is simply inadequate to meet local demands. So, what can be done? One strategy has been to impose dramatic restrictions on domestic use of water. Watering of lawns and washing of driveways are both prohibited on a state-wide level, for example. Programs to encourage replacement of inefficient fixtures have been initiated as well. These new laws are somewhat controversial, however, because analogous reductions have not been required of the agricultural sector, even though it uses far more water than does the domestic sector. Further aggravating matters, California enacted several programs—largely paid for with taxpayer dollars—during the past century to redistribute water from northern and eastern regions that enjoy a surplus to central and southern farming regions that are in a deficit. These and other reasons have enabled the state to become a major producer of fruits and other crops. Even rice, a plant that requires enormous amounts of water, is cultivated in desert-like climates in the state. Whether or not public funds should be used to support private agriculture, as well as the perception that the cost of the drought is not fairly borne by urban and rural water users, continues to be a major political problem[2].

2. Aral Sea, Central Asia: what happens when rivers are diverted?

The Aral Sea was once an enormous inland freshwater lake, home to complex, rich ecosystems and a sizeable and successful fishing industry. Starting in about 1918, though, the then new Soviet government devised a plan to use water from two major rivers that fed the lake to irrigate cotton fields in other regions. Cotton production certainly increased as hoped, but the change in water distribution led to dramatic decreases in the amount of water in the Aral. The surface area of the lake decreased by approximately 50%, salt concentration increased until the water became like a typical ocean, most of the freshwater organisms disappeared, and the human residents in the area lost their livelihoods. Making matters worse, the drying water exposed the sandy bottom of the former lake, leading to toxic dust storms and increased prevalence of human diseases. Since the Soviet Union was very secretive during much of the 20th Century, outsiders had little knowledge of the gravity of the situation, and no external political pressure was applied to reverse the problem. Only after the opening and subsequent collapse of the USSR in the late 1980s was the full extent

2. ca.gov and nytimes.com, Your contribution to the California drought, Larry Buchanan Josh Keller, Haeyoun Park

of the disaster revealed. Some attempts have been made to undo the damage, but only limited progress has been made (Figure 11.4)[3].

Figure 11.4. *The shrinking Aral Sea: map view of water coverage in 1989 vs. 2014 (a); one of many former fishing boats that can be seen abandoned on the dry sea bed (b). NASA, Public Domain (a); Staecker, Public Domain (b).*

3. Would you fight for water?

Seemingly endless conflicts in and around the Middle East are obvious to almost anyone who pays even cursory attention to the news. What is less well known, though, is that a major cause of the fighting has been scarce water resources. The Arab-Israeli conflict is certainly rooted in religious and cultural differences as well as a long history of rivalries and disputes over land,

3. NASA. Earth Observatory. earthobservatory.nasa.gov/world-of-change/AralSea

but access to water is an additional source of tension. For example, control of the West Bank brings with it a good aquifer and some water security. Similarly, wars over the Golan Heights have, in part, been driven by aspirations to control the Jordan River. Relationships among Egypt, Sudan, Ethiopia, and Uganda also have been strained for many decades as these countries vie for access to the Nile River. Various agreements and treaties have helped, but war over the Nile's water is hardly out of the question.

We use water we can access

Water is recycled by the natural processes of the hydrologic cycle. It is a finite resource, though, and time is required before water will return to any reservoir. As the number of people on Earth increases, it is becoming more and more difficult to meet daily demands for water and other limited resources (Chapter 8). Additionally, there is a non-uniform distribution of water. Even as various natural pathways move it around continuously, some areas always seem to have a relatively small amount of water relative to demand for it (e.g., much of China, the Middle East, northern Africa, the Southwestern United States), whereas others appear fortunate enough to have more than enough water (e.g., parts of Canada, the Northeastern United States, much of South America). Water stored both below and above ground is tapped to meet our demands.

Groundwater reservoirs. As we saw in Chapter 4, this underground water is not readily accessible: wells must be constructed to bring it to the surface. The challenges do not end with just accessing the water, though. Aquifers, no matter how large, are finite. In the case of groundwater, **overdraft** occurs when withdrawal through wells removes water faster than infiltration can replace it. Chronic overdrafting in some areas has led to long-term lowering (called depression) of water tables and increased likelihood that wells will run dry. The resulting shape of the water table around a well (in 360°) is called a **cone of depression**, and the resulting depth to the water table is referred to as **drawdown**. These are important considerations because they affect well design. Keep in mind that an imbalance between rates of input (i.e., **recharge**) and output (i.e., pumping) could be caused by many factors, including natural phenomena like drought, as well as anthropogenic activities that place excessive demands on water. Figure 11.5 shows a model of an aquifer after pumping from it has begun (compare it to Figure 4.16).

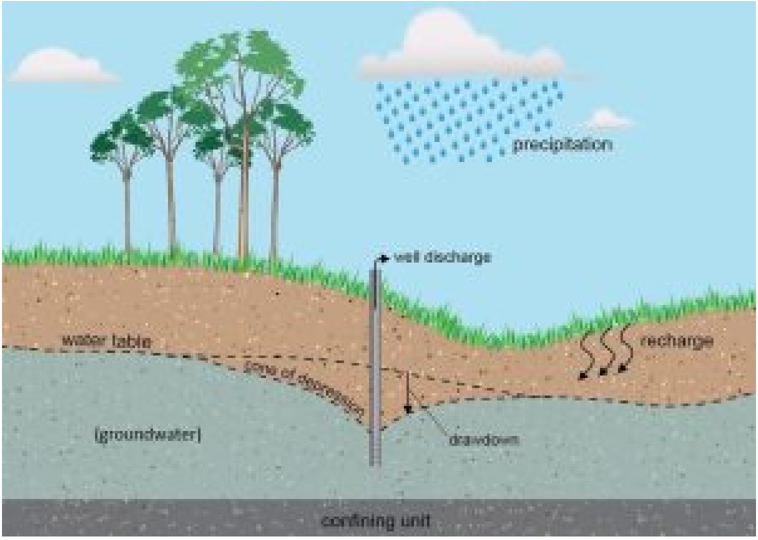


Figure 11.5. *Extraction from the groundwater reservoir changes the shape of the water table. Note the cone-shaped depression centered around the well. In cross section. USGS, Public Domain.*

Surface freshwater reservoirs. Rivers and lakes are often used as sources of water for agriculture, domestic use, and industry, particularly when groundwater is either limited or unavailable. As long as rates of water inputs to these reservoirs are equal to rates of outputs from these reservoirs, demand for water will be met. If water is removed faster than it is replaced, however, the reservoirs will become depleted, aquatic ecosystems will be damaged, and restrictions may be placed on the ways humans are allowed to use water. The water in surface reservoirs is also relatively vulnerable to pollution and must be cleaned sufficiently before it can be used (described later).

11.2 WATER POLLUTION

Clean water is indispensable to our survival, yet we have been using rivers, lakes, oceans, and other reservoirs as repositories for our unwanted waste for centuries. Why would we do this? The reasons broadly include ignorance, expedience, and convenience. History is filled with suffering brought about by seemingly illogical disregard for the importance of water, including the devastating outbreak of cholera described in Chapter 3 (Box 3.6). In that famous case, ignorance was largely to blame. But even now we try to justify water pollution with two, largely obsolete, reasons. First, reservoirs are viewed as infinitely capable of absorbing and diluting contaminants to concentrations too low to be of concern. The widely held belief captured in the phrase **dilution is the solution to pollution** was somewhat reasonable prior to the rapid growth of the human population during the past 100 years (Chapter 8) because the density of pollution sources was relatively low. Today, though, few bodies of water can absorb all the poisons people release. Second, since flowing water visibly carries toxic substances away, we can convince ourselves that the problem has been solved by dumping into a river. Appropriately summed up as **out of sight, out of mind**, this misconception often guides human responses to waste management, even if nearly everybody in the modern world is downstream of a pollution source sooner or later.

11.2.1. Sources of pollution

Water quality, essentially the suitability of a volume of water for its intended use, can be degraded by many human practices, including agriculture (Chapter 9), mining and power generation (Chapter 10), as well as gardening and other household activities, industry, healthcare, recreation (e.g., golf courses, other sports fields), and the military. Natural phenomena can also pollute water (more below). Their wide diversity has led to the creation of general classes of pollution sources based on some shared characteristics.

Point sources of pollution

These are defined as **discrete** (i.e., singular or finite—*not* the same meaning as "discreet") outlets from which pollutants enter a waterway. **Point sources** include pipes running from a factory or sewage treatment facility (more below) and leaking underground tanks that store fuels or solvents (Figure 11.6).

Figure 11.6. *Two point sources of water pollution: a pipe releasing sewage into a river, left (a); a diagram of a leaking underground storage tank releasing fuel that could contaminate groundwater, right (b). Right diagram in cross section. John Collins, CC BY-SA (a); Kelsey, CC BY-NC-SA (b).*

They release materials at relatively high concentration into water, and then those pollutants spread out away from their point of entry (i.e., as they move downstream) in what is known as a **pollutant plume** (Figure 11.7).

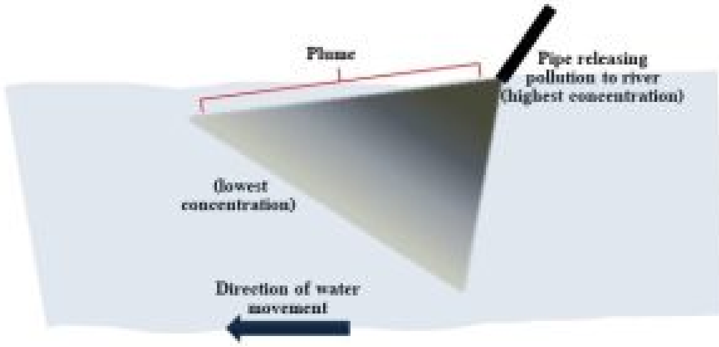


Figure 11.7. *An idealized sketch of a pollutant plume in a river. The concentration is highest at the source (the mouth of the pipe) and declines as the pollutant moves downstream. Map view. Kelsey*, CC BY-NC-SA.

These types of sources are regulated, not prohibited, in the United States by both federal and local laws. For example, a National Pollutant Discharge Elimination System (NPDES) permit allows the release of certain materials into natural waters from point sources provided the guidelines regarding how much and how frequently waste is dumped are followed. Given the nature of point sources as singular outlets, regular monitoring of them can be readily accomplished—material flowing from a pipe, say, can be directly analyzed. Whether or not a given entity is complying with water quality laws is then assessed.

Non-point sources of pollution (NPS)

These are diffuse and non-discrete outlets. From a legal standpoint, they are simply defined as anything that is *not* a point source.

Rather than flowing out of a single pipe or other structure, pollutants originate from numerous and hard-to-quantify sources such as agricultural fields, cities, and suburban developments. Recall that some fraction of the precipitation striking the Earth moves as runoff (Chapter 4). When this water flows across land that has been contaminated with pesticides or fertilizers (Chapter 9), buried garbage (Chapter 13), motor vehicle fluids from urban areas, or household chemicals and lawn fertilizers from neighborhoods, it can pick up some of those pollutants and carry them downhill to both above- and, potentially, below-ground reservoirs (Figure 11.8).

Figure 11.8. *Urban runoff is one non-point source of pollution: a city storm drain collects materials from a wide area and sends them all to a nearby natural body of water. Robert Lawton, CC BY.*

Sometimes overlooked is the way combustion contributes to non-point-source pollution. The burning of fuels, garbage, and other materials can produce air pollution, some of which mixes with water in the atmosphere and falls to the surface with precipitation. Since runoff can travel some distance before it intersects with a body of water, it has the potential to encounter a diversity of pollutants from many different sources as it moves. Although it is possible to analyze the composition of a natural body of water, it is exceedingly difficult to identify non-point sources of any unwanted compounds found in it. We could, for example, quantify the amount of a certain pesticide (Chapter 9) present in a hypothetical river, but we would have no simple way to determine which of the six farms in a 1000-acre upstream area is most responsible for releasing it. Similarly, water on city streets can carry gasoline and radiator fluid into storm drains, and subsequently to a river or ocean, but it would be nearly impossible to ascertain which of the 2,000,000 cars driving around that area are actually leaking those pollutants (see Figure 11.9 for a diagram of non-point sources of pollution).

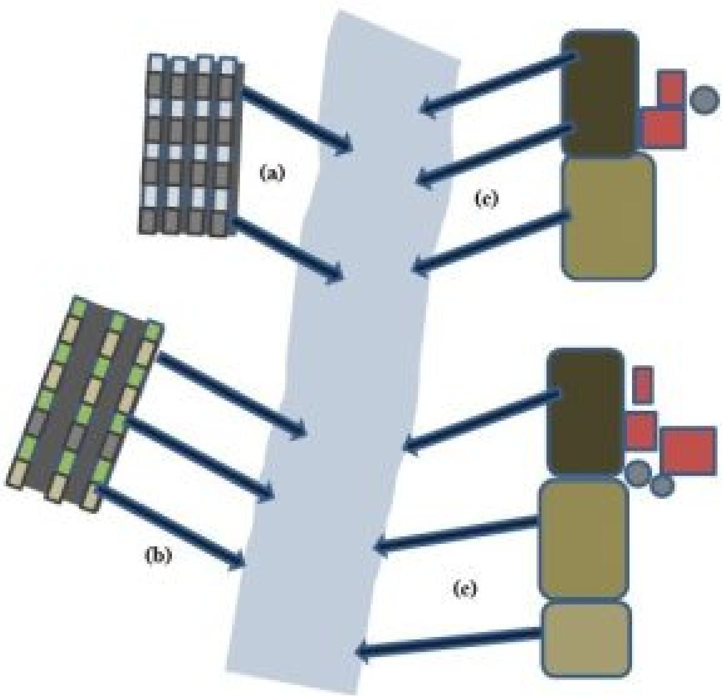


Figure 11.9. *An idealized sketch of non-point-source pollution to a river: urban runoff* (a); *suburban neighborhood runoff* (b); *agricultural runoff* (c). *Map view. Kelsey*, CC BY-NC-SA.

In short, regulation of water pollution is far more complex for non-point than for point sources. Education of polluters, economic incentives, and other means are generally employed about as often as is enforcement. In addition, the kinds of chemicals available for use, such as which pesticides farmers are allowed to purchase, can help control NPS pollution. Perhaps the most notable instance of this approach was seen in 1972 when DDT, a widely used insecticide (Chapter 1), was banned for sale in the United States. That step had a profound effect on environmental systems, as we will see in Chapter 15.

Non-human sources

Natural sources are important sources of harmful substances. For example, particles launched into the atmosphere by volcanic eruptions and spray caused by precipitation striking surfaces of soil and rocks can be deposited into waterways. Additionally, as water infiltrates soil, it can dislodge metals and other natural toxins and carry them downward toward the groundwater reservoir. Note that source of pollution does not affect the regulation of water quality: to be fit for a particular use such as drinking, water must still meet the standards for biological, chemical, and physical contaminants (next section).

11.2.2. Types of pollutants

A great diversity of biological, chemical, and physical agents can degrade water quality. In response, the United States Environmental Protection Agency (EPA) maintains a list of approximately 100 regulated substances, those for which maximum allowable limits in drinking water have been established (varying approaches are used around the world). As we will see below, these are broadly grouped as biological, chemical, or physical pollutants. You should realize that the EPA also tracks many other materials not on the list that are suspected to pose risk but for which legal standards have not been established. Water intended for uses other than human consumption is subject to similar kinds of regulation (see Box 11.3 for a brief discussion of this topic).

Box 11.3. A sliding scale for contamination

How do we determine if water is safe? The answer depends on what we intend to do with that water. Consider that streams, swimming pools, car washes, and other settings will be held to lower standards than water slated for domestic use. The number and diversity of potential pollutants and their allowable levels is enormously high: metals, pesticides, organisms, and physical materials are all of concern. In any case, you should realize that for many reasons it is next to impossible to purify water. Even drinking water has some low levels of toxins present in it.

Biological agents

Unless it has been sterilized by boiling or other means, water will contain some number of live microbes in it. For the most part, these organisms are not harmful. However, some **bacteria**, **protozoa**, and **viruses** cause diseases, and their presence in drinking water is concerning. A small sampling of illnesses that can be transmitted through polluted water includes dysentery, cholera, polio, and hepatitis. In the U.S., federal and local laws provide strict guidelines governing the numbers of pathogens allowable in drinking water because of the potential threat they pose to human health.

Chemical agents

Many chemical compounds are of concern to environmental scientists and governmental regulators. A list of potential contaminants includes pesticides and fertilizers, solvents, fuels, cleaning agents, drugs, and metals. These compounds come from a variety of sources and can bring about adverse effects to human and environmental health. Note that, although it is not always viewed as toxic, various types of salt can also cause serious water pollution issues (more below).

Physical agents

Some pollutants are best described as physical entities because they are not living, and the threat they pose is not necessarily linked to their chemical properties. For example, floating trash, sediments, and other agents that are not dissolved in water are categorized here. They are problematic for several reasons, including their ability to block sunlight, directly damage sensitive organisms, and render drinking water unpalatable (and possibly toxic).

11.2.3. Surface water pollution

Rivers, lakes, and oceans are particularly vulnerable to contamination because they tend to have large, unprotected surfaces; pollutants carried in precipitation and runoff (Chapter 4) can flow into them freely. Salts also can accumulate in these bodies due to several processes. Such reservoirs can only be used as intended if the pollutants in them are removed, a task that can be very difficult to carry out (see the section near the end of this chapter for more about strategies used to clean up contaminated surface water).

11.2.4. Groundwater pollution

Leaching from the surface

If you recall that approximately 98% of Earth's fresh, unfrozen water is found beneath the surface (Chapter 4), it should be clear why pollution of this reservoir is of great concern to environmental scientists. Like the surface waterways described in the previous paragraph, groundwater can be contaminated from many sources. However, unlike rivers, lakes, and oceans, it is somewhat less vulnerable because of its position underground. Anything released at the surface must travel vertically downward with flowing water (i.e., infiltration), a process known as **leaching**, through the unsaturated zone before it can reach the saturated zone (Chapter 4). During leaching, a pollutant may undergo chemical changes that render it non-toxic (**degradation**) (more in Chapter 15), or it could interact with the soil and other materials through which it moves in such a way that it is slowed or even stopped before it travels very far. A number of factors influence the extent to which materials are able to reach the water table. Depth is the first important consideration. As a rule, shallow groundwater is more likely to be polluted than is deep because there is less time to remove or alter the chemical before it gets to the saturated zone (Figure 11.10).

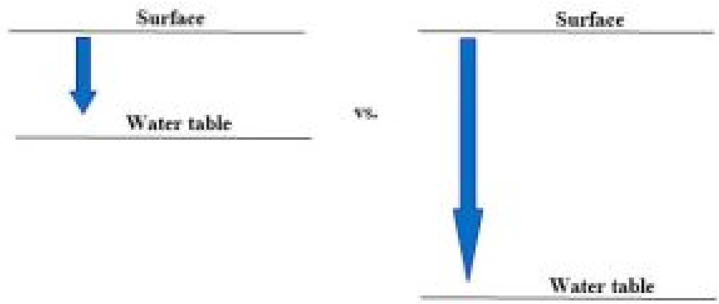


Figure 11.10. *A simplified diagram of the effect of depth on the likelihood of groundwater contamination (in cross section). Pollution moving from the Earth's surface is more likely to affect shallower reservoirs (a), than deeper ones (b). Kelsey, CC BY-NC-SA.*

Secondly, environmental properties can control the success or failure of **microorganisms** living in soil that might be able to **decompose** some pollutants. If conditions are too harsh to allow their growth, bacteria that might be capable of eating a pollutant in principle will not be able to actually do so in practice. Furthermore, given their different properties, soils vary in their capacity to bind and slow compounds flowing in water. A final variable is related to the chemical properties of the pollutant in question. Simply put, some compounds are intrinsically difficult to break down, whereas others are unstable and can be degraded easily.

Saltwater intrusion

An aquifer near a body of saltwater, such as one in a coastal region, can become contaminated if wells constructed in it withdraw water faster from the saturated zone than can be replaced by recharge (i.e., overdraft). In this case, salty water below the bottom of the ocean is drawn through the subsurface into the zone that was solely occupied by freshwater before the presence of wells (Fig. 11.11).

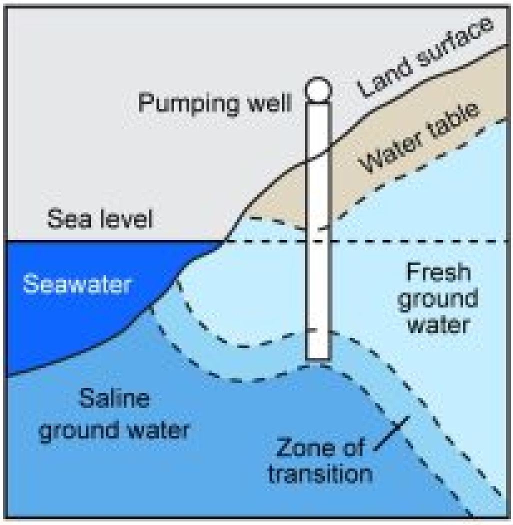


Figure 11.11. *Diagram of saltwater intrusion. Pumping water to the surface draws saltwater from an adjacent ocean into a well. US EPA, Public Domain.*

Saltwater intrusion often renders an aquifer unfit to meet demand and necessitates the importation of water from an external source. For example, a network of aqueducts carries water as much as 200 km from places like the Catskill Mountains in New York State (U.S.A.) to New York City in response to saltwater intrusion and other problems with the water near the city. We will see more about salt pollution near the end of this chapter.

> *What might be considered a pollutant in one place might not be defined as such in another. Put another way, "pollutants" could be used to refer to* ***resources out of place****. Gasoline, for instance, is called a pollutant if it*

leaves a car's fuel tank and ends up in a natural system such as a river; otherwise, it is given a more neutral name such as "fuel".

11.2.5. Humans can reduce the amount of pollutants in waterways

Speaking broadly, the amount of waste in natural systems, including reservoirs of the hydrologic cycle, can be minimized in two ways: control of releases before they occur, and clean up, known as **remediation**, of contaminated systems after pollutants have entered them. Although prevention is often less costly than remediation in the long run, it is not always possible to be proactive. Many problems were created years ago, so clean-up is the only option available to us.

Wastewater management: an overview

Industry, agriculture, healthcare, and sewage are among the sources of what is generally known as **wastewater**, that is, water-based outputs containing potentially harmful products. In simple terms, we clean, or **treat**, wastewater to remove materials that render water unfit for whatever use we intend for it (review Box 11.3). Here we will explore sewage treatment, just one example of this type of remediation.

Sewage treatment

Human sewage is not the topic of conversation at many dinner parties. It is probably fair to say that few people ponder, or frankly even wish to ponder, what happens to the water they flush down a toilet or drain from a shower. Like it or not, though, the excrement and other waste generated by each of us contributes to a continuous and enormous problem that must be addressed if we want our water to be safe for the various demands we place on it: consider that about 360 billion cubic meters (95,000 billion gallons) of wastewater is produced worldwide each year[4]. It varies by country, as you likely expect, but it is safe to ay that those toilets never stop flushing! Additionally, clean-up is enormously challenging because sewage contains so many kinds of pollutants. Multiple items from each of the three categories we established earlier—biological, chemical, and physical—are present and pose substantial risks to water quality. The list of materials that worry us includes far more than human feces and urine. Paper, drugs, plastic, latex, cleaning products, cotton swabs, pets, and anything else a person might flush down a toilet must be removed from the water. Even currency and other valuables occasionally end up in the waste stream (a $20 bill found by a sewage plant worker is still worth the same as one in your wallet, although clearly neither is fit to be put into your mouth). See Table 11.1 for a list of some of the most important contaminants found in sewage along with example adverse effects caused by them.

4. United Nations University. Institute for Water, Environment, and Health. 2019. https://inweh.unu.edu/half-of-global-wastewater-treated-rates-in-developing-countries-still-lagging/#:~:text=%E2%80%9CGlobally%2C%20about%20359%20billion%20cubic,water%20is%20currently%20released%20untreated.

Table 11.1. Example water pollutants and their potential adverse effects.[5]

Contaminant	Category	Example adverse effects
Microbial pathogens (Ch. 3)	Biological	Dysentery, cholera, typhoid fever, hepatitis, polio, meningitis, giardiasis
Pathogenic worms	Biological	Diarrhea, malnutrition
Antibiotics	Chemical (drugs)	Antibiotic-resistant bacteria (Ch. 3)
Hormones, steroids	Chemical (drugs)	Disruption of animal reproduction, possible human growth effects
Pesticides (Ch. 9)	Chemical (agriculture)	Cancer, reproductive effects
Fertilizers (Ch. 9)	Chemical (agriculture)	Oxygen deprivation to the brain
Cleaners, soaps, shampoos	Chemical (domestic)	Damage to ecosystems
Metals	Chemical (industry)	Birth defects, cognitive deficits
Solvents, fuels	Chemical (industry, transportation)	Cancer, respiratory damage
Sediment	Physical	Limits on primary production (Ch.4)
Paper, plastic, Styrofoam, wood	Physical	Various disruptions to ecosystems

Given this complex composition, it is impossible for just one

5. Most of this information comes from US Environmental Protection Agency. 2009, accessed 2023. National Primary Drinking Water Regulations. epa.gov

strategy or method to remove all the pollutants of concern; many steps are typically required to clean sewage. In addition, not all homes, neighborhoods, and cities use the same approach to sewage treatment. Put succinctly, it is a very complicated and extremely important environmental problem. Three methods commonly used in the United States are summarized here.

Backyard septic systems. Homes in areas that are relatively sparsely populated tend to employ this strategy to treat sewage (wastewater from roughly 20% of homes in the United States is treated via septic systems[6]). Waste travels from a house to a small, buried septic tank (dimensions on the order of a couple meters) where it is held long enough for large objects to sink while liquids rise. Some of the solids are partially digested by **microorganisms** that are adapted to eat such things, but some are stored and must be pumped out periodically. The contaminated liquid is slowly released from the tank into a drain or **leach field**, a flat underground pit filled with gravel, soil, and other porous materials. Here solids are physically separated from the flowing water, and additional microbial activity breaks down more of the digestible waste products. Figure 11.12 provides an idealized diagram of a typical septic system.

6. *U.S. Environmental Protection Agency. 2023. The Sources and Solutions: Wastewater. https://www.epa.gov/nutrientpollution/sources-and-solutions-wastewater*

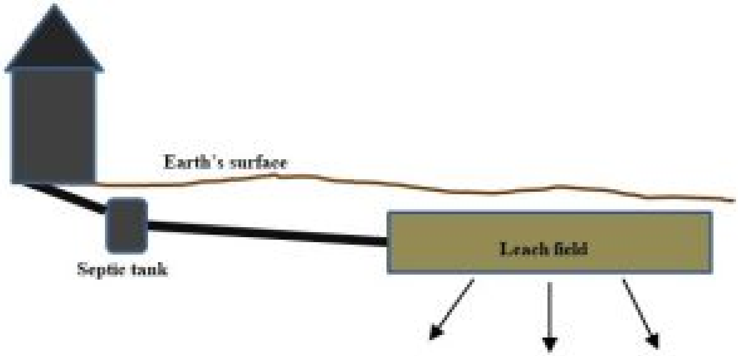


Figure 11.12. *Idealized diagram of a septic system. Sewage from the house moves to a buried tank where it is partially digested. Liquid from the tank moves to a buried leach field. In cross section, not to scale. Kelsey, CC* BY-NC-SA.

Note that these localized treatment facilities must be carefully designed and installed to ensure that surface and underground water reservoirs are not contaminated and that toxins do not come into contact with humans or other organisms. Among many considerations, sufficient space must be available to make them work as intended. In urban and other densely populated areas, septic systems are inadequate to safely treat sewage, so one of following two strategies must be employed instead.

Municipal sewage treatment plants. Where small, individual septic tanks are impractical due to population density, such as in a city, sewage is collected from homes and sent via sewer pipes to a centralized **sewage treatment plant.** This is the strategy used by about 80% of U.S. homes to manage 34 billion gallons (approximately 129 billion liters) of sewage water *every day*[7]. The

7. U.S. Environmental Protection Agency. 2023. The Sources and

specific details of how wastewater is treated vary from place to place, but these facilities tend broadly to use the same strategy. Generally, the mixture of materials described above goes through multiple steps to remove harmful materials and yield water clean enough to be released into a natural reservoir (see Figures 11.14 and 11.15, located below the descriptions of the many steps, for a diagram and a photo of a sewage treatment plant).

1. Preliminary treatment

The first step at a municipal plant is to remove large, solid materials by passing raw sewage through a screen. The result is an eclectic mound of trash that is usually dried and buried in a landfill (Chapter 13). The water that has been separated from the trash passes to the next stage. Gravel and other solids also may be removed.

2. Primary treatment

The goal of this step is like that of the screening step: further separation of solids from liquids. Liquid passing preliminary treatment is collected in **settling tanks** where heavy materials sink and liquids rise to the top. These tanks are large, at least 15 meters wide and 2 meters deep, and accomplish what occurs in the smaller septic tanks described above. The resulting liquid is still not ready to be released into a natural waterway, but it is much clearer than the initial sewage. It is sent on to the next step of treatment, described below. Some of the solids are **inorganic grit**, like gravel, and cannot be further altered. Many of the solids are **organic**, though, and are collected in tanks where **anaerobic** microorganisms can partially break them down. The gaseous products of this digestion, known as biogas, can be collected, processed into methane, and burned to generate electricity (Chapter 10). The

Solutions: Wastewater. https://www.epa.gov/nutrientpollution/sources-and-solutions-wastewater undigested solid products may be dried and applied as fertilizer to farmlands (Chapter 9).

3. Secondary treatment

The water left after primary treatment is put into contact with specialized bacteria capable of digesting many of the materials still present. As a result, a substantial reduction in the amount of organic pollutants present in the sewage is accomplished. Lower levels of organic compounds decrease demand for oxygen by aerobic microorganisms in the water because their food source has diminished in size. The logic may seem a little convoluted, but less food in the water means less demand for dioxygen and therefore higher amounts of dissolved O_2 remaining in the water. The effectiveness of secondary treatment can be assessed through the measurement of a critical property known as **biochemical oxygen demand (BOD)**; if BOD is too high, the water is deemed to be polluted.

The now-cleaner liquid moves on to additional treatment (step 4, below, may occur, or the liquid will move directly to step 5), and the solids generated in this step (dead microbial cells, for example) are sent to the same tanks used to digest the materials left after primary treatment (step 2, above). Secondary treatment is only as effective as the bacteria responsible for it, so a great deal of effort is exerted to maintain optimal environmental conditions for these organisms to grow and thrive.

4. Advanced treatment

Additional processing may, in rare cases, target remaining pollutants. The need for these steps varies from place to place, but generally they are designed to remove problematic nitrogen-containing compounds (Chapter 4), pesticides (Chapter 9), or other pollutants present in unacceptably high concentrations.

5. Disinfection

The number of disease-causing microorganisms present in the sewage is reduced through one of several specific mechanisms. Chlorination of the water is the most commonly employed technique because it is typically the least expensive option and it effectively destroys many bacteria. There are some disadvantages, though, including the facts that some important **pathogenic protozoa** are resistant to chlorine compounds, and remaining chlorine can lead to the production of potentially harmful compounds. Ozone gas (Chapter 14) or ultraviolet light can be used instead of chlorine, but they have limitations as well. By whatever method, though, disinfection is critical to reducing the risk of water-borne diseases.

6. Release of treated water

Finally, the treated water can be returned to a natural reservoir through a pipe, known as the **sewage outfall**, that exits the facility (see Figure 11.13).

Figure 11.13. *A sewage outfall on the Thames River, London (England). Treated water is released from the pipe at the bottom left of the photo into the river. Nigel Cox, CC BY-SA.*

This released liquid, known as **effluent**, is tested periodically to be sure it meets the requirements for levels of physical, chemical, and biological pollutants. Intended use will of course dictate just how strict the standards will be. Effluent released into a stream must be clean enough for swimming and fishing; notably, BOD must be low enough to ensure that fish and other aerobic organisms will have sufficient oxygen to survive. In some cases, inadequately treated sewage containing excess amounts of digestible organic material (and therefore high demand for oxygen by aerobic microorganisms) has led to fish kills in receiving waterways. Note that water destined for human consumption must pass through a separate **drinking water treatment plant**, where additional settling, filtration, and chemical additions can be used to make it potable.

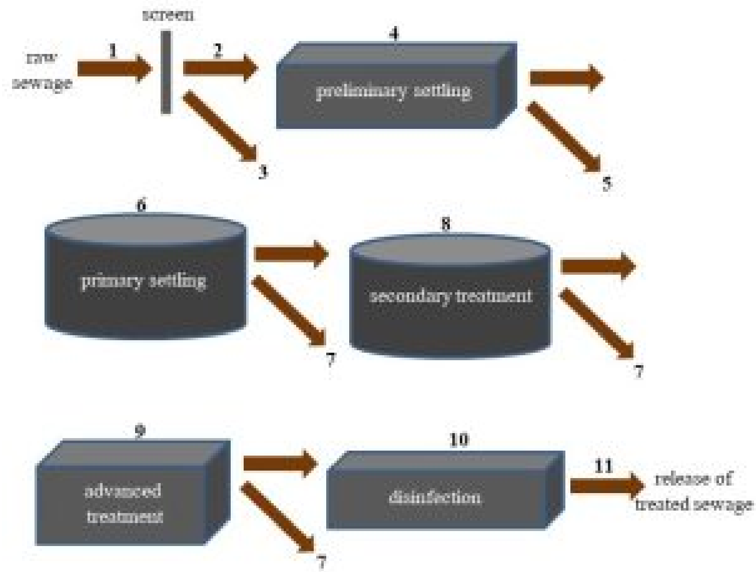


Figure 11.14. *An idealized diagram of a typical sewage treatment plant. Side view.* **1.** *Raw sewage is collected.* **2.** *Raw sewage is passed through a screen to collect large solid materials.* **3.** *Solids caught on the screen in step 2 (mostly trash) are dried and sent to a landfill.* **4.** *The mix of solids and liquids passing through the screen in step 2 may move to a preliminary settling chamber. Here, heavy grit (e.g., gravel) is collected.* **5.** *Grit at the bottom of the chamber in step 4 is sent to a landfill or experiences another fate.* **6.** *Remaining solids are allowed to settle in primary treatment tanks.* **7.** *The solids remaining after step 6 are sent to a tank in which organic materials are digested by microorganisms (the tank is not shown in the figure; solids left after steps 6, 8, and 9 end up in this same tank).* **8.** *Organic materials still in the liquid sewage after step 6 are digested by bacteria in secondary treatment. Leftover solids, including dead bacterial cells, move to the digester described in step 7.* **9.** *Additional pollutants that are still present may be removed as needed in advanced treatment. The solids from this step also move to the digester described in step 7.* **10.** *Potentially pathogenic bacteria are killed by disinfection.* **11.** *Liquid water is released to a natural body of water (e.g., river, ocean, lake). Kelsey, CC BY-NC-SA.*

Figure 11.15. *Photo of a typical municipal sewage treatment facility. Each of those settling tanks is approximately 15 meters in diameter and 3 meters deep. Everything is connected via underground pipes. Lewis Clarke,* CC BY.

Before moving forward, we should note that public facilities generally are effective in the removal of physical pollutants (steps 1 and 2, above), reduction of BOD (step 3), and killing of most pathogenic bacteria (step 5). Typically, though, wastewater treatment plants have little or no capacity to break down or remove industrial chemicals. Instead, problematic compounds are treated, by law, prior to their release into public sewage. The specifics of such Industrial Pre-treatment Programs (IPTP) vary and depend on the nature of the chemicals of concern. Broadly, commercial entities that produce toxic waste products must follow permitting rules established by the United States Environmental Protection Agency before releasing such chemicals into municipal waste streams.[8].

8. Readers interested in learning more about IPTP are encouraged

Other approaches. Septic systems and municipal facilities dominate sewage treatment in the United States. However, other options are used in a small number of cases. A constructed wetland is one example of a system that can accomplish many of the same goals as conventional treatment (Figure 11.16).

Figure 11.16. *Artificially created wetlands like this one in Poland can be effective ways to treat human sewage. Forest and Kim Starr, CC BY.*

Human waste can be piped to a relatively large and open area that has been modified to resemble, at least superficially, a swamp. Through the action of their grasses, sediments, and microorganisms, these artificial ecosystems can filter and retain toxic substances as well as facilitate decomposition of organic materials. Constructed wetlands are by no means equivalent to natural wetlands, but they offer an alternative strategy to wastewater management. Smaller-scale experimental systems that clean wastewater from a single building by using aquatic plants have been developed as well, although they too are rare.

Salt removal

Earlier in this chapter we saw that salts can accumulate in water and render it unfit for its intended use. Drinking water standards are particularly rigorous, and even a small amount of dissolved salt can leave the affected water **unpotable**. What might seem like a minor problem is actually quite troublesome: consumption of

to consult https://www.epa.gov/npdes/pretreatment-standards-and-requirements-applicability excessive amounts of saltwater can be deadly to humans. Additionally, as freshwater shortages become increasingly common in modern times, more and more people look to the world's oceans for suitable drinking water. The obvious question is: can this water be made safe to consume? Although the answer is yes, the problem is not as straightforward as we might hope. **Desalination**, or the removal of salt from water, can be readily accomplished through a process such as distillation or reverse osmosis, but a large amount of energy (and, therefore, money) is required to do it on a large enough scale to be relevant for communities of people. Some wealthy nations with no other realistic options invest enormous amounts of money in desalination, but it is generally not their first choice.

11.3 EXTENDING OUR SUPPLY OF USEABLE WATER

To end our discussion of water, we will briefly consider some of the ways people can reduce usage, conserve water, and increase availability of freshwater to help meet both present-day and future demand.

11.3.1. Improve efficiency

All three sectors can take steps to reduce the amount of water they use without compromising the integrity of the activities they perform. Since it is the biggest user of water, increased efficiency in agriculture can have a substantial impact. More careful and targeted irrigation as well as cultivation of crops in appropriate climates are among the many strategies farmers can employ to bring about savings. Industry can use less water through better management of processes as well as recycling of water in various ways. Usage by the domestic sector can be reduced by switching to bathroom fixtures, laundry machines, dishwashers, and other appliances that require low amounts of water to operate. Changes to simple habits associated with faucet use (e.g., do not leave them running unless absolutely needed) can also conserve water. Toilet design is but one example worth some consideration. Noting that flushing required far more water than was usually required, the United States Federal Government created new efficiency standards as part of the Energy Policy Act (1992). In short, household toilets were limited to 6 liters (1.6 gallons) per flush, less than half of the volume used prior to that new law[9]. Despite some early opposition to the regulations, as well as some (perhaps predictable) technological challenges, these new toilets have been in use for much of the past three decades. More efficient fixtures are available, including some no-water toilets, but they are not mandated by federal law. Remember, though, that since it uses far less water than the other two sectors, only modest overall savings can be realized though reductions in household use.

11.3.2. Raise the price of water

Many people concerned about water conservation argue that the cost of water is simply too low to effectively discourage wasteful use. According to a report published by the United States Environmental Protection Agency, water is cheaper in the U.S. than it is in other industrialized countries, giving the inaccurate

9. US Environmental Protection Agency (EPA). 2017. Water Efficiency Management Guide Bathroom Suite. EPA 832-F-17-016d. epa.gov

impression that water is more abundant than it really is. You are encouraged to consult the EPA website[10] for more information.

THE CHAPTER ESSENCE IN BRIEF [11]

Multiple and varied human demands affect the availability of clean, fresh water; as the human population grows, scarcity of this crucial limited resource will only increase. Strategies to protect our water supply include restrictions on usage and releases of pollutants, as well as treatment of wastewater.

10. US EPA. 2017. Pricing and Affordability of Water Services. https://www.epa.gov/sustainable-water-infrastructure/pricing-and-affordability-water NOTE: Various states have enacted similar measures that include higher pricing to encourage conservation by consumers.

11. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 11 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[12]

How can some types of water usage be considered consumptive if Earth is closed with respect to materials?

Why are environmental scientists concerned about water supply and quality on a planet that has such abundant water?

In your view, which sectors should be given priority access to limited freshwater supplies?

Is dilution ever a solution to pollution?

What about the content of sewage makes it particularly difficult to treat?

If it were up to you, what measures would you put in place to extend our supply of freshwater?

12. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

12. Management of Living Resources

JASON KELSEY

We have explored human use of *non-living* resources such as water, soil (ignoring that it contains many living components!), and fossil fuels in preceding chapters. Among other important ideas we noted, recall that demand for these coveted materials increases as the number of people and standards of living go up. Our focus now turns to two examples of ways *living* resources are exploited and managed by humans as well as some important consequences that come from these activities. Keep in mind that environmental scientists seek to protect more than the two specific types described here, but we will explore just forests and fisheries in Chapter 12 because they provide important examples of the kinds of challenges faced in balancing human demands against those of natural systems.

Key concepts

After reading Chapter 12, you should understand the following:

- The many services provided by natural forests
- The causes and consequences of deforestation

- The strategies used to protect forests
- How fisheries are defined and why they are highly valued
- Overfishing and other threats to the health of fisheries
- Steps that can be taken to protect fisheries
- The costs and benefits of aquaculture
- The roles played by governmental agencies in the protection of living resources such as forests and fisheries

Chapter Contents

12.1 FOREST MANAGEMENT

12.1.1. What is a forest?

Simply put, **forests** are ecosystems dominated by trees, although they can vary considerably in the organisms that inhabit them. As we have learned in Chapter 5 and elsewhere, the identity of the species found in an area depends on environmental factors such as temperature, quantity of precipitation, soil properties, and sunlight availability. **Successional stage** (Chapter 5) plays a critical role as well. Besides the trees, other appropriately adapted organisms (animals, microorganisms, plants) live within forests, employing many strategies and interactions to survive.

12.1.2. Why are forests important?

Forests provide many benefits

Here we look at some of the reasons forests are valued and why so much effort is put into their conservation.

Ecosystem services. This term refers to the many benefits forests provide by their continued existence and normal functioning.

Habitat

As noted above, forests are types of ecosystems in which many organisms live together. Clearly trees play a critical role in the health of these systems and, without them, important animals (birds, mammals, insects and so forth) would struggle to survive. Since humans are interested in protecting those *other* organisms, they are motivated to maintain forests.

Resistance to erosion

Forest soils stay in place because of the presence of trees and other plants for two important reasons. First, the branches and leaves of the trees form a physical barrier, or canopy (see Chapter 5), that blocks or **intercepts** precipitation before it strikes soil; consequently, the damage done by direct impacts of raindrops is greatly reduced. Second, the underground networks of roots provide stability and limit topsoil loss in ways we examined during the discussion of land clearing and agriculture (Chapter 9).

Enhancement of water quality and supply

Trees and other plants affect the hydrologic cycle (Chapter 4) in several ways. First, by holding soil in place and intercepting precipitation, forests minimize the amount of sediment damage to nearby waterways (Chapter 9). Second, interception slows the velocity of falling water and allows the soil surface to remain relatively permeable. As a result, infiltration is more likely than is runoff, and local supply of water is therefore enhanced (see Chapter 4 to review these pathways of the hydrologic cycle).

Storage of carbon

As primary producers, trees convert CO_2 into glucose (Chapters 4 and 5). This fixation of carbon is a necessary pathway of the C cycle because it helps recycle carbon atoms and make them available for consumers plus it provides a counterbalance to aerobic respiration and other processes that generate gaseous CO_2. Trees are therefore said to store or **sequester** large amounts of carbon that could otherwise enter the atmosphere (note that big, i.e., old, trees and forests are repositories of a large fraction of terrestrial fixed C, far more than young trees). You should realize that, despite the actions of the planet's primary producers, the amount of carbon dioxide gas in the atmosphere has been increasing since the end of the 1800s. This lack of equilibrium is largely due to anthropogenic activities such as combustion of wood and fossil fuel, that is, primary production cannot keep pace with the current rate of input. The situation only gets worse as forests are lost. We will see more about the balance between production and fixation of carbon dioxide when we consider global climate change in Chapter 14.

Other benefits. Human quality of life certainly is enhanced by the services described above, although the benefits are derived indirectly. Here we see some *direct* ways humans use forests.

Recreation

Many forested areas are maintained for activities such as hiking, camping, and picnicking.

Source of raw material

Commercial forestry is a large industry that takes trees from both private and public lands. The wood collected is processed and used as building material and in the production of paper products. Some of the wood burned as a fuel, a practice that extends far back into human history. These days, wood still provides about 6% of our worldwide energy supply (serving as the primary fuel source some

2 billion people[1]). The consequences of logging vary and depend on the techniques used to cut down trees.

Selective cutting. This strategy removes only some of the trees in an area. Since many trees remain standing after harvest, the forest can continue to provide valued services. If done appropriately, a forest of trees of different ages can be developed. Figure 12.1a shows a typical outcome of selective cutting.

Figure 12.1.a. *Selective cutting leaves many trees standing. US Forest Service, Public Domain.*

Clearcutting. This second approach to harvesting removes all the trees present in an area at one time (Figure 12.1b).

1. Food and Agricultural Organization of the United Nations. 2021. Wood energy

Figure 12.1.b. *Clearcutting removes all standing vegetation in an area at one time. Siegmund W, CC SA-BY.*

Certainly quicker and more efficient than selective cutting, it also brings immediate and dramatic changes to an existing ecosystem. Many possible outcomes can follow clearcutting, including landslides (see Chapter 7). In some cases, entire forests are deliberately burned instead of being harvested. Among other consequences of these intentional fires: the organic carbon stored in trees is converted to CO_2 gas.

12.1.3. Deforestation: causes and consequences

Deforestation leads to losses of services

It should go without saying that humans and natural systems can only continue to benefit from forests as long as forests persist. Losses, broadly known as **deforestation**, can be caused by many anthropogenic and natural forces.

Anthropogenic. Humans can reduce the amount of forested land on Earth in direct and indirect ways.

Direct

Harvesting of wood for building materials and fuel as well as land clearing (using fire or other methods) to create space for agriculture, housing, transportation, and the like require the intentional removal of the trees and other organisms in a forest. The consequences of direct removal of forests vary and depend on the size and extent of the damage done. For example, clearing of an acre or two within a forest will likely bring about less pronounced effects than will the sudden transformation of a thousand acres from thick forest to barren land. Keep in mind, though, that even seemingly subtle changes, like those associated with the construction of a highway (with the resultant habitat fragmentation) could profoundly impact a threatened or endangered species (Chapter 6).

Indirect

Without intending to remove trees or clear space, humans can still adversely affect forests. Emissions from fossil fuel combustion, for instance, can lead to the formation of acid rain and increased stress on trees (Chapter 14). So, because it is done in the interest of electricity generation, coal burning is classified as an indirect cause of deforestation. Climate change due to the release of carbon dioxide and other air pollutants (Chapter 14) can also reduce the survival of certain tree species (climate changes from natural causes, common before humans industrialized, caused deforestation as well). Diversion of water away from forests for agriculture or other human activities belongs in this category, as does the introduction of invasive species (Chapter 6).

Natural. We learned in Chapter 7 that humans are not the only sources of stress on Earth. Forests, like all systems, must also contend with natural hazards.

Tectonic hazards

Earthquakes and volcanoes can be very destructive, as we saw in Chapter 7.

Water

Both too much and insufficient water can damage forest ecosystems. The effects of drought are likely familiar to you: trees (and all plants) will not survive if precipitation drops below critical levels. Remember, though, that the amount of water available to plants is also a function of soil properties and that those soil properties are affected by land use decisions (review Chapter 9 for more about soil). On the other hand, flooding is problematic because standing water will ultimately kill most trees adapted to grow in unsaturated soils. Climate change and rising sea levels have worsened the effects of flooding in coastal regions, submerging forests under saltwater (Figure 12.2).

Figure 12.2. *The effect of flooding on forests. Left (a): standing water covers trees. Right (b): dead trees remain after water recedes. Compare to Figure 7.14.* Poon W-C, CC BY-SA *(a);* NOAA, *Public Domain (b).*

Parasites

Microbial and animal parasites have the capacity to kill large numbers of trees and bring about wide-spread deforestation. For example, starting in about 1940, populations of Dutch elm trees in North America were devastated by a fungal pathogen that was accidently introduced from Europe[2], and forests were dramatically

2. Marcotrigiano, Michael. 2017. Elms and Dutch elm disease: a quick overview. In: Pinchot, Cornelia C.; Knight, Kathleen S.; Haugen, Linda M.; Flower, Charles E.; Slavicek, James M., eds. Proceedings of the American elm restoration workshop 2016; 2016 October 25–27; Lewis Center, OH. Gen. Tech. Rep. NRS-P-174. Newtown Square, PA:

affected as a result. Infestations by the insect known as the woolly adelgid similarly affects hemlock trees in the eastern United States (Figure 12.3).

Figure 12.3. *The effect of parasites on trees. Left (a): dead Dutch elm trees. Middle (b): fir trees killed by woolly adelgid. Right (c): magnified image of woolly adelgid infestation on a tree. Webb, R. CC SA (a); Phenz at en.wikipedia, CC BY-SA (b); Connecticut Agricultural Experiment Station, CC BY (c).*

Fire

Very large fires, whether caused by natural forces such as lightning strikes or human actions, can devastate hundreds of thousands of acres of forests. Consequently, forest management for most of the past century or so has featured concerted efforts to prevent all forest fires. Unfortunately, this policy has actually increased the frequency of catastrophic fires. As counterintuitive as it may seem, some small fires are necessary to protect forests. See Box 12.1 for a look at ways fire can both destroy *and* protect forests.

Box 12.1. We can (and should) fight fire with fire.

U.S. Department of Agriculture, Forest Service, Northern Research Station: 2–5

Most people view fire as the mortal enemy of forests. Yes, a large, out-of-control fire surely is destructive, but periodic small fires tend to protect forests from infrequent but catastrophic fires. If this sounds illogical to you, you are not alone. In fact, professional forest management has featured aggressive fire suppression for much of the past century for the simple reason that trees are made of wood…and wood burns easily. Policy makers reasonably assumed that the best way to protect forests was to prevent all fires. Alas, as usual, the situation is not nearly so straight forward.

Trees shed all sorts of products onto the floor of their forest, including leaves, branches, and bark. This debris dries and becomes readily combustible, meaning it serves as fuel for any fire that might be initiated by whatever cause. Under natural circumstances (i.e., without humans), fires start regularly—every few years or so—and those flammable materials do not accumulate into particularly large piles. Such small fires burn out quickly because they use up their fuel before they can get big and hot enough to ignite the large, standing trees. Saplings and other small plants are often killed, but the tall, old trees escape with little more than charred trunks. Furthermore, fire can stimulate the release of seeds from certain plants and otherwise enables new growth.

What's the problem? By working to prevent any fires, big or small, from affecting a forest, well-intentioned forest managers encourage the accumulation of unnaturally large piles of extremely flammable debris. And, unfortunately, something will ignite that fuel sooner or later. Imagine, for example, what could happen if a drought followed a couple of decades (or more) of successful fire suppression. One bolt of lightning could lead to an immense and highly destructive fire (actually, no imagination is necessary, as this scenario plays out with some frequency, including during the summer of 2021). Instead of staying low to the ground, the flames reach up to burn the tree crowns (i.e., the high branches where leaves are concentrated) and wipe out an entire forest.

In the past several years, policy makers and managers have incorporated knowledge about the importance of fire into their maintenance schemes. Controlled burns and natural fires have been used to some extent, but decades worth of accumulated fuel still imperils many forests.

12.1.4. How do we protect forests?

Put simply, we can consult our tools of systems analysis: if **reforestation** and **afforestation**, that is, recovery of previously forested area and an increase in the amount of land that is forested, respectively, can compensate for losses due to harvesting and land clearing, then there will be no net loss of forests. Approaches to achieving an acceptable balance vary within and among countries, and both public- and private-sector decision makers exert influence over the disposition of forested land. For example, management of forests in the United States is the responsibility of two federal departments (although many local and state agencies play important and, at times, conflicting, roles).

The National Parks Service (NPS) of the Department of the Interior.

The agenda of the NPS is to manage the 85 million acres designated as National Parks[3]. These areas are maintained as recreational sites (camping, hiking, etc.) but are closed to timber harvesting. Thus, it would be fair to say some measure of both conservation and preservation (Chapter 6) is accomplished in these parks: yes, humans are clearly allowed to enter them, but forests here are not used as sources of wood.

The Forest Service (FS) of the Department of Agriculture

The FS manages 193 million acres of land, including forest and grassland areas[4]. The name of the agency might suggest otherwise, but National Forests are open to logging—timber companies buy rights to harvest trees from these federal lands. National forests are also open for mining as well as recreation. The job of the FS is clearly different than that of the NPS, and, although it attempts to manage the forests sustainably, its mission and methods are controversial among some environmental activists and scientists.

3. US National Parks Service. NPS.gov; the total land area of the United States is around 2.3 billion acres.

4. The Forest Service, US Department of Agriculture. fs.usda.gov

Not directly related to the discussion of forests but still important: grasslands managed by the Forest Service are leased to ranchers for animal grazing.

Striving for sustainability

So, what is the verdict: are forests increasing or decreasing in size? As is (too?) often the case for us, the answer is not as definitive as we would like it to be.

More trees does not necessarily equate to more forests. You may hear the oft-quoted statistic that there are more trees today than there were a century ago—it is a common talking point among those who believe deforestation is not a problem (or, who wish to make it appear so). Evidence does support the broad claim of an increase in the total number of trees since about 1900, but we need to be careful to distinguish between *forest ecosystems*, which feature natural biodiversity, and *tree plantations*, essentially agroecosystems designed to maximize the growth of one tree species and minimize the growth of everything else (refer to Chapter 9). Tree farms, in other words, may support large numbers of trees but they do not provide the same services as forests. Furthermore, long-term trends are very different than those of the past several decades: recent estimates suggest there are half as many trees on Earth today than there were at the start of human civilization[5].

Earth is experiencing net deforestation. According to United

5. Crowther, T., Glick, H., Covey, K. et al. 2015. Mapping tree density at a global scale. Nature 525:201–205. doi.org/10.1038/nature14967

Nations scientists, the rate of deforestation outpaced forest growth between 1990 and 2020, even as the rate of loss of forests has slowed in recent years[6]. Among other important consequences of this change: there has been net movement of carbon from the terrestrial biosphere to the atmosphere. Unsurprisingly, neither forest cover nor rates of deforestation are evenly distributed globally, with just five countries (Russia, Brazil, Canada, the U.S., and China) serving as home for more than half of Earth's forests and Africa and Asia experiencing the highest and lowest rates, respectively, of deforestation[7].

12.2 FISHERIES MANAGEMENT

12.2.1. What is a fishery?

The definition of these aquatic systems is rather broad. According to the Food and Agriculture Organization of the United Nations (UN FAO), a fishery can be defined by one or more of the following: a region of the ocean, the organisms caught, the type of equipment

6. FAO and UNEP. 2020. The State of the World's Forests 2020. Forests, biodiversity and people. Rome. doi.org/10.4060/ca8642en https://doi.org/10.4060/ca8642en

7. FAO and UNEP. 2020. The State of the World's Forests 2020. Forests, biodiversity and people. Rome. doi.org/10.4060/ca8642en https://doi.org/10.4060/ca8642en

used, or the people involved in the work[8]. In any case, it should be clear that these systems are managed to provide food and income to people.

> *Yes, people extract fish from both freshwater and marine fisheries, but since the latter are so much larger than the former, we will focus our attention on oceans.*

12.2.2. Why are fisheries important?

Fisheries provide valuable services

Nutrients. However we define them, fisheries harvest organisms that are overwhelmingly consumed by humans. Fish contribute an average of about 15% of the dietary animal protein consumed by half of the people in the world[9]. As you likely expect, those numbers vary with culture, location, and level of development.

Livelihoods and income. The commercial fishing industry provides employment to approximately 10% of the world's population[10].

8. Fisheries and Aquaculture Department (FAO). 2014. United Nations. fao.org
9. Fisheries and Aquaculture Department (FAO, 2021). United Nations. fao.org./rural-employment/agricultural-sub-sectors/fisheries-and-aquaculture/en/
10. Fisheries and Aquaculture Department (FAO, 2021). United Nations.

Maintenance of organisms and their ecosystems. Although not the primary objective of fisheries management (which can reasonably be understood as the advancement and wellbeing of humans) the health of ecosystems that support commercially important fish is clearly relevant to the industry. Practices that threaten aquatic habitats, such as land clearing, with its resultant sediment damage, for agriculture (Chapter 9) and forestry (above), dumping of waste (Chapters 11 and 13), climate change and sea-level rise (Chapter 14), and the intentional or inadvertent release of chemical and biological pollutants (Chapter 15) are of great concern to those who make their living from fishing. As we will explore in some detail below, commercial fishing itself also can cause substantial degradation of marine ecosystems and the species which rely on them.

12.2.3. Threats to fisheries

The fundamentals of this story are like those of so many others we have encountered: if the rate of removal of a marine organism from the ocean exceeds the rate of replacement of it to the ocean, there will be a decrease in the size of the population of that organism. Natural processes certainly do play roles, but anthropogenic activities have been largely responsible for declines in both commercially important fisheries and, more broadly, marine ecosystems.

fao.org./rural-employment/agricultural-sub-sectors/fisheries-and-aquaculture/en/

Overfishing

This phenomenon involves the unsustainable use of marine resources. A species (fish, lobster, etc.) is said to be overfished if the rate of reproduction of it cannot keep pace with the rate at which it is caught by humans. Here we briefly consider some of the explanations for this important phenomenon.

> ***Maximum sustainable harvest*** *refers to the total amount of a species that can be taken without leading to its decline. Compare this concept to* ***maximum sustainable yield****, a term we used in the context of agriculture* (Chapter 9).

Unrealistic views of ocean resources. The vastness of the oceans contributes to the faulty conclusion—widely held historically—that no amount of fishing by humans could appreciably affect the number of marine organisms. The fact that many aquatic organisms are **r strategists** (Chapter 5) can lead one to assume that reproduction is more than sufficient to keep up with harvests. To cite just two examples, an average adult female Atlantic salmon releases about 7,500 eggs upon spawning (after it spends two years at sea), and the female shortnose sturgeon releases somewhere between 30,000 and 200,000 eggs every three to five years[11]. Despite these seemingly high numbers, both species are endangered. We will see other species that have been brought to the brink of extinction by unsustainable harvesting shortly.

11. National Oceanic and Atmospheric Administration, U.S. Department of Commerce. fisheries.noaa.gov/species

Efficient and indiscriminate technology. The commercial fishing industry is like most enterprises in that it strives to be as cost effective and profitable as possible. Technological advances during the past several decades have enabled the capture of large numbers of marine organisms with, relatively speaking, low energy expenditure, effort, and financial investment. For the most part, commercial fishing does not involve people using fishing poles to catch one fish at a time (recreational and subsistence fishing do rely on these approaches). Unfortunately, increased efficiency comes with adverse consequences. First, as stated above, organisms that are commercially important can be extracted from the oceans at very high rates. Second, organisms that are *not* of interest can also be captured and killed because many of the commonly used fishing techniques do not discriminate among species. That is, in addition to the target fish (for example, bluefin tuna or flounder) many non-target animals (to name just a few, turtles, dolphins, and whales) can be inadvertently taken. This second group of organisms, collectively referred to as **by-catch**, is typically hauled aboard fishing vessels and then thrown back into the water (often, dead) because it has little commercial value (Figure 12.4).

F***igure 12.4.*** *By-catch. Left (a): most of the organisms on the deck are not the target so will be discarded overboard. Right (b): a shark caught in a net intended to capture different organisms will likely die before it is released.* NOAA, *Public Domain (a and b).*

An understanding of the details about equipment used by commercial fisheries is unnecessary to appreciate the problems of overfishing. Keeping things simple, devices that capture target organisms include baited hooks floating at various depths for extended periods of time, purse seine nets that encircle a school of fish (after the fish have been located), trawl nets that are dragged through water or ocean sediment, and box-like devices known as traps that are left and checked periodically (Figure 12.5; see also Box 12.2 for some information about a tool known as a driftnet). The strategy chosen depends on the target species, but all the gear can potentially lead to by-catch. For example, dolphins drown when trapped under water by nets or long lines (remember that they are mammals and breathe air).

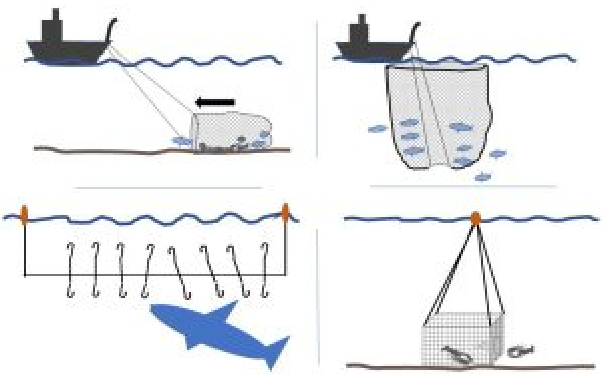
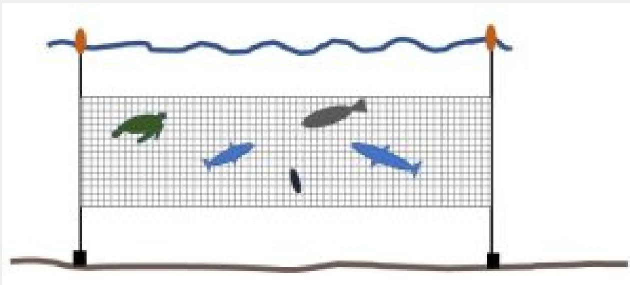


Figure 12.5. *Four commonly used types of fishing gear. Clockwise, starting with upper left: trawl net, baited hooks, trap, purse-seine net. Not to scale. Kelsey,* CC BY-NC-SA.

Box 12.2. Driftnets: passive, efficient, deadly

A particularly controversial fishing technique involves the use of **driftnets**. As the name suggests, these devices are not steered or pulled through the ocean, rather, they are set up in a location (with anchors and floats) and left to passively catch whatever organisms get tangled in them (refer to the diagram, below, a cross-sectional view of a hypothetical net). They are quite efficient at catching fish, but they also lead to a tremendous amount of by-catch. Scientists, policy makers, and activists have concerns about these tools. Among the many potential issues, they are sometimes forgotten or abandoned, meaning they can kill both target and non-target organisms for extended periods. Rules about their use have gotten increasingly strict since the 1980s, although nets as large as 2.5 km in length are still legal in some fisheries[12]. Illegal use continues as well.

12. United Nations. 1998. FAO Fisheries Department. Report of the Food and Agriculture Organization concerning United Nations General Assembly Resolution 52/29 entitled “Large-scale pelagic drift-net fishing; unauthorized fishing in zones of national jurisdiction and on the high seas; fisheries by-catch and discards; and other developments



Environmental disruptions

Changes in environmental conditions caused by anthropogenic activities also exert stress on fisheries and can lead to their decline.

We will focus on human sources of stress here, but keep in mind that natural forces also can alter environments in important ways (Chapter 7).

Pollution. We have seen that many human activities release physical, biological, and chemical pollutants into natural systems. For example, sediment pollution from agriculture and deforestation can degrade coastal waters and harm fisheries. Ocean dumping of industrial, domestic, and power-generation-related waste introduces many toxic substances as well. Air pollutants from various types of combustion are also of concern because they can mix with precipitation that falls into the oceans. The release of

CO_2 from fossil fuel combustion is important for multiple reasons, including the way it ultimately leads to acidification of ocean water (we will see more about this and related phenomena immediately below and in Chapter 14). The relationship between environments and organisms is well known to us by now, so it should come as no surprise that a change to the pH of the water in which organisms spend their lives can affect the structure and function of marine ecosystems (see more about pH and acidity in Chapter 9).

Climate change. Since this stressor is addressed in detail in Chapter 14, we will not devote much space to it here. Very briefly, human activities are driving, among other things, increases in the average global temperature, changes in climate on local scales, shifting ocean currents, and rises in sea level. Scientists continue to collect data on the likely consequences of climate change, but it is clear that many economically important marine species have already changed their habits and ranges in response to their altered environments. Monitoring of the distribution of American lobsters in the coastal waters of New England (U.S.A.), for instance, has revealed a distinct shift of population centers further from the shore during the past few decades as these mobile organisms search for cooler waters. Black sea bass, found along the entire eastern coast of the United States, have similarly moved north with time as water temperatures continue to rise[13]. Put into practical terms, owners of commercial vessels need to spend time and money tracking their target organisms and move with them into new locations, often expending more fuel to do so, to maintain their harvests and livelihoods.

Habitat alteration and fragmentation. Changes to habitat overlap with the previous issues about climate change in that they broadly affect the environment of an organism. Humans can alter

13. NOAA. 2021. “The northeast shelf: a changing ecosystem.” fisheries.noaa.gov

physical habitats in many ways, for example, through the installation of offshore oil platforms, wind farms, and sea walls. These and similar structures can force organisms to change location and may split a population that was once continuous (more about fragmentation can be found in Chapter 6).

Invasive species. Competition and predation by introduced species can threaten the survival of certain marine species. Review Chapter 6 for more about invasives.

12.2.4. Examples of managed fisheries

The number of protected marine organisms and fisheries is quite high. As usual, space considerations necessitate a limited look at some notable examples. In other words, the few cases described here are intended to provide a sense of the variety of challenges to management and are by no means exhaustive. If you find yourself wanting more, consult the United States NOAA website, fisheries.noaa.gov.

Atlantic salmon

Most people know the fascinating and somewhat unusual salmon story: these are the fish that hatch and spend their early years in freshwater rivers, swim to the ocean for a year or two as adults, and then return to the stream from which they originated to spawn (i.e., mate) and produce the next generation of eggs. They travel great distances, moving back and forth across boundaries among states and countries and, thus, face diverse threats to their survival and success. In addition to natural predation during all phases of their lives, hazards from human activities have hindered them. The several causes of fishery decline described above affect salmon, but some are more important than others. Habitat fragmentation resulting from dam and bridge construction, as well as development of hydroelectricity power generation dams (Chapter 10), impose major impediments to the necessary migration of these fish. Also, excessive harvesting exerted so much pressure on natural salmon populations during the past century that commercial and recreational fishing of wild Atlantic salmon has been banned in the United States since 1948. International cooperation to protect this species from further decline (and the risk of extinction) has increased, yet overfishing in waters outside the jurisdiction of the U.S. government continues to threaten the long-term viability of the fishery[14] (see Box 12.4 for a little more about salmon).

Box 12.4. Wait, what? No Atlantic salmon??

Right: wild Atlantic salmon may not be legally caught in U.S. waters by either commercial or recreational fishers. That fact might surprise you if you have ever bought salmon at an American store or restaurant, where salmon seem to be readily available. So, where do those fish come from? They are raised in fish farms, a topic we will explore later in this chapter.

Sharks

This term refers to hundreds of fish species that, although sharing certain fundamental characteristics, range widely in numerous

14. National Oceanic and Atmospheric Administration, U.S. Department of Commerce. fisheries.noaa.gov/species

ways, including size, habitat, prey, and risk posed to humans. In other words, the one-size-fits-all definition used by many people is inadequate and leads to misplaced fears about and attitudes toward these important organisms. It also has fueled the killing of about a hundred million sharks each year (yes, you read that right: 100,000,000).

Modern-day sharks are closely related to organisms that have been in the oceans for hundreds of millions of years. In other words, they both arrived before and outlived dinosaurs. How have they survived for so long? As good environmental scientists we know the short answer: they are extremely well adapted to live in ocean environments. We might say, in a somewhat informal way, they found strategies that worked long ago and have stuck with them. These days, many shark species are struggling against extinction due to pressures from anthropogenic activity that remove them from the ocean faster than they are replaced. Sharks are especially vulnerable to overfishing because they tend to have lower-than-usual reproductive rates relative to other fish. Most of them do not reach sexual maturity for several years and only produce one or two offspring annually. It is difficult to count the numbers of individuals of any ocean species, but data from many different scientific sources indicate that shark populations have declined precipitously during the past several decades. One recently published analysis concludes that the number of known oceanic sharks (and rays, closely related organisms) has decreased by three quarters since 1970[15]. Some species are in better shape than others, however the risk of extinction of these ancient organisms is high overall.

People hunt for sharks for two main reasons: their fins are highly prized for the perceived (but not-supported-by-scientific data) medicinal benefits they provide and, to some extent, for food (e.g.,

15. Pacoureau, N., Rigby, C.L., Kyne, P.M. et al. 2021. Half a century of global decline in oceanic sharks and rays. Nature 589, 567–571

an expensive concoction known as shark fin soup). That second reason has been responsible for the bulk of the loss of sharks from the oceans. It is banned in some countries (including the United States) and highly regulated by international conventions, but **shark finning**, or the practice of capturing an individual, cutting off its fins, and then throwing the live fish back into the water where it slowly sinks and dies, continues to threaten the survival of sharks worldwide (Figure 12.6). We should also add fear, hate, and misunderstanding to the list of reasons people hunt sharks.

Figure 12.6. *Two images of shark finning. Top (a): a dorsal fin that was just cut off a live shark; the finless animal will be discarded into the ocean. Bottom (b): multiple shark fins drying on a sidewalk in Hong Kong. Naka9707, CC BY-SA (a); Cloneofsnake, CC BY SA (b).*

Contrary to popular belief, these animals are hardly our enemies, and you should care about the loss of sharks from the oceans. First, only a very small number of species (among them, great white, bull, white tip, nurse, and tiger) pose a threat to humans, and shark attacks generally occur when a land-based person entering a water-based shark's habitat is mistaken for prey. Second, sharks play critical roles in their ecosystems because, as top-level predators, they help maintain the size of their prey populations.

In an average year, four or five humans are killed by *sharks whereas about 100,000,000 million sharks are killed by humans in the same period. Think of it this way: for every* ***one*** *human death due to sharks, around* ***20 million*** *sharks die at the hands of humans.*

American Lobster

This is not a fish, of course, but is an important marine species nonetheless. Lobsters have relatively long life spans (possibly as much as a century, if they escape harvesting), beginning as tiny free-swimming larvae and eventually settling to the seafloor after a few months. They provide a great example of r-selected reproduction: of the many thousands of eggs that can be released by a female at one time, literally just two or three are likely to survive to adulthood[16].

These organisms can be found in the Atlantic Ocean off the northeastern states of the U.S., although the bulk of the population is in the cooler waters near Maine. In recent years, 93% of the commercially harvested American lobsters have come from the Gulf of Maine[17]. This fishery has enormous economic importance: in 2019, for example, it was valued at about $625 million. Many thousands of people depend on lobster for their livelihoods and many millions more expect their demand for lobster as food to be

16. Fishwatch, fishwatch.gov. Part of the National Oceanographic and Atmospheric Administration. NOAA.gov

17. NOAA Fisheries. fisheries.noaa.gov National Oceanographic and Atmospheric Administration. NOAA.gov

met. Consequently, the ongoing health of these organisms is the focus of much effort and concern.

Management of this fishery is complicated by the fact that it cuts across multiple political boundaries and jurisdictions, and each may have different priorities and agendas. In the U.S., states control the water out to 3 miles from shore whereas NOAA (federal) is responsible for the area between 3 to 200 miles out to sea. A third entity, known as the Atlantic States Marine Fisheries Commission (ASMFC), is a regional organization working to manage many migratory species, including the lobster. It consists of representatives from the fifteen coastal U.S. states that run north-south from Maine to Florida and works cooperatively with individual states, NOAA, and the U.S. Fish and Wildlife Service (FWS)[18]. Together, these groups enact and enforce regulations to protect the fishery. For example, individual lobsters must meet minimum *and* maximum size requirements (size is a proxy for age) if they are to be kept. Very small lobsters are too immature to have mated and large ones are thrown back because they have the potential to produce very high numbers of eggs (older females produce about 20 times more eggs per breeding event than do younger ones[19]). Other rules aimed to increase sustainability have been enacted as well.

Opinions vary on whether efforts to conserve this fishery will be successful. Harvests in northern waters near Maine have been high for decades and even increased in recent years, leading some people to conclude that the proper balance of input and output has been reached. Others point out that the situation will be difficult to sustain in the long run because harvest rates likely exceed hatching and survival rates for new lobsters. The substantial decline in the fishery in waters south of Maine, for example, near New York and

18. Atlantic States Marine Fisheries Commission. asmfc.org

19. Fishwatch, fishwatch.gov. Part of the National Oceanographic and Atmospheric Administration. NOAA.gov

New Jersey, also has raised some concerns. Further complicating matters, climate change exerts substantial pressure on populations (see the note about shifts in lobster population in response to rising water temperature, above), even in the absence of intensive fishing.

Coral reefs

Our final example is different than the others in that it is not a particular organism but is a type of ecosystem that supports multiple species. They are not exactly fisheries, but nevertheless are critical to the health of many marine systems, including those that support the growth of organisms targeted for human harvest. What is a coral reef? They might appear to be built of rocks and similar non-biological materials, but reefs are constructed from the skeletons are living organisms called corals (Figure 12.7).

There are hundreds of different known species of corals, and the ones that form the basis of shallow-water reefs host mutualistic algae inside their tissues[20] (the coral provides habitat and the algae provides fixed carbon—consult Chapter 5 for more about this kind of relationship).

Figure 12.7. *A group of coral-forming organisms (each is called a polyp).* NOAA, *Public Domain.*

Corals reefs provide many important services. First, they act as a kind of protective barrier between the open ocean and nearby shorelines; reefs dissipate wave energy somewhat, an especially critical function during storms.

20. USGS. Coral reefs. Coastal and marine hazards and resources program. usgs.gov

Second, they are home to a wide variety of marine species adapted to the conditions provided by the reef—these areas are thought to support approximately 25% of the oceans' biodiversity[21] (Figure 12.8). Third, they are a source of recreation and income for many people. Maintaining healthy reefs is, therefore, of great value on many levels.

Figure 12.8. *The skeletons of corals make up the reef structure, around and within which a community forms. Wise Hok Wai Lum, CC BY-SA.*

Unfortunately, despite regulations and other types of management, the number of reef-building organisms in the world's oceans has declined precipitously in recent decades: in U.S. waters alone, the loss is estimated to be between 70 – 80% since the 1970s[22]. Certainly natural stressors such as hurricanes can damage reefs, but several anthropogenic activities play major roles in their decline as well. For example, agriculture has increased the amount of sediment and nutrients flowing to coastal waters (Chapter 9). Trash and other debris dumped into rivers or the ocean (Chapter 13) can similarly affect these sensitive ecosystems. Increases in greenhouse gases in the atmosphere (noted above and Chapter 14) also pose an ongoing threat to reef health. Since coral skeletons are built of calcium carbonate, rising ocean acidity slows coral growth, plus warmer waters damage corals. Finally, the harvesting of fish and the corals themselves can be very destructive, particularly when explosives or

21. USGS. Coral reefs. Coastal and marine hazards and resources program. usgs.gov

22. USGS. Coral reefs. Coastal and marine hazards and resources program. usgs.gov

cyanide is used (common practices). One important consequence of these and other stressors is a phenomenon known as **coral bleaching**, that is, the corals become pale as they lose their algal symbionts and approach death[23] (Figure 12.9).

Figure 12.9. *Three images showing the progression from a healthy reef to bleached coral to a dead reef.* NOAA, *Public Domain.*

12.2.5. Measures taken to protect fisheries

Protecting fisheries is a high priority for many people, but success requires a combination of science-based laws, enforcement, compliance, cooperation, and innovation.

Laws

Regulating fisheries is complicated for many reasons, not the least of which is the shared nature of ocean resources. Coastal waters

23. National Ocean Service. oceanservice.noaa.gov Part of National Oceanographic and Atmospheric Administration. NOAA.gov

are controlled by multiple states and countries having different agendas and priorities, and the open ocean—areas that are outside the jurisdiction of individual countries—is governed by various international agreements. Although we already encountered many of these ideas already, a terse summary of some of the management strategies used to conserve marine organisms is included here.

Several organizations and agencies conserve fisheries. The Food and Agricultural Organization of the United Nations works with member nations to develop and enforce international agreements, and countries (and states, where applicable) have their own regulatory bodies.

Permitting and registration requirements. Those wishing to legally operate a commercial fishing vessel must obtain the appropriate permits to provide accountability and track activity.

Catch limits. Along the coastlines of the United States, NOAA and states set annual limits on how many individuals can be harvested from fisheries. The limits are different for each species and can vary from year to year in response to population assessments. The rules for international waters are different, and some fisheries are not subject to limits.

Gear restrictions. Laws are in place to both limit the catch of a target species and reduce by-catch. In other words, some gear is prohibited in some fisheries (see above for more about common gear used in commercial fishing).

Size restrictions. As noted in the discussion of the American lobster, captured individuals from most fisheries must meet minimum and maximum size requirements if they are to be legally caught. The intention here is to optimize reproduction of a target organism.

Seasonal closings. Some fisheries are only open during a portion of the calendar year to protect species during their spawning seasons. Like so many other rules, these dates vary with organism and from year to year.

Enforcement, compliance, cooperation

The large size of Earth's oceans makes enforcement of rules, no matter how strict they may be, very difficult and necessarily uneven. It is simply impracticable to keep track of every ship, especially in international waters. Clearly it is in the best interest of fishers to voluntarily adhere to regulations—and most do so—yet illegal fishing (in violation of one or more laws regarding size, season, or gear, for example) does occur. Authorities may board and inspect ships when they are encountered at sea or when returning to port. If evidence of banned practices is found, say, quantities of shark fins without shark bodies, lobsters that exceed the size limit, or individuals of species that are protected from harvesting, impoundments and arrests are possible. Self-reporting and policing by those working in the industry also are important pieces of the compliance puzzle.

Alternatives to commercial marine fisheries

The intense pressure on organisms living in the oceans has spurred people to explore other ways to meet human demand for fish.

Inland freshwater fisheries. The fish in rivers and lakes are taken and used as food, but they are simply not abundant enough to meet more than a small fraction of worldwide demand. They are still regulated by local, state, and national agencies, though. In the U.S., the Fish and Wildlife service is responsible for freshwater fisheries.

Aquaculture. Fish can be intentionally bred and raised in various types of controlled environments in ways that are analogous to those used to grow cattle, chickens, and other animals (Chapter 9). Tree plantations (earlier in this chapter) are comparable as well.

Several strategies are used

Approaches to aquaculture range from the raising of fish in ponds or tanks that are separate from natural marine systems to cages submerged in ocean waters. In many cases, young are bred in hatcheries before they are introduced into wild fisheries or a fish farm enclosure (see Figure 12.10).

Figure 12.10. *Two common approaches to fish farming. Left (a): cages within the ocean water are used to grow and confine a species of interest. Right (b): tanks that are maintained away from the ocean's water are packed with fish.* Asc1733, CC BY-SA (*a*); Narek75, CC BY-SA (*b*).

Many types of organisms are farmed

Finfish (e.g., salmon), shellfish (e.g., oysters), and seaweed are all grown in aquaculture systems.

Food production via aquaculture is increasing worldwide

As the human population grows, the need for food, including fish, also goes up. In recent years, aquaculture has met an increasing proportion of that demand: worldwide, about half of all seafood consumed comes from fish farms (although, in the U.S., wild-caught fisheries still dominate)[24].

There are pros and cons of aquaculture

Like land-based agriculture, fish farming offers benefits and brings about adverse consequences. On the plus side it provides food for people and helps protect natural fisheries. Its costs, on

24. Growth in aquaculture contributes to rise in fish consumption. United Nations. news.un.org

the other hand, are not trivial. First, raising fish is very resource intensive. Although research to improve efficiency is ongoing, a lot of pressure is put on natural ecosystems, both marine and terrestrial, to obtain food for farmed organisms. Second, the high density of fish in an enclosure leads to the release of large quantities of nutrients (in the form of fish excrement) and an increase in **eutrophication** and other types of water pollution. Finally, submerged cages do not provide a fool-proof barrier against movement of fish into the open ocean; farmed individuals that escape can introduce harmful diseases and genetic traits into wild populations.

12.2.6. Oceans are not just sources of human food

The focus of this chapter is the management of natural resources for the purposes of meeting human demands for food and other materials. Before we move on, though, a few words about other reasons to protect marine ecosystems are in order.

Earth is sometimes called "The Blue Planet" for a very good reason: around two thirds of its surface is covered with water. As we learned in Chapter 4, the oceans hold about 97% of that water, and they make up about 99% of the planet's habitable space[25]. Evidence that life began under water is strong (Chapter 6), although, as most of us reading this book can attest, it has since moved onto land as well. Organisms living in the oceans are critical to the maintenance of natural systems as we know them, including photosynthetic phytoplankton that provide approximately half of Earth's oxygen[26]. Phytoplankton are also vital because they occupy

25. Living Ocean. Science.NASA.gov

26. Living Ocean. Science.NASA.gov

the first trophic level of marine food webs (including those that support commercially important organisms).

Much research has been conducted on, and much written about, the status of the world's oceans. Conclusions and predictions vary to some extent, but most people working in the field warn that marine systems are in decline. United Nations scientists, for instance, note several concerns[27], including the following three. First, about 60% of marine ecosystems of direct importance to humans are not being used sustainably, and as many as 13% of commercial fisheries may have collapsed (i.e., are no longer able to support relevant harvesting). Without substantive changes, half of the species in the oceans could face extinction by the year 2100. Second, releases of nutrient pollution, including sewage, ocean dumping, and agriculture have led to hundreds of low-oxygen areas, known as **dead zones,** in which aerobic organisms (like fish) cannot survive. One of the largest of these zones is found downstream of the Mississippi River (United States) in the Gulf of Mexico. It varies in size but in 2022 covered an area of about 8500 square km[28]. Third, the conditions of mangrove swamps and salt marches, vital coastal ecosystems because they protect the mainland, are breeding areas for many organisms, and are vital carbon sinks (i.e., store carbon that could otherwise move to the atmosphere—more in Chapter 14), have deteriorated precipitously due to chemical pollution and human activities such as boating[29]. Unless steps are

27. United Nations educational, scientific, and cultural organization. 2017. unesco.org/

28. What is a dead zone? National Oceanographic and Atmospheric Administration. NOAA.gov

29. Human activities such as dredging and careless boating are threatening South Florida's mangroves and seagrass floridakeys.noaa.gov

taken to protect these and other threatened marine systems, all of Earth's aquatic and terrestrial biosphere could undergo dramatic changes.

THE CHAPTER ESSENCE IN BRIEF [30]

Ever-increasing demand for food, materials, space, and livelihoods threatens the long-term supply of many coveted resources. Forest and fisheries management serve as two important examples of strategies that can be employed to mitigate declining availability of valued natural products; like all such efforts, they have promise but face multiple challenges.

---

30. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 12 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more…[31]

How are clearcutting and conventional farming (Chapter 9) related and how does each contribute to mass wasting (Chapter 7)?

Can resource management counteract extinction (Chapter 6)?

How does deforestation affect the carbon cycle? Spoiler: you will get a similar question after Chapter 14!

How might deforestation harm the livelihood of the captain of a fishing boat? Might this be another episode of the principle of environmental unity???

How can the establishment of a minimum size requirement for a fishery (i.e., a size below which an individual fish is deemed too small to keep) be understood using input-output analysis and a systems model?

So what if we clear the oceans of aggressive sharks? Isn't that good riddance? Who needs 'em?

31. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

I don't eat fish, nobody I know earns a living through a fishery, and I live far from a coastline: so, why should I care about oceans?

Why is the protection of fisheries so difficult?

CHAPTER APPENDIX

Alphabet soup: a guide to some relevant agencies

Resource management is carried out by many agencies/groups at the local, state, national, and international level. For your convenience, those most relevant to the topics in this chapter are very briefly described below. Among other things you should note as you look through this list is its disparate nature: the science and business of conservation cuts across many different agencies and jurisdictions.

United States federal agencies

Department of Agriculture (USDA).

This Cabinet-level department is responsible for the very general task of managing food, nutrition, agriculture, and natural resources. The USDA also works to provide economic opportunity for rural areas in America[32].

National Forest Service (NFS). The NFS is part of the USDA and is responsible for the management of forest and grasslands[33]. See the section on forests and forest management, above, for more.

Department of the Interior (DOI).

The secretary of the DOI also is part of the President's Cabinet. Its mission, as described in its official documents, is to "manage America's vast natural and cultural resources[34]." Eleven different bureaus fall under the DOI, including the three listed here.

Geological Survey (USGS). The USGS is responsible for the monitoring of natural hazards (e.g., earthquakes), management of water and other essential resources, and the study of ecosystems and the effects of climate change on human and natural systems[35].

Fish and Wildlife Service (FWS). This bureau works to conserve endangered and threatened species (mostly domestically, but contributes to international efforts as well), monitors migratory fish and freshwater fisheries, oversees wetland habitats, and enforces wildlife laws[36].

32. About the US Department of Agriculture. usda.gov/our-agency/about-usda

33. This is Who We Are. The Forest Service, US Department of Agriculture. fs.usda.gov

34. About Interior. doi.gov

35. Who We Are. usgs.gov

36. Abpout us. fws.gov/help/about_us.html

National Parks Service (NPS). The NPS primarily works to protect natural forests and other areas as described in the section on forests and forest management, above[37].

Department of Commerce (DOC)

You might be surprised to see this department on our list, another whose Secretary is a member of the President's cabinet. Despite its mission to "create economic growth and opportunity[38] ", it oversees a bureau that comes up frequently here and in other chapters of this textbook.

National Oceanographic and Atmospheric Administration (NOAA). Its website states "NOAA is an agency that enriches life through science. Our reach goes from the surface of the sun to the depths of the ocean floor as we work to keep the public informed of the changing environment around them[39]." Among NOAA's responsibilities are management of marine fisheries and coastlines, weather forecasting, and climate monitoring. Its work is done in the context of expanding economic development and opportunities.

37. About Us. US National Parks Service. NPS.gov

38. About Commerce. commerce.gov

39. About Our Agency. NOAA.gov/about-our-agency

International organizations

United Nations Food and Agricultural Organization (UNFAO).

This group is part of the United Nations, meaning it is made up of 193 member states (as of July, 2021). Its mission is to promote food security for all people[40].

State and regional organizations

Atlantic States Marine Fisheries Commission (ASMFC). As noted in the section on fisheries management, above, the ASMFC is a group made up of representatives from several U.S. states, NOAA, and FWS to manage many fisheries along the east coast of the United States[41].

Note also that U.S. states and other countries have offices of environmental protection, fisheries, natural resources, and so forth, although names differ (e.g., Department of Fisheries, Department of Inland Fish and Wildlife, etc.)

40. About FAO. fao.org/

41. About Us. Atlantic States Marine Fisheries Commission. asmfc.org

13. Waste Management

JASON KELSEY

Earth is closed with respect to materials, as we know very well by now. Thus far we have explored the first of the two consequences of this fundamental principle described in Chapter 1: the supply of living space, arable land, food, energy sources, forests, clean water, and other valued resources is not infinite. In Chapter 13 we will look at the second consequence, namely, that waste cannot be eliminated, rather, unwanted products generated by the billions of people on this planet must be managed within the confines of our system. In other words, when you *throw something away* it does not disappear—what you are really doing is sending your waste on a long journey that moves it among multiple receptacles, processing centers, and storage facilities.

Key concepts

After reading Chapter 13, you should understand the following:

- The many sources and types of human-generated waste
- How and why waste management was different in the past than it is today
- The costs, benefits, and challenges of the multiple

modern strategies used to manage waste

- The meaning of cradle-to-grave management of hazardous waste
- Some of the strategies used to reduce waste generated by humans
- Why proper waste management is important

Chapter contents

13.1 SOURCES OF WASTE

13.1.1. Natural

All living systems release waste as part of their normal functioning. Organisms extract nutrients from the food they ingest and then release unwanted products into their environments. Although some of those leftover materials are toxic to the individual releasing them, they will make a fine meal for another organism that is adapted to live off biological waste. Fungi and certain bacteria, for example, utilize remains and excrement to gain energy and materials they need. Of course, these decomposers also perform a critical function in that they free up carbon, nitrogen, phosphorus and the other elements vital to the continued survival of the biosphere (Chapter 4). Importantly, natural products rarely accumulate for extended periods but instead are fully recycled (with a few exceptions—we saw in Chapter 10 that petroleum, for example, persists in the lithosphere for hundreds of millions of years).

> *We are using "food" as a rather general stand-in for the sources of energy and materials that organisms acquire through many possible specific mechanisms. Consult Chapter 5 to review the different strategies used by members of the biosphere.*

13.1.2. Anthropogenic

Human activities generate a great variety of waste. Some of these materials may be decomposed into simple products which return to Earth's biogeochemical cycles (Chapter 4), but many are not so easily broken down by natural chemical reactions. As we will see shortly, the accumulation of anthropogenic substances can lead to several undesirable consequences, including exposure of human and natural systems to toxins, degradation of soil, air, and water, the reduction in available land appropriate for housing, farming, and other uses, and the removal of valuable resources from natural systems indefinitely. Because they pose different levels of risk and require different management approaches, anthropogenic waste is assigned to one of several categories.

Municipal solid waste (MSW)

Generally known as trash or garbage, this first group includes unwanted products people generate during their regular, daily routines. Relatively innocuous materials like glass, paper, plastic, fabric, and food scraps are among the constituents of MSW (consult Figure 13.1 to learn more). We will focus the bulk of our attention on this category of anthropogenic waste.

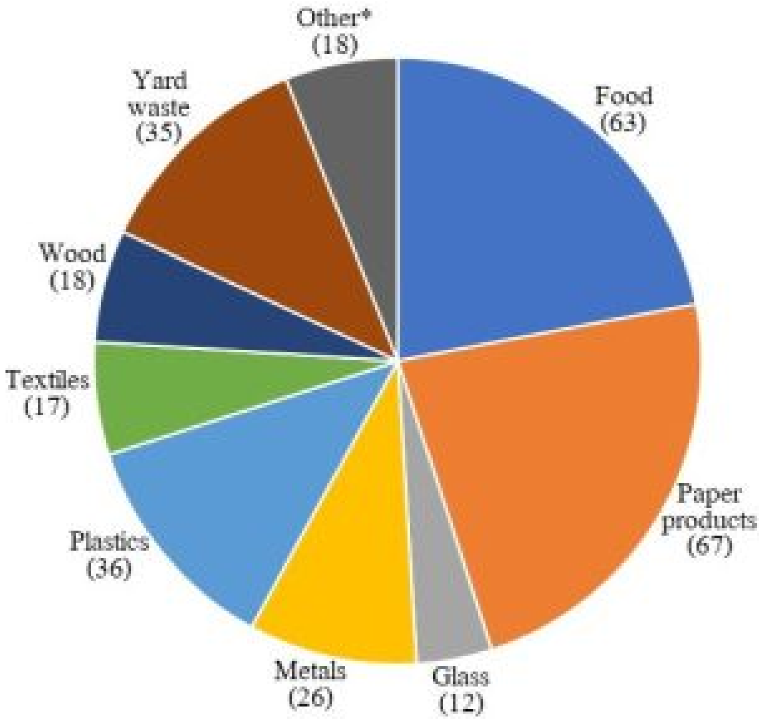


Figure 13.1. *The composition of municipal solid waste in the U.S. Numbers are tons (out of a total of 262.4 million tons generated in 2015). *Includes rubber, leather, inorganic waste, and other materials. Data from US EPA, Public Domain; figure by Kelsey, CC BY-NC-SA.*

Hazardous waste

Solvents, metals, explosives, drugs, pesticides, radioactive materials, and many other potentially harmful substances are released from homes, offices, factories, hospitals, universities, and military facilities. These substances require special handling and cannot be combined with MSW.

Biomedical waste

Bodily fluids, tissues, infectious agents, and objects that have come into contact with these biological substances (syringes, broken glass, sample vials, bandages, and related items) are included here. Like hazardous waste, biomedical waste is subject to stricter rules than those for household trash.

13.2. PAST PRACTICES

The problems associated with human-generated waste were not always quite as pressing or complex as they are currently. Prior to the middle part of the 19th Century, the volume of waste generated was small and easy to handle relative to what we see today. The reasons for this difference are straight forward: far fewer people were present on Earth (between 1 and 2 billion, roughly 1/5th of the current population) and resource use per person was also generally lower due to limited access to technology and lower standards of living (review Chapter 8 for more information about the history of human population growth). For example, in the United States, both the number of people and the amount of trash generated per person increased steadily from 1960 to about 2000, although the number has levelled off in recent years (Figure 13.2).

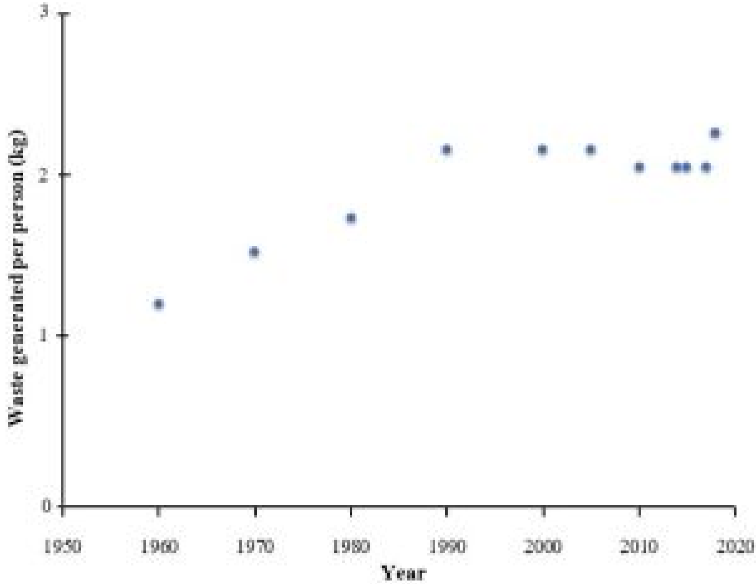


Figure 13.2. The amount of municipal solid waste generated per person in the United States from 1960 to 2018. Data from US EPA, Public Domain; replotted by Kelsey, CC BY-NC-SA.

Furthermore, waste generated today is broadly more diverse (i.e., it contains many more objects with different properties) and more toxic than it was in the past, making modern management both increasingly difficult and urgent.

It would be fair to say that, until the middle parts of the 1900s, most people were guided by the familiar adages "dilution is the solution to pollution" and "out of sight, out of mind" (Chapter 11) when it came to managing their unwanted products: garbage, sewage, and so forth were routinely left in open pits, buried in shallow holes, pumped into surface and underground waters, burned, or moved to a remote or sparsely populated location. Four specific approaches are briefly described here. Note that, although replacements for them have been developed during the past several decades, these older methods persist. Abandoned sites can still exert adverse effects (more below) and, in some areas, they are still actively used.

13.2.1. Open dumps

Some people equate the term “dump” with “landfill,” although the two terms are far from synonymous. The former refers to a place in which many types of garbage are left uncovered on the Earth’s surface with few or no barriers, or **containment** structures, to restrict access to and movement of the trash away from the site (Figure 13.3).

***Figure* 13.3.** *An open dump like this one in Indonesia lacks security and containment. McIntosh, CC BY.*

In other words, animals (including humans!) can scavenge in dumps, likely exposing themselves to hazardous substances, and toxic materials can readily flow via runoff and leaching to streams, lakes, oceans, and groundwater (Chapter 4 and Chapter 11). The later term, landfill, is associated with a modern approach to waste that has replaced dumps in many places, as we will see shortly.

13.2.2. Open burning

Another strategy used in the past was simply to ignite piles of household trash, tires, clothing, plastics, chemicals, and so forth. This practice, often referred to as **open burning**, is restricted in the United States by several laws linked to the **Clean Air Act**. You should realize, though, that this approach to waste management is still allowed (in the U.S. and elsewhere) under certain conditions.

13.2.3. Ocean dumping

Here, as the name suggests, trash is transported out to sea and released into the water (Figure 13.4). This strategy was once widespread, and many nations routinely discarded sewage, along with domestic, medical, construction, dredging, military, and even nuclear waste into the ocean.

Figure 13.4*. A barge transports trash to be dumped into the ocean. US EPA, Public Domain.*

Laws enacted during the past several decades have reduced, if not eliminated, the release of trash into oceans. For example, the United States passed the Marine Protection, Research, and Sanctuaries Act in 1988 to regulate this practice. In short, it largely prohibits dumping of several kinds of waste (e.g., that from hospitals, industry, and **sewage treatment facilities**—Chapter 11), although exceptions are made for some materials and circumstances (e.g., dredging waste).

13.2.4. Deep-well injection

Simply put, very deep wells transport liquids to rock layers thought to lie underneath groundwater reservoirs. This once common strategy is still used to some extent, but it become more tightly regulated by federal laws enacted in the 1970s[1]. See Fig. 13.5.

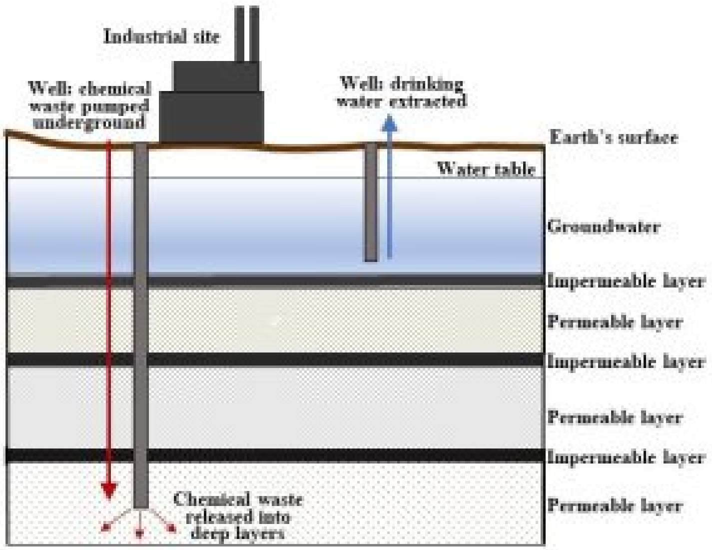


Figure 13.5. *Idealized diagram of deep-well injection. Waste from the factory is pumped via a well to a geologic layer that is about 1000 meters below and isolated from groundwater. In cross section, not to scale. Compare to Figures* [4.16](#) *and* [11.5](#)*. Kelsey,* CC BY-NC-SA.

13.3. CURRENT PRACTICES

Changes to the size, lifestyles, and expectations of the human population during the past 100 or so years, combined with

1. US Environmental Protection Agency (EPA). Protecting Underground Sources of Drinking Water from Underground Injection (UIC). https://www.epa.gov/uic/underground-injection-control-regulations-and-safe-drinking-water-act-provisions

increasing understanding of the adverse consequences of mismanagement of waste, rendered practices used historically inappropriate in today's world. We have learned that air, soil, and water are *not* infinitely able to dilute substances added to them. Indeed, by dumping, burning, and the like, humans have brought about appreciable changes to the chemical make up of Earth's atmosphere and hydrosphere since the middle of the 19th Century. As we will see in Chapter 14, combustion of fuels and other materials has led to the accumulation of toxins in the atmosphere. Oceans, lakes, rivers, and other water reservoirs have similarly been polluted by various anthropogenic activities. Furthermore, a decreasing number of places on this planet are suitable locations for long-term storage of waste. In short, population growth and industrialization have led to increasing demand for resources and living space such that new approaches to waste management have become necessary. Here we explore some of the common modern strategies used to handle the three classes of anthropogenic waste defined above and meet the standards mandated by federal, state, and local laws (see Figure 13.6 for a summary of the fate of MSW in the U.S.).

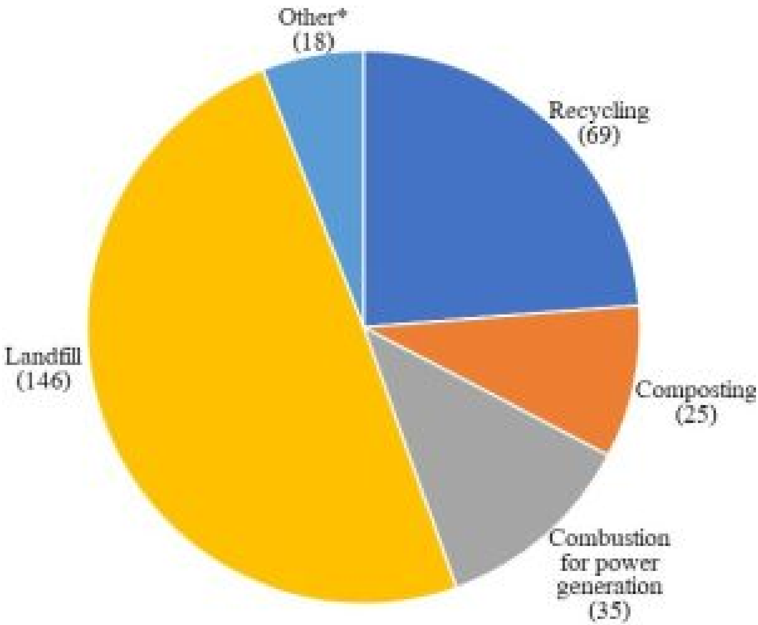


Figure 13.6. *Summary of the fates of municipal solid waste in the U.S., million tons (293 million tons were generated in 2018). *Includes food-waste handling. Data from US EPA, Public Domain; figure by Kelsey, CC BY-NC-SA.*

13.3.1. Municipal waste

Sanitary landfills

These underground storage sites isolate, or **contain**, certain types of low-toxicity trash such as that produced by households, office buildings, and schools. They employ a strategy that is functionally the opposite of the traditional dilution-is-the-solution-to-pollution approach because, rather than spreading waste out, landfills collect and concentrate it in relatively limited areas.

Landfill design and operation.

It is challenging to find appropriate sites

It likely goes without saying that few people get excited about the prospect of living near a landfill. Among the reasons for such reticence are concerns about the noise and congestion caused by large garbage trucks traveling on local roads, an increase in the number of scavengers (e.g., seagulls and turkey vultures) attracted to food waste, trash-related odors, aesthetic degradation of landscapes, and the decline of property values in affected areas. Furthermore, federal and state laws prohibit the construction of landfills in various places deemed inappropriate or particularly sensitive such as floodplains (Chapter 7) and wetlands (i.e., marshes, swamps, and the like). Companies seeking to open a landfill therefore encounter many political, economic, and legal barriers.

> *It is not uncommon for a party wishing to construct a new landfill (or other potentially controversial facility) to gravitate toward locations that are already economically depressed or populated by people who are disenfranchised (or at least are perceived to be so).*

Large areas are divided into individual cells

Contrary to the belief held by some, landfills are not simply giant holes into which trash is dumped and buried. Instead, a large tract of land—generally several hundred acres—is divided into smaller subunits known as **cells.** Each cell is on the order of 10 – 20 acres in size and managed separately. Depth varies with location and can reach hundreds of meters or more.

Cells are lined with impermeable layers

According to federal law[2], modern landfills in the United States

2. US Environmental Protection Agency (EPA). Municipal Solid Waste

must be bounded on all sides by layers of clay and plastic to minimize the movement of water through garbage and the leaching of toxins to groundwater (Chapter 11). See Figure 13.7 for a diagram of a typical landfill cell. This practice is effective in restricting the movement of pollutants, but it creates some challenges (explained in "Disadvantages of landfills", below).

Note that many landfills were built before laws regarding containment were in place. In some cases, such sites are responsible for substantial environmental contamination.

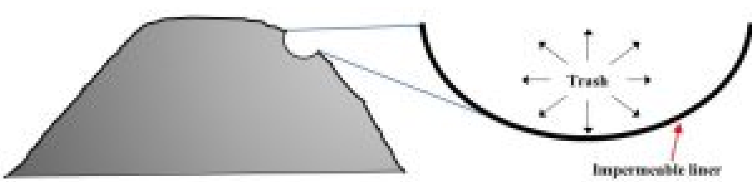


Figure 13.7. *Simplified diagram of a landfill. Left: the larger site in which multiple cells are constructed. Right: a single active cell is filled with trash. The surface marked "impermeable liner" is made up of plastic on top of clay. Both are in cross section. See main text for more details. Kelsey,* BY BY-NC-SA.

Materials are received and compressed to maximum density

A relatively small area within the landfill actively receives trash. At this location, known as the **working face,** trucks dump their

Landfills. https://www.epa.gov/landfills/municipal-solid-waste-landfills contents. Large vehicles (weighing 40 – 50 tons) then repeatedly drive over the garbage to spread it into even and highly compressed layers (Figure 13.8).

Figure 13.8. *A compactor, left (a) and the working face of a landfill, right (b). Ropable, Public Domain (a); Eric Guinther,* CC BY-SA.

Cells are capped when full

The elevation to which a landfill may grow is dictated by the permit granted to the responsible company. In other words, a limited mass of trash is allowed before a site is considered filled. So, once waste has been compacted as much as is possible, the top of a cell is lined with additional clay and plastic to seal it on all sides indefinitely. Soil and grass seeds are applied to the surface above the cap (grass provides stability and resistance to erosion), and the site is then monitored for changes or breaches that could lead to environmental contamination (Figure 13.9). A new cell is opened and the process repeated until the entire landfill site has been filled to capacity.

Advantages of landfills.

Waste is contained

Modern landfills sequester unwanted trash and generally protect human and environmental health. Without them, unsightly and toxic materials could damage terrestrial and aquatic systems and pollute drinking water. Importantly, they use simple and mature technology and are relatively cost effective (compared to other methods discussed below).

Methane can be collected to generate electricity

Oxygen availability inside landfills is extremely limited. Consequently, organisms that can grow in the absence of O_2 gas dominate these artificially created environments. Many anaerobic bacteria—particularly those thriving in landfills—release CH_4, **methane**, as a product of their normal metabolism (review anaerobic and aerobic respiration in Chapter 4). Arguably, methane production is both a disadvantage and an advantage of landfills as a waste management strategy. It is problematic because this explosive gas could accumulate to dangerous levels within sealed cells if not removed by a vacuum system (Figure 13.9.a,b).

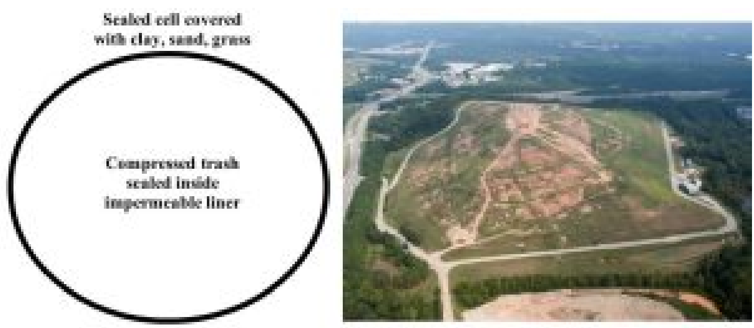


Figure 13.9. *Left (a): conceptual diagram of a sealed landfill cell (in cross section—compare to Figure 13.7). Right (b): a landfill made up of many sealed cells. Note the grass growing on top of older cells (aerial photo). Kelsey, CC BY-NC-SA (a); CarlisleEnergy, CC BY-SA (b).*

Historically, the collected gas was vented directly to the atmosphere, contributing to air pollution (Chapter 14), or burned at a large flare (Figure 13.10) (CO_2, the dominant product of this combustion, is a pollutant that poses a lower risk than does CH_4—more in Chapter 14).

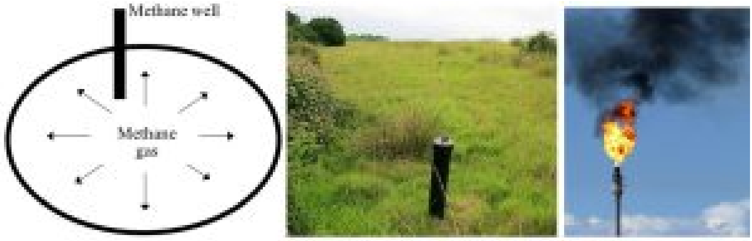


Figure 13.10. *Fates of methane generated by anaerobic decomposers in landfills. Simplified diagram of sealed cell with methane well, (cross section, not to scale) left (a); photo of methane well in a capped landfill (approximately 1 meter high), middle (b); photo of a methane flare (they range 5 – 10 meters high), right (c).* Kelsey, CC BY-NC-SA. *(a); Peter Facey,* CC BY-SA *(b); W.carter, Pubic Domain (c).*

On the plus side, combustion of the methane can be used to produce electricity (consult Chapter 10 for more about electricity generation). This process, an example of what is often called **waste-to-energy**, has become increasingly popular during recent decades. According to the U.S. EPA, about 75% of active landfills in the United States now use captured methane in power generation[3].

Disadvantages and challenges of landfills.

Decomposition is very slow

Natural **decomposition** that would otherwise break down and recycle waste is severely limited in landfill cells because the **aerobic** microorganisms capable of efficiently carrying out this function are deprived of necessary oxygen. Sealing garbage in water- and air-tight chambers minimizes water pollution, but it also creates an anoxic environment (i.e., lacks available O_2). So, much of the paper,

3. US Environmental Protection Agency (EPA). Basic Information about Landfill Gas. https://www.epa.gov/lmop/basic-information-about-landfill-gas

food, yard waste, cotton, and other such materials tend to persist in landfills for decades or centuries. If you were to tunnel deep into an old landfill (yes, some scientists do this willingly) you would very likely find many readily degradable items, for instance, newspapers and partially eaten hotdogs discarded in the early 1900s in near-pristine condition. In other words, trash is stored, not eliminated, in landfills.

Long-term containment and monitoring are required

The operator of a landfill is legally obligated to maintain the site for many decades after it has ceased receiving trash and been sealed. In addition to the methane collection system described above, special wells are drilled to detect potential leaching of toxins to **groundwater**. Overall site integrity is also monitored to protect human and environmental health.

Land use options are limited

Landfills take up an enormous amount of space. Sealed landfills may appear to be solid grassy mountains, but their surfaces are not appropriate for housing, agriculture, and the like because they change shape and lose stability as trash undergoes slow decomposition, shifts, and settles. In addition, concerns about contamination of water reservoirs and exposure of humans and animals to harmful materials substantially limit future use. Various recreational and industrial facilities—as well as **solar farms** (Chapter 10) have been constructed on top of closed landfills, although this practice is not commonplace (Figure 13.11).

Figure 13.11. *Examples of how surfaces of closed landfills may be used: left (a), public park in England; right (b), site on which large-scale solar energy collectors have been installed in Georgia, U.S.A. (aerial view). Tdorante10, CC BY-SA (a); CarlisleEnergy, CC BY-SA (b).*

Incineration

In this second approach, waste is combusted to reduce its mass and volume. Contrary to the vision some people have, this is a complex process that is quite distinct from open burning (review above).

Incinerator design and operation.

Waste is combusted at high temperature

Municipal solid waste (MSW) is burned inside specialized facilities that resemble large, closed furnaces. Importantly, these incinerators operate at temperatures high enough to maximize destruction of trash (in the range of 1400 °C or 2500 °F, far higher than, say, a fireplace or oven[4]).

Ash is collected and landfilled

Combustion does not destroy trash, rather, it greatly reduces the

4. Mass.gov. Municipal Waste Combustors. https://www.mass.gov/guides/municipal-waste-combustors#-about-combustion-facilities-

volume of MSW. Results vary, but a decrease of 85 – 95% is typical[5]. The leftover ash must still be managed, so it is generally transported to a sanitary landfill as described in the previous section.

Toxic emissions are minimized

Heavy metals, dioxin, particulates, and other hazardous substances are produced when MSW is burned. Early incinerators were not equipped to control the release of such toxins, but modern incinerators employ mechanisms to largely prevent the movement of harmful products into Earth's atmosphere.

Advantages of incineration.

Incineration reduces the need for landfill space

Incineration helps alleviate the problem of shrinking landfill availability because it dramatically reduces the volume of MSW needing to be managed.

Trash can be used as fuel to generate electricity

When MSW is burned, the heat released can be used to boil water, produce steam, and generate electricity (Chapter 10). Thus, solid waste incineration is another example of the waste-to-energy strategy described above.

Disadvantages and challenges of incineration.

Risk of air pollution

We saw earlier that the burning of solid waste produces many toxins. Since these substances threaten human and environmental health, public opposition to construction and operation of incinerators tends to be intense. It turns out that laws governing MSW combustion in the U.S. are rather strict, and new facilities use several chemical and physical methods to stay within regulatory

5. US Environmental Protection Agency (EPA). Combustion with Energy Recovery. https://archive.epa.gov/epawaste/nonhaz/municipal/web/html/basic.html

limits. Accidental releases are unlikely, but not impossible. On the other hand, emissions of carbon dioxide and water vapor are not controlled. Although trash incineration puts relatively small amounts of it into the atmosphere, CO_2 is of interest because of its role in climate change (Chapter 14).

Need for management of toxic solids

The incombustible ash left after incineration is generally sent to a landfill. Since it may contain highly concentrated toxins, it must be handled carefully.

High start-up costs

Incinerators are very expensive to build. According to the U.S. EPA, construction of a typical facility requires at least $100 million[6]. Revenue is generated from fees for services as well as electricity sold to power companies, however, recovery of the initial investment may take decades.

Recycling

Many products that are integral to modern life—things such as containers, books, magazines, clothing, electronic devices, vehicles, buildings, and so forth—are collected and reprocessed after they are no longer functional, and the materials that made them up are incorporated into new objects. This notion that resources can be salvaged and used again should remind you of our earlier discussions of the natural pathways that move materials within and among spheres of the Earth. Processes such as **fixation**, **consumption**, and decomposition drive the constant **translocation**, transformation, and repurposing of **atoms** throughout living and

6. US Environmental Protection Agency (EPA). Combustion with Energy Recovery. https://archive.epa.gov/epawaste/nonhaz/municipal/web/html/basic.html

non-living reservoirs (Chapter 4). As noted earlier, materials do not accumulate in natural systems and are not otherwise discarded. Instead, they are recycled. Although the principle behind this waste management strategy is simple, the practice of synthetic recycling is rather complex.

Overview of the process.

Initial separation and collection of materials

You have likely seen the various containers intended to receive different waste materials: recyclable paper, glass, metals, and certain plastics are isolated from non-recyclables such as food, soiled paper, Styrofoam, and certain (other) plastics. Waste management companies collect the waste and haul it to the appropriate processing facilities. Trash that will not be recycled is transported to a landfill or incinerator, whereas the recyclables move on to the next step in a long journey.

> *These distinct categories of MSW are often referred to as **waste streams**.*

Further separation of comingled recyclables

Materials must be separated from each other if they are to be efficiently recycled, and various strategies are used[7]. Typically, complex assembly lines made up of people picking objects off a conveyor belt, physical screens of varying sizes, and robotic devices separate papers, metals, plastics, and glasses from each other (Figures 13.12 and 13.13).

7. US Environmental Protection Agency (EPA). Reduce, Reuse, Recycle. https://www.epa.gov/recycle

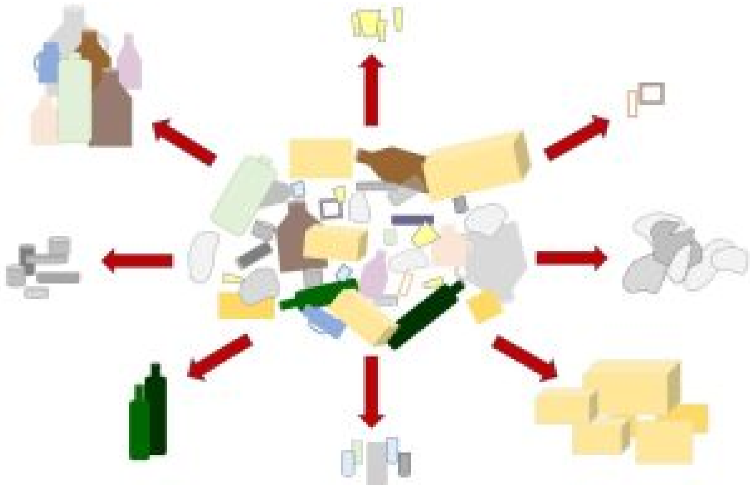


Figure 13.12. *Mixed collections of papers, cardboards, plastics, glasses, and metals must be separated into individual materials for recycling to be efficient and practical.* Kelsey, CC BY-NC-SA.

Figure 13.13. *Mixed recyclables move on conveyor belts as people sort them by hand. Automated sorting often follows the type of initial processing shown here.* US EPA, Public Domain.

Each of those groups can then be further sorted into distinct, specialized streams. For example, magnets can isolate steel-containing cans from other metals. The many types of plastics (which correspond to the familiar numbers on the bottoms of soda, milk, yogurt, and detergent containers) are distinguished from one another by passing them in single file past a beam of light; since different plastics affect the light in different ways, an automated processor can push each object into a separate pile of like materials. Glasses of different colors may be similarly sorted by optical properties. Papers and boxes are handled by a combination of manual and other strategies. Sorting of comingled recyclables is a crucial step but is clearly quite time consuming and labor intensive.

Sorted recyclable materials can be sold, reprocessed, and reused

Individual groups of paper, plastic, and so forth are compressed and baled (Figure 13.14) so they can be easily transported and sold.

Figure 13.14. *Once sorted, objects are compressed and baled in preparation for sale. Pictured here: bales of plastic bottles. Michal Maňas, CC BY.*

Like other commodities, a market of buyers and sellers influences their supply, demand, and price. These materials are used to manufacture a wide range of new products, but they are far from equally recyclable.

Glass. This is one of the few materials that we can reprocess and reuse an indefinite number of times with no appreciable degradation in its quality. Put another way, glass can repeatedly be melted down and serve the same function over and over.

Aluminum. Like glass, this metal is also highly reusable. Although some loss of quality does occur with each subsequent generation, it is fair to say that it can be made into new, functionally equivalent, products indefinitely. In fact, construction of new cans from recycled aluminum is easier and cheaper than it is from virgin material (the mineral bauxite) mined from the lithosphere.

Paper. Used paper is far less recyclable than glass or aluminum because its quality is substantially degraded with each generation: after six or so rounds, it simply can no longer maintain enough physical integrity to function adequately. Furthermore, the grade of products made from recycled paper diminishes with repeated recycling. Formal water-marked writing paper, for example, is generally made from fresh materials only. After this high-quality material is discarded, its reuse is limited to lower-grade newspaper and the like. Paper board and cardboard have similarly finite recyclability.

Textiles. Clothing and shoes can be recovered and reprocessed in different ways. Many of these items are cut up into rags or otherwise reused, but a very small fraction of the materials in textiles is converted to raw material to fabricate new articles[8].

Plastic. This is a more complicated story than that for other materials because the chemical and physical properties of different plastics vary so much (as noted above in the discussion of sorting). Although some can be reprocessed and used to construct new products like the old ones from which the materials were derived, most plastics cannot serve the same purpose more than one time. For instance, the plastic in food and drink containers is rarely made into new, equally usable bottles with equivalent properties, instead, it may be converted into unrelated items like jackets, furniture, and housing (Figure 13.15).

8. National Institute of Standards and Technology (NIST). 2022. Your Clothes can have an Afterlife. www.nist.gov/news-events/news/2022/05/your-clothes-can-have-afterlife

Figure 13.15. *Plastic from bottles and other containers can be recycled into new objects as shown here: left (a), a park bench in Portugal and, right (b), a house in Nigeria. Kolforn, CC BY-SA (a); Marlenenapoli, Public Domain (b).*

We will explore the challenges to effective plastic recycling in more detail shortly. In any case, research on plastics that can be better recycled is ongoing, and advancements could change the way plastic waste is managed.

Electronics. Mobile phones, computers, televisions, and related devices are built of multiple materials, including plastics, glass, and metals like lead, mercury, nickel, and zinc[9]. The last group—the metals—is particularly noteworthy. If recovered from electronic waste streams, these valuable materials can be sold to electronics manufacturers for reuse. Note, though, that collecting them from used devices is difficult and risky because many, like lead and mercury, are quite toxic to the humans that sort through them (Figure 13.16). We will return briefly to this topic near the end of the chapter.

9. US Environmental Protection Agency (EPA). Electronic Waste and Demolition. https://www.epa.gov/large-scale-residential-demolition/electronic-waste-and-demolition

Figure 13.16. *Electronic, or e waste, is often shipped from high-income nations to low-income nations: left (a), containers of discarded mobile phones and, right (b), workers recovering valuable, yet hazardous, metals from various electronic and other devices (both in Ghana). Fairphone, CC BY-SA (a); Marlenenapoli, CC0 (b).*

Advantages of recycling. Recycling provides many benefits. For example, materials that are recycled do not end up taking up space in landfills. The potential for pollution of soil, water, and air that can arise from mismanaged landfills and incineration is also reduced by recycling. Furthermore, recycling reduces the need to obtain new, or virgin, materials through processes such as mining (e.g., for aluminum) and logging (i.e., for paper). Rather than being discarded, materials can be reclaimed for future use.

> *Recycling is a way to extend our access to resources on a planet that is closed with respect to materials.*

Challenges of recycling.

Start-up costs and infrastructure

The processes described above require a sizable commitment from municipalities: money must be earmarked for construction, collection, transportation, separation, and maintenance of the many facilities needed. Although income can be generated from the sale of recycled materials, additional funding from taxes and grants is often needed to sustain recycling operations.

Contamination of waste streams

If you recall how the sorting strategies described above are intended to produce separate commodities that can be sold and recycled, you likely see the problem caused by contamination. Something like food waste in a batch of, say, recyclable paper, often renders the otherwise valuable material worthless. In such a case all the materials in a bundle would likely be landfilled or incinerated because the cost to clean the paper to an acceptable level is simply too high. In practical terms, educating and encouraging people to comply with labels like "Recyclables Only" is an important piece of this complex puzzle.

Manufacturing products so they can be recycled at the ends of their useful lives

Companies generally design, build, and market their goods with maximum utility and profitability in mind, but they do not always consider how to facilitate the recovery of materials that comprise them once these products cease to function. Among the ways recycling could be enhanced include simplified dismantling of complex objects like computers, cars, and the like and standardization of materials used within an industry. Furthermore, the problems of sorting and separation would be minimized if the number of different types of materials making up any manufactured product was reduced. Readers interested in learning more about overcoming design challenges should consult the books and articles written about these and related topics[10].

10. William McDonough has written much on these topics, including Cradle to Cradle: Remaking the Way We Make Things (McDonough and Braungart, 2008) and The Upcycle: Beyond Sustainability – Designing for Abundance (McDonough and Braungart, 2013).

Final words on recycling. Before moving on to the final waste-management strategy on our list, let's consider a few closing thoughts about recycling.

Reuse is not equivalent to recycling

Although these concepts do overlap, important differences between them should be noted. Simply put, if an object is repurposed without being broken apart and stripped of its raw materials, we will consider that to be *reuse*. So, refilling existing bottles, using old newspaper to line a birdcage, converting plastics or metal into artwork, altering existing clothing to produce new clothing, and even pretending a corrugated refrigerator box is a fort are all examples of reuse. On the other hand, when materials in existing objects are extensively reprocessed and used to construct completely new products, say when glass is melted and reformed into new jars, printer paper is shredded, pulped, and pressed into toilet paper, or plastic from two-liter bottles is chopped and reduced to pellets that are remolded into new containers, furniture, or clothing, then we use the term *recycling*. Of course, both strategies reduce the amount of waste entering landfills and incinerators, albeit in different ways.

Benches are not the same as bottles

We have already seen that not all materials are equally recyclable: some, like glass, can be reprocessed and reformed into new and equivalent objects whereas plastics (and some others) often cannot be used to construct objects that are the same as the source from which they were taken. The plastic recovered from beverage bottles, to recall an important example we noted previously, is generally used to make things such as park benches or warm-up jackets instead of products that serve the same purpose again. This practice of recycling materials into lower-quality goods is called **downcycling** by some because the materials are not able to accomplish the same task more than once. Importantly, downcycling necessitates large amounts of new raw materials to meet demand for high-grade containers and the like.

Alternatively, converting used materials into products that are the same as, or even superior to, the original, is often termed **upcycling**. This later outcome more closely resembles natural biogeochemical cycles (Chapter 4) in which carbon and other atoms can be truly recycled without loss in their ability to perform the same tasks over and over and is, therefore, generally more desired than the former. Keep in mind, though, that upcycling often requires a very large energy input and may come with prohibitively high economic and environmental costs.

Materials intended for recycling may end up elsewhere

Objects made of recyclable materials can get mixed into waste streams heading for landfills or incinerators due to negligence, ignorance, expedience, or simple mistakes. These recyclables can also enter waterways or other natural systems, thereby damaging ecosystems and human health.

Despite recycling, Earth STILL is closed with respect to materials

Recycling and reuse can shrink the need for landfill space and extraction of new raw materials, however, they do not entirely overcome the constraints posed by the finite nature of Earth's resources. These waste-management strategies certainly have an effect, but they alone will not shift us toward sustainability (Chapter 1). Arguably, for some people, recycling creates a perception of abundance and implies a license to use excess paper, plastic, glass, and so forth. The psychology of this phenomenon is clearly outside our study area, but it is context that the reader is encouraged to ponder. In any case, the need for less consumption—or, a lower rate of output in our familiar systems analysis model (Chapter 2)—will become ever-more urgent as standards of living and demands for goods and services rise worldwide. In short, there is good reason those three 'R' words, **Reduce**, **Reuse**, **Recycle,** are so often bound together.

Composting

We encountered **composting** in Chapter 9 when we studied agriculture (review Box 9.1 for more information). To remind you, natural processes are manipulated and harnessed to facilitate the rapid decomposition of plant remains from food, clothing, and paper into an organic-rich product that is very much like soil organic matter. It can be carried out in small, backyard bins or large, industrial-scale tanks that might be used by a restaurant, farm, or college dining hall (Figure 13.17).

Figure 13.17. *Composters range in size from large, like this facility in Edmonton, Canada, left (a), to small, like this backyard unit, right (b). Wasteman2009, Public Domain (a); Kelsey, CC BY-NC-SA (b).*

Bacteria and **fungi**, along with various **detritovores** (Chapter 5), quickly break down the scraps because the environmental conditions within composters are very favorable (e.g., high moisture, temperature, and nutrient availability). The product of this accelerated decomposition, **compost**, can then be applied to gardens and agricultural fields to increase soil fertility (Figure 13.18).

Figure 13.18. *Under the proper environmental conditions, decomposers and detritovores quickly convert scraps of food, paper, and related materials, left (a) into compost, a nutrient-rich product that resembles soil organic matter, right (b).* SuSanA, Secretariat CC BY *(a and b).*

Composting brings benefits and challenges. On the upside, it reduces the amount of material going to landfills, a particularly important outcome because even food waste will not readily decompose in a sealed cell (as noted above). Furthermore, natural processes can be used to recycle nutrients to support crop growth in a relatively inexpensive manner (Chapter 9). There are challenges and disadvantages, though. Among other problems, compliance of users is needed, space is required, offensive odors may be released, and rotting food can attract unwanted rodents. Consult Box 13.1 to read about another potential problem with composting.

Box 13.1. Compost-ABLE vs. Compost-ED

Compostable bowls, forks, cups, and the like are becoming increasingly common. They may sound brilliant, but you should keep in mind that they will only break down as advertised if they are carefully collected and transferred to a composting facility. If, however, they are buried in a conventional landfill, they will persist like the food and similar materials we encountered back in the description of sanitary landfills in 13.3.1. In other words, compostable waste will only become compost under appropriate environmental conditions.

Methods of MSW management are typically used in tandem

No single strategy is appropriate to manage all the waste humans generate, so an approach known as **Integrated Waste Management** (IWM) is often employed. With IWM, a municipality develops a customized plan that combines two or more of the methods described in the previous section to meet its needs. One city might use a combination of incineration and recycling whereas another could rely more on landfills, composting, and recycling. In short, plans differ from place to place (and might even evolve with time), but most endeavor to minimize both financial and environmental cost.

13.3.2. Hazardous waste

Since these materials are highly toxic, they cannot be sent to MSW landfills or incinerators and are, instead, separated and managed in facilities especially designed for them. The guiding principles are the same here, though, as we saw for municipal waste: protect human and environmental health in the most cost-effective way possible. Several high-profile cases of environmental contamination in the 1970s led to strict federal laws in the United States regarding hazardous waste (review Chapter 1, including Table 1.1). The Resource Conservation and Recovery Act (RCRA, 1976 and later amendments), for example, mandates so-called **cradle-to-grave management** of hazardous waste, meaning producers, transporters, and disposal facilities are required to track every step of the management journey of regulated chemicals. Laws such as RCRA are intended to protect human and environmental health by establishing uniform standards of handling and record keeping so responsible parties are held accountable in the case of accidental releases (and, to avoid another case of abandoned chemicals such as the one in **Love Canal** described in Chapter 1). Here we briefly consider some of the common strategies used to manage these risky materials.

Hazardous Waste Landfills

Some hazardous waste end ups in **secure landfills**, structures that are conceptually like sanitary landfills in that they are intended to contain toxins and protect water, humans, and natural ecosystems. However, these structures are more complex than the ones we saw above, containing extensive monitoring systems to detect leaks, multiple layers of concrete and other strong materials (unlike the plastic liners used in MSW landfills), and more durable construction to maintain their structural integrity indefinitely.

Incineration

As with MSW, some hazardous waste is burned in facilities that are specially designed to operate at very high temperature and capture toxic emissions.

Recycling

Some solvents and other hazardous materials can be processed and recycled, thus reducing the amount of materials going to landfills and incinerators. As it can be difficult and expensive to separate recyclables from the rest, this option can only be used in certain situations.

13.3.3. Biomedical waste

Since human fluids and tissues (and the plastics, glasses, metals, and other solids contaminated with them) have the potential to spread infectious diseases, they are kept separate from municipal waste. Biomedical waste, like hazardous waste, is subject to careful monitoring, containment, and disposal. Generally, materials are first sterilized by heat or other means and then burned in specially designed incinerators. The resulting ash may be sent to municipal landfills if it is deemed free of contaminants or stored in a more secure site if it still contains hazardous substances.

> *Infectious diseases can move among people and are caused by bacteria, protozoa, viruses, and other parasitic organisms we encountered in Chapter 3.*

13.4. CAN WE REDUCE WASTE?

It would be fair to say that **waste minimization**—the production of the smallest volume of unwanted products—it a goal of most stakeholders in the public and private sector. Some people take waste minimization a step further and seek to eliminate the need for landfills or incinerators. As its name suggests, the **Zero Waste Movement** aspires to a system in which 100% of our materials are recycled or reused indefinitely. Achieving its goal would necessitate adjustments to consumption habits and product design, as we considered above, and a transformation to a system modelled on Earth's natural processes (refer to the opening page of this chapter). The challenges impeding a move to zero waste are numerous, however, such a system would alleviate many of the problems associated with waste generation we have already considered, including finite landfill space, air and water pollution, and loss of useable materials that are discarded.

13.5. CONSEQUENCES OF MISMANAGEMENT

As we have implied throughout this chapter, a lack of appropriate and effective waste management can cause water, air, and soil pollution. Also worrying is the potential for direct exposure of human and non-human organisms to poisonous materials in various waste streams (more in Chapter 15). Animals scavenging in open dumps, as well as people trying to recover valuable metals in electronic waste, are among those that can be adversely affected by poor management of discarded materials.

We will close this chapter with a few words about a serious threat that has gained prominence in recent years. Littering, accidental releases, and intentional dumping have introduced great quantities of physical pollutants into oceans and other bodies of water. Plastics are particularly important because they tend to resist decomposition and persist in waterways indefinitely. Large floating debris can entangle, suffocate, or otherwise hinder organisms that encounter it. In addition, very small fragments—known as **microplastics**—can cause damage to individuals, ecosystems, and humans (see Box 13.2 for more about this growing problem).

Box 13.2. Large ocean, tiny particles

Microplastics—defined as objects between 1 nm[11] and 5 mm in any dimension[12]—end up in the oceans for a couple of reasons. First, large bottles and related objects can undergo some degradation in the intense sun along coastlines and start to break up into small pieces that wash into water (Figure 13.19.a,b).

Figure 13.19. *Plastic debris commonly accumulates on beaches, like this one in Hawaii, left (a); large pieces can break down into smaller particles that enter the ocean, right (b).* NOAA, *Public domain (a and b).*

11. nanometer, 0.000000001 meters

12. US EPA. Microplastics Research. epa.gov

Second, small plastic beads intentionally added to some consumer products like toothpaste and cosmetics pass through the sewage treatment plants we encountered in Chapter 11 (Figure 13.20).

Figure 13.20. *Microplastics intentionally added to toothpaste and cosmetics (pictured here next to a US dime for scale) are small enough to move through sewage treatment facilities and enter waterways.* US NOAA, *Public Domain.*

Although research on this phenomenon is ongoing, data suggest that these microplastics are often mistakenly seen as food by fish. So, what's the big deal with eating tiny debris? Among other concerns, consumption of small particles can fill an animal's gut with non-nutritive solids and lead to starvation. Furthermore, since some chemical pollutants (e.g., industrial and agricultural products) often bind to these plastics, exposure to toxins can increase in affected organisms. The adverse consequences experienced by an individual may then be passed to and magnified in consumers at higher trophic levels, including humans (review trophic interactions in Chapter 5)[13, 14, 15]. More about the movement of pollutants through food webs is presented in Chapter 15.

THE CHAPTER ESSENCE IN BRIEF[16]

Increasing population size, standards of living, and resource use have led to a growing need for effective waste management. Several adverse consequences result from a lack of appropriate measures, including the loss of discarded materials for future use, the filling

13. What are Microplastics? National Oceanographic and Atmospheric Administration (NOAA). https://oceanservice.noaa.gov/facts/microplastics.html

14. US Environmental Protection Agency (EPA). State of the Science White Paper: Effects of Plastics Pollution on Aquatic Life and Aquatic-Dependent Wildlife. EPA-822-R-16-009

15. www.epa.gov/trash-free-waters/state-science-white-paper-effects-plastics-pollution-aquatic-life-and-aquatic

16. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 13 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

of Earth's finite space with trash, and the contamination and degradation of natural systems. Several strategies are used to minimize the problems of waste, but all bring both costs and benefits.

Think about it some more...[17]

Thinking back to our definitions from Chapter 7, what are the important differences between the fates of natural waste and those of waste produced by human activities?

Why is "disposal" a misleading word to use in the context of waste management? Think about our fundamental principles in environmental science (Chapter 1) as you ponder your answer.

Why is containment of waste such a critical concern (think back to what you learned about groundwater in Chapters 4 and 11)?

17. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

Thinking like an environmental scientist and remembering the ways environmental conditions influence organisms, why is decomposition in landfills so slow?

What strategies could we use to improve recycling so it better mimics decomposition and reuse in natural ecosystems?

If you were in charge of making the relevant decisions, what would you do to reduce our waste management problems?

14. Air Pollution and Climate Change

JASON KELSEY

Many of the natural processes and human activities we have studied in previous chapters release pollutants into Earth's atmosphere. Here we examine several important consequences of these emissions. Note that some of the most politically charged topics in this textbook are part of Chapter 14, so we should be particularly careful to maintain our objectivity, seek evidence-based explanations, and stay firmly grounded in science. Various scientific arguments will be considered, but *un*scientific arguments will not receive equal attention (review Chapter 2 for more about science as a way of knowing).

Key concepts

After reading Chapter 14, you should understand the following:

- The many sources and types of atmospheric pollutants
- Why humans intentionally pollute the atmosphere
- The three scales of air pollution
- The causes and consequences of smog

- How acid precipitation forms, why it is of concern, and how it can be mitigated
- The evidence supporting the claim that human activity is changing global climate, why climate change is problematic, and how it could be combatted
- The causes and consequences of ozone depletion, a phenomenon largely separate from climate change
- How global cooperation among Earth's nations has allowed the ozone layer to partially recover
- The connections among fossil fuel combustion (Chapter 10) and outdoor air pollution
- How indoor environments are subject to air pollution

Chapter contents

14.1. AN INTRODUCTION TO ATMOSPHERIC POLLUTION

14.1.1. The atmosphere is dynamic

We begin with a brief reminder of the characteristics of the atmosphere. In simplest terms, it can be visualized as several layers of gases that envelop the solid planet. It is intimately linked to the **hydrosphere**, **biosphere**, and **lithosphere**, as the four spheres affect and are affected by each other. Gases here move and mix, and their composition has changed throughout history due to the activities of **terrestrial** and **aquatic** organisms. You are strongly encouraged to review the first several pages of Chapter 4, including Figure 4.1, for more information about the structure, chemistry, and processes active in the atmosphere.

14.1.2. Many substances can pollute the air

Air pollutants can be defined as substances that enter the atmosphere and lead to adverse consequences. Since their origins and properties are so varied, environmental scientists have established categories of pollutants based on answers to several simple questions.

Chemical, physical, or biological?

This categorization scheme groups substances based on their fundamental nature and the resultant types of risk they pose when in the atmosphere.

Chemical. Chiefly due to their chemical properties, these pollutants can directly or indirectly affect organisms and ecosystems.

Direct effects

Many substances are emitted from sources in forms that are immediately toxic. For example, mercury and sulfur dioxide from coal combustion can cause brain damage and lung irritation, respectively.

Indirect effects

Some substances are transformed to new chemical pollutants via reactions they undergo in the atmosphere. Sulfur dioxide, seen above as acting directly, also can react to form sulfuric acid and cause acid rain (described shortly). Alternatively, a substance in the air could be of concern not because it reacts to form a new chemical entity, but because its presence leads to important modifications to environmental conditions. **Carbon dioxide**, for example, is generally innocuous—we humans spend our lives in the constant presence of this gas without experiencing any adverse health effects because of it. However, if an unusually large amount is suddenly introduced into a system, it temporarily displaces O_2 such that all **aerobic** life forms in the affected region die from asphyxiation. A small number of volcanic events (Chapter 7) have killed herds of animals, as well as their human shepherds, through this unusual and deadly mechanism. On a much larger scale, CO_2 also influences the average temperature of the Earth. As we will learn in some detail near the end of this chapter, increases in the amount of carbon dioxide in the atmosphere can be linked to changes in climate across the planet. Again, the gas does not act like a poison *per se*, but it still has the potential to appreciably affect living and non-living systems.

Physical. These substances also act directly and indirectly on the biosphere. For example, certain particles (e.g., fibers, fine dust, soot) can directly damage lung tissues, whereas large clouds of dust (e.g., from a volcanic eruption) can partially block incoming sunlight, drive changes to climate and, thus, indirectly affect organisms.

Biological. Cells of microorganisms (Chapter 3) may be sent into the air by natural forces such as sneezing, coughing, and the crashing of waves onto rocks, or through anthropogenic activity like manure spreading for agriculture (Chapter 9). Additionally, fungi and many plants launch spores and other structures into the atmosphere as part of their reproduction. If inhaled, these biological entities can cause harm to animals, including diseases and allergic reactions.

Natural or anthropogenic?

Air pollutants can originate from natural processes like **decomposition** and volcanic eruptions as well as from agriculture (Chapter 9), fuel combustion (Chapter 10), waste management (Chapter 13), and other human activities. This distinction is important for two reasons. First, many assume that harmful emissions come exclusively from people, but air pollution is a phenomenon that predates human civilization. Second, governments can only enact laws and policies to regulate **anthropogenic** sources of emissions—it likely goes without saying that we cannot prosecute or fine volcanoes whenever they release excessive amounts of hazardous gases. The focus of this chapter will be human sources of pollutants because those are the ones over which we can exert some control.

Mobile or stationary?

Here we distinguish between sources that move around, such as cars and trucks, and those that stay in one place, such as coal-burning power plants and volcanos. Why do we care about this difference? It comes down to ease of monitoring and controlling. Since, by design and intention, vehicles travel from place to place, it can be difficult to track the pollutants they are releasing. Standards vary among countries and U.S. states, but it is safe to say that a car could produce an enormous amount of pollution, far more than is legal, before the problem is discovered. On the other hand, emissions from a fixed smokestack are relatively easy to test, meaning violations of regulations can be readily detected.

Primary or secondary?

This categorization scheme reflects whether a substance has been chemically transformed after release from its source. A **primary pollutant** is one that has not been altered from its original form, whereas a **secondary pollutant** is the product of reactions involving primary pollutants. Consider, for example, the exhaust from a car: it consists of several primary pollutants, including carbon dioxide, carbon monoxide, various **hydrocarbons**, and compounds containing nitrogen or sulfur. It turns out that each of these chemical substances is potentially problematic in its own right. However, on hot, sunny days, a secondary pollutant called ozone gas can be produced in reactions involving some of the components in the tailpipe emissions (Figure 14.1; we will also return to this point shortly).

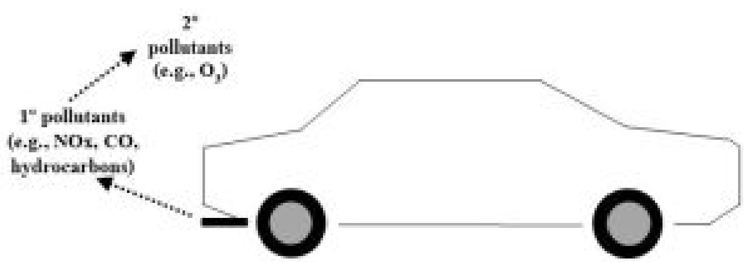


Figure 14.1. *Products of fossil fuel combustion enter the atmosphere as primary pollutants. Reactions in the atmosphere transform some of those primary pollutants into secondary pollutants. Kelsey, CC BY-NC-SA.*

As we saw above, the utility of these designations can be understood in the context of regulations. Laws about allowable amounts and types of primary pollutants from different sources are influenced by many factors, including the likelihood that secondary pollutants will be produced. These pollutant types play roles in upcoming sections of this chapter.

Indoor or outdoor?

Pollution is not a phenomenon affecting only open-air environments, as some of the most serious air quality problems are found in closed, interior spaces like factories, laboratories, and offices. As you likely expect, rules about the composition of indoor and outdoor air vary considerably. We will see specific examples of both types of air pollution later.

14.1.3. Pollutants are shared

Air moves fast

Everything on Earth is connected, as we have seen in many places throughout this book (review the **principle of environmental unity**, Chapter 1). However, since air is the fastest moving medium on this planet, atmospheric gases are more directly and obviously shared than are any other materials on the planet. In other words, substances emitted into the air can quickly spread from their sources and affect ecosystems and people thousands of kilometers away.

Air readily crosses political boundaries

A somewhat simple and obvious fact is critical to our discussion of air pollution: lines drawn on maps do not stop air and the pollutants carried in it from moving among cities, states, and countries. Consequently, even if a local government passes the most rigorous air quality regulations, it will still be subject to the pollutants emitted by its less careful neighbors (Figure 14.2).

Figure 14.2. *Air and the pollutants in it can move in the direction of the arrow without interference from borders among states and countries.* US NOAA, *Public Domain (modified by Kelsey).*

14.1.4. Erroneous assumptions about the atmosphere influence our actions

Humans have used the atmosphere as a receptacle for waste materials for centuries, despite knowing that polluted air is detrimental to human health. Why would we deliberately degrade such an important resource? The answer is driven by several assumptions, notions that are at least partially false.

Assumption 1: the atmosphere is so vast we cannot affect its composition

Recalling the adage "dilution is the solution to pollution" (Chapter 11), many people assume that human activity is incapable of appreciably changing the chemistry of Earth's enormous atmosphere. The fact is, though, the atmosphere is not infinitely able to dilute air pollutants. We will see that the concentrations of numerous substances released from anthropogenic activities have measurably increased since the latter part of the Nineteenth Century.

Assumption 2: fast-moving air will make pollutants go away

Air currents have the capacity to rapidly carry pollutants away from their sources. The assumption, then, is that the risk of poisonous substances disappears in the wind. Two important problems challenge the validity of this expectation, however. First, again recalling the discussion of water pollution in Chapter 11, "out of sight, out of mind" fails to account for the reality of a shared atmosphere. Second, as we will see when we explore the problem of urban smog, later, pollutants do not always move away as we hope.

Assumption 3: chemical reactions reduce the toxicity of air pollutants

It is certainly true that chemical pollutants can be transformed via naturally occurring **chemical reactions** in the atmosphere, but those reactions do not necessarily eliminate toxic substances. Sometimes, they make matters worse. Urban smog and acid precipitation are two examples of problems associated with products of reactions, that is, secondary pollutants.

Assumption 4: the risk of air pollution is greatly exaggerated

Finally, many people are comfortable emitting potentially toxic substances into the atmosphere because they simply do not accept the conclusions of scientists. Resistance to reducing or eliminating certain pollutants is often driven by denial of data and evidence, an issue that is particularly problematic in the public discourse about global climate change (which we will explore later). Moreover, the effects of air pollution are sometimes seen as remote, trivial, or just a vague concern for subsequent human generations. Convenience and expedience in the here and now (along with, frankly, ignorance) can outweigh concerns about future adverse consequences.

14.1.5. Effects of air pollution vary with scale

The effects of outdoor air pollution are grouped into one of three categories: **local** (i.e., felt in the immediate vicinity of a source), **regional** (i.e., within several hundred kilometers of a source), and **global** (i.e., the entire planet). Specific examples of each level will be presented in the next section. Before we proceed with that discussion, though, you should realize that a single source of pollution can bring on effects at all three scales and that the nature of the observed consequences will likely be different at those different scales. We do not simply see the same phenomena affecting increasingly large areas, instead, we can observe one type of problem at the local scale, another at the local scale, and a third at the global scale. For example, emissions from one coal-fired powerplant (Chapter 10) can cause smog at the local level, acid precipitation regionally, and contribute to global warming, ocean acidification, and planet-wide climate change.

14.2. SOURCES AND CONSEQUENCES OF OUTDOOR AIR POLLUTION

The bulk of this chapter will address several local, regional, and global consequences that arise from the emission of pollutants into Earth's atmosphere.

I: Local-scale effects

14.2.1. Smog

Local-scale effects are limited to areas close to air pollution sources. Of the phenomena that can be categorized here, we will consider only smog, an important problem seen in many urban areas.

What is smog?

The word "smog" was coined about a century ago as a contraction of "smoke" and "fog". Since it is a moist, smoky, and cloudy mix of substances that impairs breathing and vision, this word is reasonably descriptive, if inexact. These days we use a more rigorous definition of smog, one that consists of two parts.

It accumulates near pollution sources. Smog defies the expectation described above that air currents will carry pollutants away from their sources. Several forces can trap smog in an area, as we will see below.

It consists of primary and secondary pollutants. Emissions from motor vehicles, factories, and power plants, including sulfur and nitrogen gases, hydrocarbons, and particulates, make up the bulk of the primary pollutants in smog. Ozone gas, O_3, is a secondary pollutant derived from nitrogen dioxide (NO_2) and hydrocarbons in intense sunlight. Again, we hope chemical reactions will reduce or eliminate the hazards of air-borne toxic substances, but in this case, they generate a new danger.

Factors promoting smog formation

High source density. Smog is most likely to develop where air pollution sources are concentrated. Cities, with their dense populations, large number of fossil-fuel burning vehicles, and manufacturing and power-generating facilities, are susceptible to smog.

Topographical barriers. Air movement can be impeded by valley walls, mountains, and even steady wind blowing in one direction (Figure 14.3). As a result, some locations are simply more vulnerable to smog accumulation than are others.

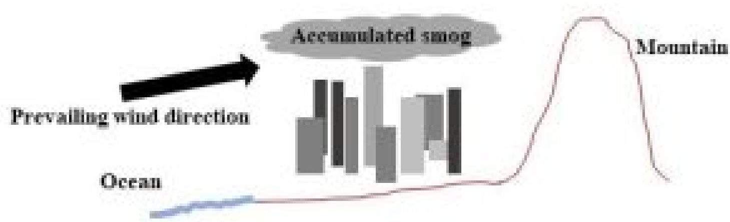


Figure 14.3. Emissions from a city can be trapped by a combination of barriers. Kelsey, CC BY-NC-SA.

Atmospheric inversions. These can restrict the movement of air pollutants away from sources and contribute to the problems of smog.

Vertical movement of air is inhibited

Air temperature usually decreases with altitude from the ground to the top of the troposphere (about 12 km high; review Figure 4.1). So, the warmest air—that at Earth's surface—carries air pollutants up as it expands and rises. Cooler air at higher altitudes then drops (since it is denser than the warm air). Sometimes, though, a layer of cool air can sink below warm air and remain at ground level. This abnormal condition, known as an **atmospheric** or **thermal inversion**, traps air pollutants near their sources (Figure 14.4). Smog continues to accumulate as long as an inversion remains in place and vertical mixing of air is prevented.

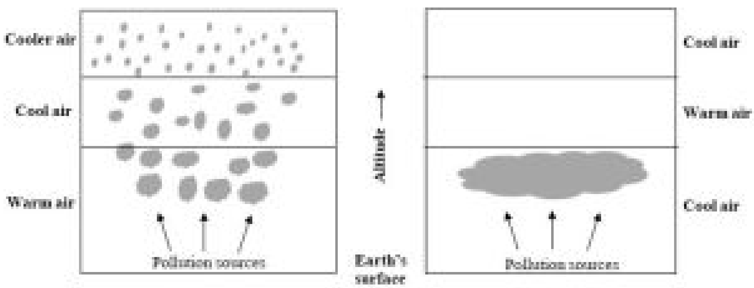


Figure 14.4. *Under normal conditions, air temperature declines with altitude and pollutants emitted from the surface move away (left). During a thermal inversion, a layer of relatively warm air is sandwiched between two layers of cooler air and pollutants are trapped near their sources (right). Kelsey,* CC BY-NC-SA.

Natural phenomena exacerbate anthropogenic activity

Thermal inversions occur naturally. For example, cooling under a clear night sky could cause cold air to sink and displace warm air sitting on the ground. A city located in an affected basin would then be subject to smog because the pollutants generated by its dense human population cannot disperse.

Appropriate climatic conditions. Since smog occurs under a specific set of atmospheric and climatic conditions, the severity and nature of it vary considerably with location, weather, and season.

Photochemical smog(Los Angeles smog)

This type of smog is favored in hot, sunny, and dry locations. It tends to form when primary pollutants from cars and trucks are released into air and react in intense sunlight to form secondary pollutants. If winds are minimal or air movement is otherwise impeded, this toxic mixture of particulates, hydrocarbons, nitrogen oxides, and ozone can be trapped over a city until a new weather pattern brings relief (individual episodes of intense smog have been known to persist for several days). Many cities around the world are susceptible to photochemical smog formation. For example, Los Angeles, California (U.S.A.) often experiences this local air pollution problem because it possesses all the necessary ingredients: high source density from automobiles, manufacturing, and power generation, topographical boundaries, and ideal climatic conditions (generally high temperatures, low moisture and cloud cover, and minimal wind) (Figure 14.5.a).

Figure 14.5.a. *Photochemical smog in Los Angeles, CA. Diliff, CC BY-SA.*

Mexico City (Mexico), Beijing (China), and Santiago (Chile) are just three of the many other places that regularly endure photochemical smog.

Sulfurous smog (London smog)

Named for a city susceptible to it, sulfurous differs from photochemical smog in two important ways.

It occurs in areas that are cold, overcast, and wet. In these cases, sunlight is partially blocked from reaching the surface by heavy cloud cover. Less sun penetration causes ground-level air temperature to decrease and that in the atmosphere, at and above the clouds, to increase. Eventually, water condenses in the cooling air, and fog forms at the surface. Any air pollutants present will be unable to move very far from their sources; instead, they will accumulate.

Coal emissions play a major role. Here, smog often consists of primary pollutants from coal combustion, namely soot, hydrocarbons, and sulfur oxides. Because the conditions are not conducive to its formation, ozone is a minor concern. However, another secondary pollutant, sulfuric acid, may be produced. Combining the previous two points, we can imagine how sulfurous smog can become a serious problem. In response to falling air temperatures, people burn more coal to generate heat. As emissions accumulate, air quality worsens until weather conditions change and the inversion is disrupted. London (England) is well known for its susceptibility to this type of smog, even as coal usage there has declined during the past several decades (Figure 14.5.b).

Figure 14.5.b. *Sulfurous smog is visible in London.* NT Stobbs, CC BY-SA.

Adverse effects of smog

Environmental health. Animals and plants can be inhibited or killed by exposure to smog. For example, **lichen** disappear from the sides of buildings in cities with serious smog problems. The loss of lichen in urban areas is often used as an indicator of poor air quality.

Human health. Several pollutants in smog are toxic to humans and can cause conditions ranging from mild irritation to death. Sulfur dioxide and ozone, for example, damage eyes and lungs and are deadly at high concentrations. Those with pre-existing respiratory conditions such as asthma are particularly vulnerable. During severe smog events, people generally avoid outside air as much as is possible.

What can be done about smog?

We will return to a broader discussion of air pollution prevention near the end of this chapter, but for now we will provide a simplistic, if reasonable, answer to the question: reduce emissions of primary pollutants such as those released during fossil fuel combustion.

II: Regional-scale effects

14.2.2. Acid precipitation

These effects are felt some distance away from the pollution sources that cause them. **Acid precipitation**, an important and familiar example of a regional-scale phenomenon, often occurs many hundreds of kilometers (or more) downwind of the sites from which the relevant emissions are released.

An introduction

Put simply, the name of this phenomenon says it all: air pollution causes rain, snow, ice, and fog to be more acidic (have a lower pH) than they would be otherwise. Dry materials can also be affected. Consult Box 14.1 for a short primer about the meaning and measurement of **acidity**.

Box 14.1. pH: how we measure and express acidity.

We already encountered the concept of acidity back in Chapter 9, but here we return to it in the context of air pollution.

The **pH scale** is used to express the acidity of water-based, or aqueous, liquids. In short, the amount of positively charged hydrogen **ions** (H+) present is measured and that number is converted to pH units. **Acids** have the *highest* concentrations of H+, and, because of the nature of the mathematical transformation used, the *lowest* pH values. **Bases**, on the other hand, contain more of a different ion, OH- (and lower amounts of H+), and have higher pH values. Although it can go beyond these limits, the scale usually is assumed to start at 0 (most acidic) and run to 14 (most basic). A liquid that is neither acidic nor basic, **neutral,** has a pH of 7 (Figure 14.6 shows the pH scale and includes values for some common substances).

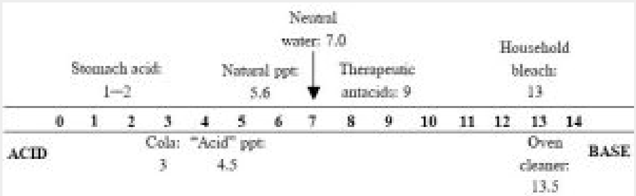


Figure 14.6. *The pH scale. Each number represents a 10-fold change in acidity—see the text in this box for a full explanation.* *Kelsey*, CC BY-NC-SA.

You should be aware of three critical facts related to the pH scale. First, the concentration of H+ ions does not change in a linear fashion, despite what Figure 14.6 suggests. Instead, each number represents a ten-fold change in acidity! Thus, the concentration of H+ in a liquid with a pH of 6 is *ten times* higher than that in a liquid having a pH of 7, pH 5 is ten times more acidic than pH 6, and so on. Furthermore, since each step in the scale multiples the difference in acidity by ten times, pH 3 indicates a concentration of H+ ions that is 1000 times higher, or is a thousand-fold more acidic than, pH 6 (10X for each step from 6 to 5 to 4 to 3). A change from, say, pH 5.5 to pH 4.5, is therefore more significant than it might appear. Second, when an acidic substance contacts a basic one, a reaction between the two will yield a more neutral (closer to pH 7) product. Anyone who has ever tried to relieve symptoms of heartburn by chewing antacids has experienced such a reaction firsthand: stomach acid escaping into the esophagus can be neutralized, albeit temporarily, by swallowing calcium carbonate or another basic substance. Finally, both acids and bases are corrosive—the further a substance is from pH 7, the more hazardous it is.

Origins and development of acid precipitation

Figure 14.7. Acid precursors. USGS, Public Domain.

Acid precursors are emitted. The formation of acid precipitation begins with the release of certain primary pollutants into the atmosphere. Sulfur-oxygen compounds like SO_2 emitted from coal-fired power plants (Figure 14.7), and nitrogen-oxygen compounds (i.e., "NOx") from gasoline, diesel, and coal burning (Chapter 10), are two important **acid precursors.** Natural sources of precursors, including volcanoes, also contribute.

Chemical reactions produce acids from acid precursors. Once in the air, the S and N gases can be converted in reactions with atmospheric water to nitric acid or sulfuric acid, respectively (Reactions 14.1, 14.2a and 14.2b).

(Reaction 14.1) $2NO_2 + H_2O \rightarrow HNO_2 + HNO_3$ ***(nitric acid)***

(Reaction 14.2.a) $SO_2 + H_2O \rightarrow H_2SO_3$ ***then,***

(Reaction 14.2.b) $2H_2SO_3 + O_2 \rightarrow 2H_2SO_4$ ***(sulfuric acid)***

Both nitric and sulfuric acid dissolve into precipitation and lower its pH. You should realize that, due to chemical conditions in the atmosphere, unpolluted natural rainwater has a pH of about 5.6; in other words, it is already somewhat acidic. To be classified as true acid precipitation, pH has to fall to 4.0 or lower (take another look at Figure 14.6, above). Figure 14.8 provides a graphical overview of the steps that lead to this regional air pollution problem.

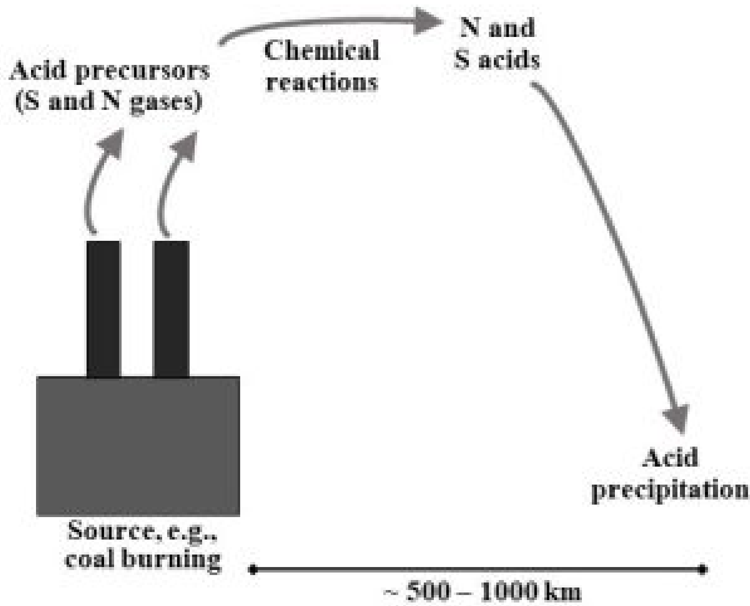


Figure 14.8. *An overview of the process from emission of acid precursors to the formation of acid precipitation. See main text for details. Kelsey,* CC BY-NC-SA.

Potential adverse effects of acid precipitation

Several consequences can be caused by acid precipitation. As we will see in the next section, though, the severity of the described effects varies from place to place.

Environmental health. Because it can dramatically alter the chemistry of soil and water, acidic precipitation may influence the **structure** and **function** of ecosystems (Chapter 5). If the changes are moderate, acid-tolerant species can gain an advantage over those that have adapted to higher (and what we could reasonably consider to be normal) pH. More severe acidification may create environmental conditions that few organisms survive.

Terrestrial

Forests, grasslands, farms, and other land-based systems rely on soil for nutrients, water, and physical support (Chapters 5 and 9). Changes to soil pH can reduce the success of plants and pose a challenge to all organisms present in an affected area. Importantly, the acids do not tend to harm plants directly, rather they introduce stress by disrupting the physical structure of soil, initiating reactions that increase the **biological availability** of toxic metals (e.g., aluminum), and decreasing the amount of essential nutrients present. Forests populated by trees that were otherwise well adapted to their environments and able to withstand stressors such as herbivory (i.e., consumption by plant-eating animals) and diseases have been severely damaged following substantial increases in the acidity of precipitation.

Portions of southern Germany and southeastern Canada, along with northeastern states such as New York and Vermont (U.S.A.), are among the places in which acid-related forest damage is prevalent (Figure 14.9). Production of crops from agricultural fields (Chapter 9) can also be reduced in acidified soils.

Figure 14.9. *A forest in the Czech Republic affected by acid precipitation. Lovecz, Public Domain.*

Aquatic

Acid precipitation can lower the pH of lakes and rivers to levels that are quite hostile to native organisms. Increased acidity can directly damage anything not adapted to such conditions, but indirect effects are at least as important. For example, the extra mucous some fish produce in response to stress can interfere with oxygen uptake through gills and even kill through asphyxiation. Acids can also lead to loss of nutrients (through chemical reactions like those seen in soils, above), inhibiting **primary producers**. Aquatic communities in the Adirondack Mountains of New York (U.S.A) provide just one example of the potential damage this regional phenomenon can cause. Historically, precursors released from coal power plants in places such as Ohio and Pennsylvania moved northeast and acidified precipitation in upstate New York. The pH of lakes fell precipitously, leaving many visibly barren. Certain rivers in Oslo (Norway) have undergone similar chemical changes, leading to declines in the numbers of important species such as salmon.

Susceptibility to acid precipitation varies

Acid precipitation does not lead to the same consequences everywhere. Due to several environmental factors, including a property known as buffering capacity, separate locations receiving the same acidified rainfall may respond quite differently. Put simply, **buffering capacity** is the amount of acid or base a material can absorb before its pH changes. So, something that can withstand the addition of large amounts of acid without itself become appreciably more acidic is said to have *high* buffering capacity. *Low* buffering capacity would be the term used to describe something that is very sensitive to (i.e., easily changed by) a small amount of acid. Limestone and related rocks are largely comprised of the compound calcium carbonate, an important natural buffer. If soil in a forest, field, or other terrestrial system was derived from such materials (see Chapter 9 for information about the contributions rocks make to soil formation), its pH will remain relatively constant even as it receives large amounts of acidic precipitation. A body of water can also be stabilized if its sediments contain calcium carbonate or similar compounds. Since granite, sandstone, and most other common rocks have very low buffering capacity, the pH of systems containing them, or products of their weathering, are easily lowered by incoming acids. Interestingly, the pH of water within the same stream can differ substantially from place to place—within just a few km—as the types of rock underlying it change. The Bushkill Creek in Eastern Pennsylvania (U.S.A.) provides a good example of this phenomenon, with its pH rising from approximately 6.5 to 7.5 as it flows a mere 20 km. Finally, keep in mind that, although resistance to pH changes can provide a benefit to humans, high buffering capacity may increase the numbers of sinkholes in an area (see Box 14.2 for more).

Box 14.2. Getting swallowed by the Earth: a price of high buffering capacity?

There can be a downside to the presence of good natural buffers like limestone in—and more importantly, under—an area: the neutralization reactions of the acids in infiltrating water with the bases in the solid rocks providing support for the surface can create big problems. Eventually, if enough of the underlying limestone is dissolved, the ground can collapse into depressions called **sinkholes** (Figure 14.10a. and 14.10.b).

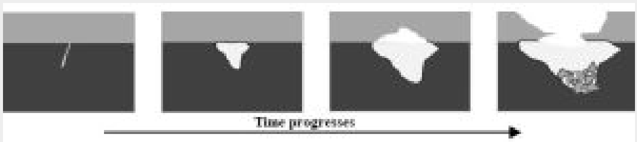


Figure 14.10.a. *As the underlying limestone (black) is dissolved by acids, support for the surface of the Earth is slowly lost; eventually, a sinkhole opens. In cross section. Kelsey,* CC BY-NC-SA.

Figure 14.10.b. A *sinkhole in the front yard of a house in Florida (U.S.A.). USGS, Public Domain.*

Sinkholes range in size from less than a meter across and a few centimeters deep to hundreds of meters in all dimensions. Sometimes they appear suddenly, imperiling unfortunate people in vehicles, buildings, or just standing at the wrong place at the wrong time. In other cases, they open gradually, allowing enough time for evacuations. You should be aware that many natural and anthropogenic activities (including unsustainable usage of groundwater—see Chapter 11) can bring about sinkholes, with acid-base reactions just one common mechanism. It may be possible to fix these holes using one of several approaches. The enormous levels of destruction and human suffering that can be caused by sinkholes should remind us yet again of the value of objective study and science: if decisions about land use, including where and how to build structures, roads, and bridges, are made without consideration of available knowledge and data, the consequences can be expensive and even deadly.

Responding to the adverse effects of acid precipitation

Humans can mitigate some of the consequences of acid precipitation in both natural and managed systems by taking advantage of the acid-base neutralization reactions described in Box 14.1, above. For example, a white, powdery substance commonly known as **lime** is often applied to acidified agricultural fields. The calcium hydroxide in this material reacts with acids, pushing soil pH up toward a more neutral value. Although it adds yet another task (and expense) to the life of a farmer, regular liming between growing seasons helps maintain crop yields that would otherwise be limited by acid-induced loses in soil fertility (Figure 14.11).

Figure 14.11. *The effect of liming on crop growth is shown in this experimental field. The soil was naturally very acidic, but lime was added to the soil in the background to raise its pH; nothing was added to the soil in the foreground. Plants were far more successful in the more basic soil. Alandmanson,* CC BY-SA.

Reducing acid precipitation

As noted at the end of the section on smog, we will take up the larger topic of reducing air pollution, including emissions of **acid precursors**, near the end of the chapter.

III: Global-scale effects

Now we consider two consequences of air pollution that affect the entire Earth, in areas both close to and remote from emissions sources: climate change and ozone depletion.

14.2.3. Global climate change

The first item on our list is familiar to most people living in today's world because it is discussed so widely. Furthermore, it is a controversial and divisive issue, eliciting intense, often values-based responses, from many people—including those with expertise in areas *other* than science. The reasons for the high levels of passion surrounding global climate change are numerous and include the fact that the problem is complex and difficult to understand, the potential consequences of it are enormous and dangerous, and the solutions we must embrace to reverse or minimize it range from inconvenient to very unpleasant. Here we will largely steer clear of politics and opinions about what *we would like to be true* about Earth's natural systems and our relationships to them and examine the *science* of climate change.

We begin with the name: global warming or climate change?

Right off the bat, the name used to describe this phenomenon is complicated and, unfortunately, a source of some confusion and even suspicion. When it first became a well-known issue over three decades ago, this problem was generally called "warming" because it referred to an increase in the average temperature of the entire Earth. In more recent years, scientists have called it "climate change", as that term more accurately reflects the relevance and consequences of global warming. In other words: we are still studying, and concerned about, changes in the planet's temperature. However, it is critically important to realize that global warming does *not* cause the temperature in every location on Earth to go up. Instead, due to an overall higher temperature, individual climates across the globe shift and change. Although most areas would indeed become hotter, others would actually get colder. Precipitation patterns would also change, increasing or decreasing water availability in different ways in different places and ecosystems. In short, the problem is far more nuanced than is reflected in the original name. Some people who do not accept that climate change is a real concern (we will examine their arguments near the end of this section) point to this new name as a sign that the science of climate change is unreliable and invalid rather than an attempt to better describe the problem.

Questions of interest to scientists

The essence of this global-scale air pollution can be captured in four essential questions.

#1: Is Earth getting warmer? The very concise, and evidence-based, answer to this question is: yes. Widely accepted data indicate that the *average* temperature of this planet has gone up by about 1.1°C (around 2° F) since the pre-industrial age (defined as 1880 – 1900)[1], but the rate of change has been neither uniform nor constant. Notably, most of the observed increase occurred during the past 35 years, with the warmest years on record (in descending order) being 2023, 2016, 2020, 2019, 2015, 2017, 2018, 2014, 2010, 2013 and 2005 (tied)[2],[3]; interestingly, the hottest single days on record

1. NASA 2022. Global climate change, vital signs of the planet. www.ncei.noaa.gov/news/global-climate-202112 AND Climate.gov. 2022. www.climate.gov/news-features/understanding-climate/climate-change-global-temperature

2. National Oceanographic and Atmospheric Administration (NOAA). 2023. www.noaa.gov/news/2022-was-worlds-6th-warmest-year-on-record

3. 2024. World Meteorological Organization (WMO)

occurred during July of 2023)[4]. Taking a somewhat longer view, most of the years 1900–1977 were *cooler* than the average for the 20th Century; since then, every year has been *warmer* than that same average (Figure 14.12).

> *Keep in mind that the averages were calculated from many places on the globe—of course, some were warmer, and some cooler, than the average.*

4. https://climatereanalyzer.org/clim/t2_daily/

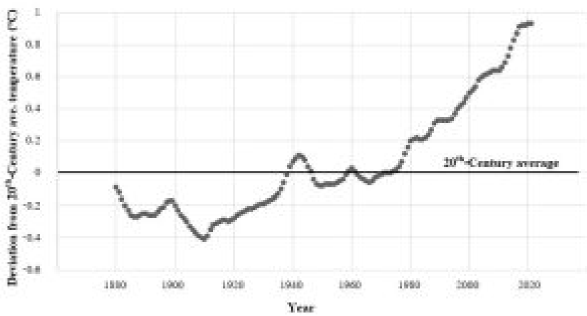


Figure 14.12. *Annual temperatures of Earth 1880-2021. The average temperature of the 20th Century is shown as a horizontal line at 0 °C, and the value for each year is plotted relative to it. For example, the temperature in 1880 was 0.11 °C below the average, 1951 was equal to the average, and 2020 was 0.97 °C above the average. Data from* US NOAA, *Public Domain; figure by* Kelsey, CC BY-NC-SA.

Will this warming trend continue? We will see more about making predictions below, but for now, careful analysis of past, current, and expected future patterns points to **an average global increase of 1.1 – 5.4 ºC (2 – 9.7 ºF) by the year 2100**[5].

#2: If the answer to question #1 is "yes", what is the cause? Although a small number of people argue otherwise, the temperature of the Earth clearly has gone up and appears likely to continue along that trajectory into the future. The data are reliable and difficult to dispute. With that first of our four questions relatively easily handled, then, we move on to our second, somewhat more difficult question. Given our observations that the average temperature on Earth has been going up for more than a century, it stands to reason that the rate at which energy enters this system currently outpaces the rate at which it exits. That is, either the amount of energy received by the planet has increased since 1880 or the amount of energy retained has increased during the same period. So, which has changed, inputs or outputs? To answer, we must examine three major factors that control Earth's energy balance.

The sun's energy output

You might be surprised to learn that the sun does not release a constant amount of energy. Instead, it cycles through periods of relatively high and low output. Decades of scientific study have revealed that visible changes to its surface are directly linked to the amount of energy it releases, with features known as **sunspots** coming and going as output waxes and wanes, respectively (Figure 14.13).

5. NASA 2022. Global climate change, vital signs of the planet. www.ncei.noaa.gov/news/global-climate-202112 AND Climate.gov. 2022. www.climate.gov/news-features/understanding-climate/climate-change-global-temperature

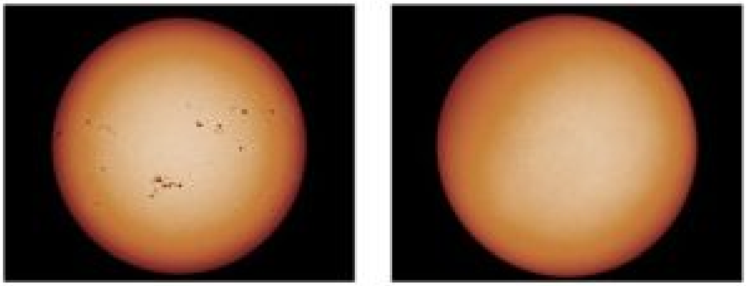


Figure 14.13. *The number of visible sunspots changes with time. The dark patches on the sun's surface (left) seen during a maximum coincide with increased energy release. No spots are visible during a minimum (right).* US NASA, *Public Domain.*

With the proper type of telescope, you can safely see and count the dark patches present at any moment. People have been doing this for centuries, and their data show something unmistakable: the cycles are regular, with the highest numbers of sunspots (referred to as **maximum sunspot activity**) appearing every 11 years. Figure 14.14.a shows the record of sunspot activity humans have kept since the early 1600s (note that modern scientists consider data from the first century of collection to be less reliable than those collected since), and Figure 14.14.b. shows the most recent data and a prediction for the next several years.

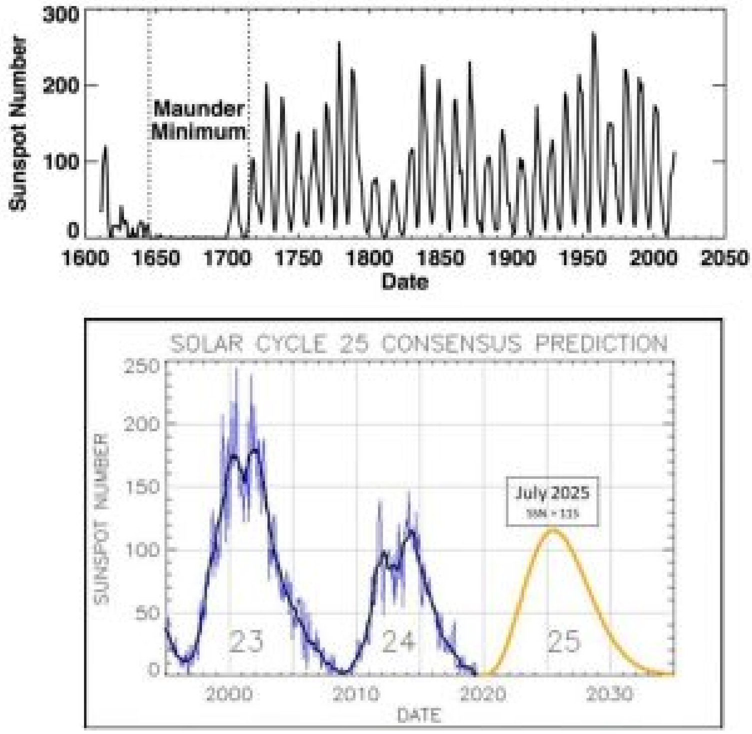


Figure 14.14. *Sunspot cycles. Upper (a): the record of sunspot activity since the 1600s. Note the regularity of the cycle, with peaks (or valleys) separated by 11 years. The absolute number of spots varies from cycle to cycle. The Maunder Minimum was a period during which sunspot counts were far lower than usual and Earth cooled somewhat. Lower (b): most recent data showing minimum number in about 2019 and a prediction of what is expected to occur during the upcoming cycle.* US NASA (a); US NOAA (b), *both Public Domain.*

Fluctuations in the sun's output are relevant to us because they might be responsible for the increased average temperature described in our first question, above. If recent changes in Earth's temperature have been exclusively caused by changes in the sun, then the answer to the second question is "global warming is the product of natural forces over which humans have no control", that is, we would be powerless to do anything about it except adapt.

Indeed, some who deny the reality of human-caused climate change make such an argument. Although there is some on-going discussion of this point among scientists, most hold that the data simply do not support the sun-as-cause hypothesis for a few reasons. First, the Earth's temperature has been going up since about 1880, and the rate of that increase has accelerated during the past few decades despite the ups and downs of energy received from the sun. Second, data from satellite sensors indicate no net change in the total solar energy received by Earth since 1978—even as Earth's temperature has gone up (Figure 14.15).

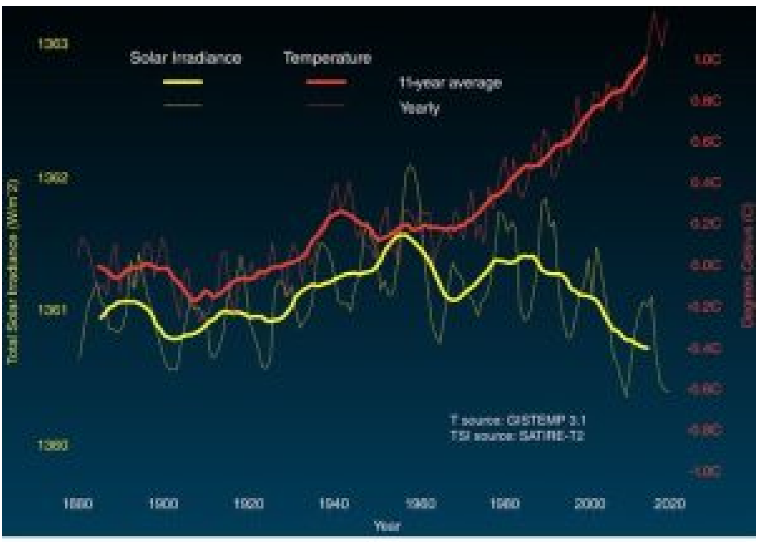


Figure 14.15. *Solar energy received by Earth (yellow) has declined in recent decades while the temperature of Earth (red) has increased. Thinner curves represent yearly levels linked to the sunspot cycle, and heavier lines show averages and overall trend.* US NASA, *Public Domain.*

Third, sun-driven warming should affect all layers of the atmosphere. However, evidence shows that only the inner atmosphere and the surface are warming (in fact, the stratosphere has been *cooling* recently), an outcome more consistent with an increase in the planet's greenhouse effect (described below)[6]. Does a change in energy received affect Earth's climate? It would be hard to imagine otherwise. However, forces that exert an even greater effect on temperature seem to be overwhelming fluctuating solar activity. Put succinctly, something else must be responsible for rising temperatures.

Earth's albedo

The amount of incoming solar energy reflected off the Earth, its **albedo**, is a second potential factor controlling temperature. As we saw in Chapter 4, the color of a surface affects how much energy it absorbs: broadly speaking, lighter colors reflect more than do darker ones. Ice and snow absorb less energy than soil, dark rocks, and asphalt (green forests are somewhere between those extremes). So, we would expect this planet's average temperature to be influenced by natural and human activities that alter surface characteristics. Below is a short list of specific phenomena that can change how much incoming solar energy is reflected off the planet to space.

Volcanic activity. We learned in Chapter 7 that eruptions release molten and solid rock, ash, and several gases into the air. Ash—a mixture of very small and light particles that can be distributed across the globe by wind and then stays aloft for years—is particularly important in this current context because it can block a fraction of sunlight reaching Earth. Major volcanic eruptions have brought about measurable cooling in their aftermaths throughout history. The 1991 eruption of Mt. Pinatubo in the Philippines serves as one example of the effect of volcanoes on albedo. The massive amount of ash and other materials ejected into the air was spread throughout the entire stratosphere within a year of the explosion and lowered average temperatures by about 0.6 °C for the next

6. NASA. 2022. Is the Sun causing global warming?. climate.nasa.gov/faq/14/is-the-sun-causing-global-warming/

15 months. Once the materials dispersed and settled, the cooling effect disappeared[7].

Cloud cover. Water vapor in the atmosphere can interact with dust particles to form clouds, and these clouds can reduce absorption—that is, increase albedo—because they block some sunlight entering the atmosphere. Changes in the amount of cloud cover have the potential to affect the amount of energy reaching the Earth and its average temperature. When we discuss variables affecting the way scientists are able to predict the future of global warming, we will return to the subject of clouds.

Anthropogenic air pollution. Certain emissions from human activity can affect albedo. For example, sulfur compounds and solid particles produced by coal combustion—sometimes referred to as **sulfate aerosols**—can enter the atmosphere and block some incoming radiation (Figure 14.16).

7. NASA. 2001. Global Effects of Mount Pinutubo. earthobservatory.nasa.gov/images/1510/global-effects-of-mount-pinatubo

In a somewhat ironic and unintended outcome, efforts to clean up emissions from coal-fired power plants during the past few decades have led to a decrease in the cooling effects caused by particulate pollution. Few would argue that more toxic air on the local scale is therefore a good thing, but it is, nonetheless, a fact that cannot be ignored.

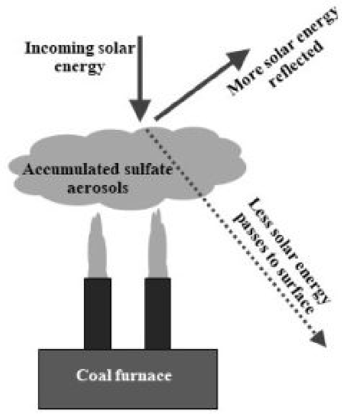


Figure 14.16. *Sulfate aerosols partially block incoming sunlight and cool the surface. Kelsey*, CC BY-NC-SA.

Land use changes. The spread of human civilization and advances in technology and standards of living (Chapter 8) have altered Earth's surface during the past millennia. Large-scale changes such as the building of cities and conversions of hundreds of thousands of acres of forests to farms are two of the important ways we could change surface colors from light to dark and reduce albedo. The extent to which these changes have appreciably lowered albedo is unclear, though, and further study is underway to improve our knowledge.

To conclude this short discussion of albedo, we should acknowledge that the relative importance of changes in energy reflection to the larger story of contemporary global warming is not well understood. The phenomena described above clearly play roles, but whether they are strongly linked to the temperature changes observed since 1880 is by no means clear. Worth noting, however, is a NASA monitoring program that has so far revealed no compelling evidence for long-term change in global albedo[8],

8. NASA. 2014. Measuring Earth's Albedo. earthobservatory.nasa.gov/images/84499/measuring-earths-albedo

despite the well-documented, widespread clearing of land, expansion of agriculture, and urbanization that have occurred while average temperatures steadily went up. In any case, scientists continue to study the problem.

Heat retained by the atmosphere

Commonly known as the **greenhouse effect**, this final factor is extremely important to the living things on Earth as well as to the process of global warming. Here we look at this large, complex process, in small, relatively simple pieces.

The atmosphere is transparent to some incoming solar energy. Earth receives several types of energy, known as **electromagnetic radiation**, from the sun. These different types are distinguished from each other by a property known as **wavelength**, the distance between the peaks or valleys of the travelling energy waves (Figure 14.17a). As represented by **the electromagnetic spectrum** (Figure 14.17b), wavelengths range from very short (about 0.000000000001 meters, m) to very long (about 10 m). Much of the energy emitted by the sun is in the range of 400 – 700 nanometers (nm, 0.000000001 m). These wavelengths can be seen by the human eye and are generally referred to as **visible light**. It is energy in this limited, visible range (Figure 14.17c), that easily passes through the gases in the atmosphere.

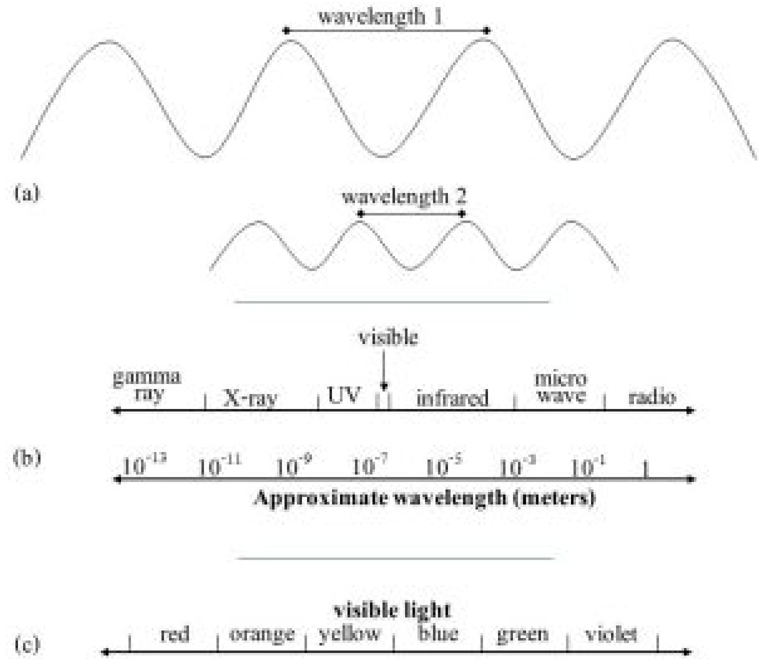


Figure 14.17. *Wavelength is defined as the distance between two peaks or valleys of travelling energy waves. In this hypothetical example, 1 is said to be longer than 2 (a). The electromagnetic spectrum is a representation of many types of energy arranged from shortest to longest wavelengths. Energy types having different wavelengths behave differently (b). The visible spectrum is a relatively narrow band of wavelengths between* 10^{-7} *and* 10^{-6} *meters; it is magnified at the base of the figure to reveal the details of the colors humans can see (c). Note that scientific notation is used to express wavelengths due to the enormous range from smallest to largest. By convention, 10 is multiplied by the exponent to give the number of meters, so* 10^{-9} = 0.000000009 *m and* 10^{-1} = 0.1 *m. Kelsey,* CC BY-NC-SA.

Incoming energy is converted to heat. Light energy striking the surface is converted into longer-wave infrared (IR), or heat energy, and then radiated up and outward toward space.

The atmosphere is relatively opaque to outgoing heat energy. Because longer-wavelength IR energy has different properties than light energy, it cannot easily pass though the atmosphere. Instead, the heat radiated from the surface is absorbed by certain **tropospheric** gases (called **greenhouse gases,** or **GHGs**) and then re-emitted in all directions—including back toward the Earth (Figure 14.18). We will return to the subject of GHGs, including their sources and concentrations in the atmosphere, shortly.

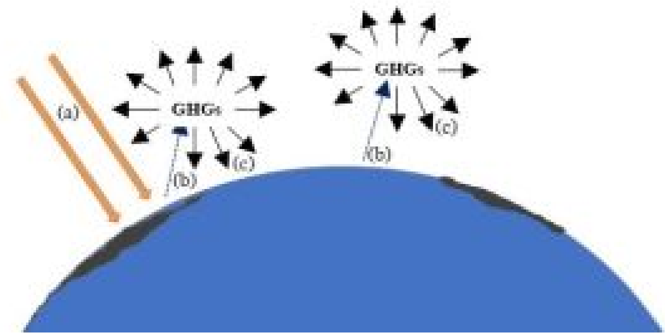


Figure 14.18. *A simplified diagram of the greenhouse effect. Incoming short-wave solar radiation (a) easily passes through the greenhouse gases, GHGs, in the atmosphere and reaches Earth's surface. Then, it is converted into heat energy (or infrared, IR) and emitted back toward space (b). The longer-wave IR behaves differently than the incoming radiation and does not easily pass through the atmosphere; instead, much of it is absorbed by GHGs and then re-radiated in all directions (c). A portion of the IR is re-radiated back to the surface and raises its temperature higher than it would be in the absence of the greenhouse effect. In cross section. Kelsey, CC BY-NC-SA.*

The greenhouse effect is natural and life giving. Although sometimes mistaken to be both a creation of and menace to humans, at the most basic level, the greenhouse effect is neither. Natural processes produced, shaped, and maintained it for many millions of years, long before any people appeared on Earth. We will see in upcoming discussions that the greenhouse effect can be *altered* by humans, but it was certainly not *created* through anthropogenic activity. Furthermore, the greenhouse effect is not a bad thing! In fact, because GHGs absorb and re-radiate heat, life as we know it can live here. Notably, the greenhouse effect increases the average temperature of the surface by about 33 °C[9], meaning that water is present in all three phases (solid, liquid, and vapor). Earth would be far too cold without it for its current biosphere to exist.

The greenhouse effect was described nearly two centuries ago. Contrary to what is believed by many, the science of the greenhouse effect is hardly new. Scientists living in the 1800s studied the relationship between GHG concentrations and Earth's temperature. In the 1850s, for example, a woman named Eunice Newton Foote was one of the first to discover that carbon dioxide plays a role in heating up a gas-filled atmosphere[10]. Others expanded on her findings, including a Swedish chemist named Svante Arrhenius, who, in 1896, conducted further work and even boldly suggested that human activity—notably, the burning of fossil fuels—would likely increase Earth's temperature by adding to the existing greenhouse effect[11]. We will see that analyses done by today's scientists build upon and confirm much of this early work.

9. NASA. What is the greenhouse effect? climate.nasa.gov/faq/19/what-is-the-greenhouse-effect/#:~:text=Credit%3A%20NASA%20Jet%20Propulsion%20Laboratory,it%20would%20be%20without%20them.

10. Climate.gov. Happy 200th birthday to Eunice Foote, hidden climate science pioneer www.climate.gov/news-features/features/happy-200th-birthday-eunice-foote-hidden-climate-science-pioneer

11. Anderson, Hawkins, Jones. 2016. CO_2, the greenhouse effect and global warming: from the pioneering work of Arrhenius and Callendar to today's Earth System Models. Endeavour 40: 178-187.

> *Despite what may be suggested by the name of the phenomenon, Earth is not heated in the same way as is a greenhouse. The greenhouse effect involves the actions of certain gases that absorb and radiate heat toward the surface. The glass in a greenhouse provides a physical barrier that inhibits heat from readily escaping.*

Multiple gases contribute to Earth's greenhouse effect

Many processes generate GHGs. Here we briefly examine those most relevant to the greenhouse effect.

Water vapor. Gaseous water is more abundant and effective than any other greenhouse gas, accounting for the majority of Earth's global warming. It is predominately produced by natural processes such as **evaporation** and **transpiration**, and cycles quickly in and out of the atmosphere. Its role in climate change is a complicated one, as the amount of water vapor present is directly linked to temperature. Importantly, the water-climate connection can be understood as a **positive feedback loop** (Chapter 2), because, as the surface gets warmer, more water moves into the atmosphere; this additional water causes temperatures to rise even more, and so forth. Note, then, that although it mostly moves along natural pathways, anthropogenic activities that release other greenhouse gases (immediately below) influence how much water is in the atmosphere. Thus, the contribution water makes to the greenhouse effect is often said to be *amplified* by other GHGs and their effect on temperature.

Carbon dioxide. The second gas on our list is less abundant than the first, but it is widely seen as the most important and problematic because of the role it plays in changing Earth's temperature and climates, AND because human actions appreciably affect its concentration. Note that carbon dioxide is used as the standard against which the effects of other greenhouse gases are compared. That is, contributions to global warming by GHGs are expressed in terms of number of CO_2 equivalents. By convention, then, each molecule of this gas is said to have a global warming potential (GWP) equal to 1. We will see the usefulness of such a standardization below. But first, we will briefly review the processes that emit and absorb this critical gas.

As we learned in previous chapters, **CO_2** is produced by many natural and anthropogenic activities. **Aerobic respiration** and **decomposition** convert **organic** carbon compounds in things like food, and soil to carbon dioxide. Other phenomena like volcanic eruptions and fires also release it. When we add them up, natural sources account for about 90% of this GHG entering the atmosphere annually. Humans release a relatively small amount of CO_2, mostly through combustion of **fossil fuels**, although **agriculture**, **deforestation**, and other practices play roles as well. In total, anthropogenic activity is responsible for about 10% of the carbon dioxide moved to the atmosphere each year. It may be seemingly small, but this human contribution has made a substantial difference, as we will see shortly. Keep in mind that, of the GHGs released by humans, CO_2 is by far the dominant one by quantity, making up about 65% of worldwide GHG emissions (it is 79% of the United States' emissions)[12].

What happens to the carbon dioxide put into the atmosphere? The short answer is: much is removed via various mechanisms or sinks (environmental scientists use the general term **sink** to refer to processes that absorb or otherwise remove materials from any reservoir). Clearly, some of the carbon dioxide entering the atmosphere from whatever source is withdrawn by an opposing

12. U.S. Environmental Protection Agency (EPA). Global greenhouse gas emissions data. www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data

force—otherwise, levels would have built up to far higher levels than we observe currently. For example, **fixation** into **glucose** by **algae**, plants, and some **bacteria** is a major pathway of the carbon cycle (Chapter 5), one that provides biologically available C to organisms while it pulls CO_2 out of the atmosphere. The rate at which that carbon returns to the atmosphere varies. Some is released quickly through the respiration of **aerobes**, but much persists in long-lived organisms like trees (decades or centuries), in **soil** (millennia), and in **fossil fuels** (tens of millions of years). Some CO_2 also moves into the hydrosphere via **diffusion** (Chapter 4), staying there for varying lengths of time.

We can determine the relationship between inputs and outputs if we measure what happens to the quantity of a material in any system (review Chapter 2 for more about systems analysis). In this case, CO_2 concentration has gone up steadily (that is, the gas has been accumulating) since about the year 1880, leading us to conclude that the rate at which the greenhouse gas enters the atmosphere exceeds the rate at which it exits.

At this point it would be fair to wonder: hang on, just *how* do we know about carbon dioxide levels from the past? The answer is, we combine evidence from direct and indirect measurements. The first type is relatively straight forward. Researchers at a site near Mauna Lau, Hawaii (U.S.A) have directly collected and analyzed air samples since 1958. That work is ongoing and has produced one of the most famous (as these things go) graphs to be produced by scientists (Figure 14.19).

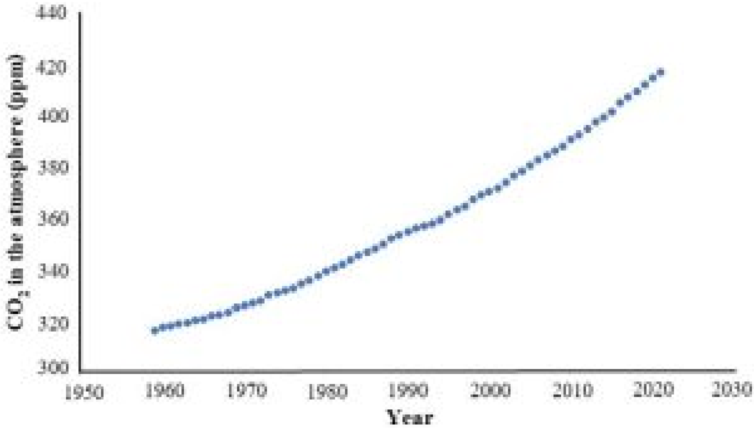


Figure 14.19. *The increase in carbon dioxide in the atmosphere from 1958 – 2021. Data points are yearly. Data were obtained by direct sampling at Mauna Loa, Hawaii (U.S.A.). Data from Tans, US NOAA / ESRL and Keeling, Scripps Institution of Oceanography, Public Domain; replotted by Kelsey.*

The second type of data collection is more complicated than the first and involves indirect measurements (also called **proxy measurements**) to give us information for the years before we sampled the atmosphere in real time (i.e., prior to 1958). Proxy measurements utilize samples of ancient atmospheres that have been trapped in bodies of ice, **glaciers**, that persist for hundreds of thousands of years. Very briefly, long cylindrical vertical cores are taken from thick sections of **glaciers**, and air in pores from thin sections is analyzed to determine concentrations of CO_2 and other gases in them (Figure 14.20).

Figure 14.20. *Ice cores are extracted from ancient glaciers, and the air trapped in them is analyzed to determine past concentrations of GHGs. Here is a portion of a core after it has been removed from ice.* US NOAA, *Public Domain.*

From here, we make two assumptions. First, it has been well-established that a glacier gets older with depth. Plus, we can use various tools of geology and chemistry to establish the age, in years, of different layers of ice. Second, air in the bubbles we find today are tiny samples of Earth's atmosphere from the past—they represent a series of snapshots, of sorts, of the air that must have been present when the ice formed. By combining age data with chemical analyses we can reconstruct the history of the composition of the atmosphere (Figure 14.21).

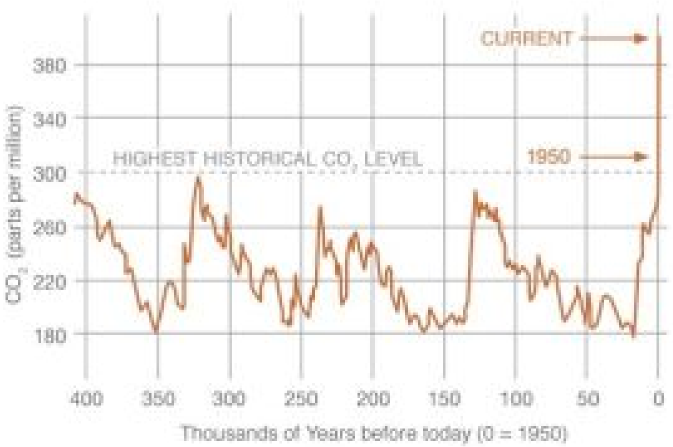


Figure 14.21. *Atmospheric carbon dioxide concentration during the past* 400,000 *years (ice core data).* US NOAA, *Public Domain.*

We will return to the topic of CO_2, including the long-term trends in its atmospheric concentration, shortly.

Methane. Emissions to and concentrations of CH_4 in the atmosphere are far lower than those for carbon dioxide, but this gas is a much more powerful GHG: one molecule of methane contributes as much to global warming as about 24 molecules of carbon dioxide (i.e., it has a GWP of 24 CO_2). What are the sources of methane? As with CO_2, both natural and anthropogenic processes release CH_4 to the atmosphere. This gas is produced during **anaerobic decomposition** of organic carbon compounds in swamps and other natural low-oxygen environments (Chapter 4). It is also released from the digestive systems of animals such as cows, camels, elk, and termites. In those cases involving animals, **mutualism** between a larger organism and microorganisms living inside of it is responsible for digestion of plant material, survival of both participants, and production of methane. The worldwide breeding of billions of dairy and beef cattle by farmers adds a substantial amount of CH_4 to the atmosphere (see Chapter 9), an activity that is likely to increase as both the number and standards of living of humans go up (Chapter 8). Burning of fossil fuels, deforestation, and **landfills**, also linked to increasing resource use by people, produce methane as well. Methane is the second-most-important anthropogenic GHG by quantity, accounting for approximately 15% of our emissions[13]. Several phenomena reduce the amount of methane in the atmosphere (i.e., are sinks). For instance, certain specialized bacteria convert it into CO_2 as part of their normal metabolism. In Chapter 13 we saw that combustion of methane produced in landfills is similarly transformed. Yes, carbon dioxide is still a greenhouse gas, but since it is a far less effective

13. U.S. Environmental Protection Agency (EPA). Global greenhouse gas emissions data. www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data

GHG than methane, we generally view these reactions as beneficial (in terms of effect on greenhouse warming). Overall, though, processes that remove methane cannot keep up with inputs of the gas, so it continues to accumulate. Like those for CO_2, atmospheric concentrations of CH_4 have gone up precipitously in recent centuries: today's levels are more than twice as high as those from the preindustrial age (again, as indicated by analyses of glaciers and direct measurements). Data suggest that the rate of increase has accelerated in recent years[14].

Nitrous oxide. The atmospheric concentration of this gas, N_2O, is considerably lower than that of either CO_2 or CH_4, but it has a **GWP** of approximately 3007. A little more than half of it enters the atmosphere from natural cycling of nitrogen (Chapter 4). Agriculture is responsible for the bulk of anthropogenic nitrous oxide, with fuel combustion and industrial processes making contributions as well. All told, N_2O is the third highest GHG on the list[15]. Chemical reactions convert it to other gases, but since its concentration is going up, those sinks do not currently keep pace with sources. Also of interest is the role nitrous oxide plays in ozone depletion, the second global-scale air pollution issue we will consider (below).

Fluorinated and chlorinated gases. This broad group includes several manufactured compounds that consist of fluorine (Fl), chlorine (Cl), carbon, and other atoms in various combinations. These days, emissions of fluorinated gases far outpace those of

14. National Oceanographic and Atmospheric Administration (NOAA). Increase in atmospheric methane set another record in 2021. www.noaa.gov/news-release/increase-in-atmospheric-methane-set-another-record-during-2021

15. U.S. Environmental Protection Agency (EPA). Global greenhouse gas emissions data. www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data

chlorinated gases because the latter have been largely phased out of production and use (more can be found below, including figures of them, in the discussion of ozone depletion). The relative concentrations of these are very low, yet they still can make a substantial contribution to warming because they are at least 10,000 fold more potent as greenhouse gases than carbon dioxide (sulfur hexafluoride, SF_6, tops the list with a GWP of 22,500[16]). Another very important consideration is that the residence time of these gases in the atmosphere tends to be very long, likely thousands of years in some cases. Note that no natural source produces any relevant amount of these types of gases, even if natural processes can remove them from the atmosphere (more shortly).

Carbon dioxide is linked to fossil fuel combustion: a closer look

A balance between natural inputs and outputs meant that concentrations of CO_2 in the atmosphere remained constant during much of the past thousand years. However, anthropogenic activity—notably, the burning of coal, oil, and natural gas—since the start of the Industrial Revolution has disrupted that long-held equilibrium: the organic carbon that had been stored in the lithosphere as fossil fuels for millions of years was converted to CO_2 and began moving into, and accumulating in, the atmosphere (Chapter 10). Keep in mind that a measurable increase in the concentrations of atmospheric GHGs is of concern because more outgoing heat energy is absorbed and re-emitted toward Earth (i.e., the greenhouse effect is enhanced). The planet's surface then gets even warmer than it was in the presence of a natural greenhouse effect. In other words, the start of the upward trend in average temperature we noted at the beginning of this discussion—and which we are trying to answer in question #2—coincided with

16. U.S. Environmental Protection Agency (EPA). Global greenhouse gas emissions data. www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data

industrialization, and industrialization resulted in increased emissions and atmospheric concentrations of CO_2. These links between human activity and changes in atmospheric chemistry are crucial pieces of evidence used to both explain what we have seen during the past 150 or so years and predict what is likely to happen in the future. Most of this discussion is centered on carbon dioxide, although the concentration of methane followed a similar trend since the 1800s. Due to differences in patterns of human activities and emissions, nitrous oxide and fluorinated / chlorinated gases did not begin to increase until the 1960s.

Atmospheric concentrations of some GHGs are higher now than in 800,000 years

Using ice core analyses, scientists have reconstructed the 800,000-year history of atmospheric carbon dioxide and methane. The concentration of each gas is higher now than at any time during that long period for which we can collect data. We will see a graphical representation of that trend shortly.

GHG concentrations have been linked to temperatures in the past

As we learned above, rising concentrations of gases like CO_2 are of concern on their own because of their chemical properties: as GHGs accumulate in the atmosphere, more heat will be re-emitted to Earth's surface. Scientists have also noted compelling evidence from ice-core data of a close link between temperature and atmospheric greenhouse gas concentrations during the past 800,000 years. Using chemical analysis of certain atoms present in glaciers, a timeline of past temperatures has been reconstructed and it follows the same trend as those for CO_2 (Figure 14.22.a).

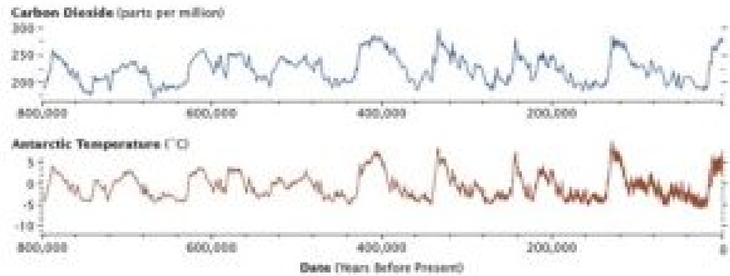


Figure 14.22.a. *Atmospheric CO_2 and temperature for most of the past 800,000 years (the most recent data are not included). Note that they follow the same trend, that is, their levels have risen and fallen together. Data were obtained from ice core sampling (see the main text). US NASA, Public Domain.*

If we assume the rules governing the universe have not changed with time, we should expect that the relationship between GHGs and temperature is the same today as it was historically. The temperature-CO_2 link since 1880 is evident as well (Figure 14.22.b).

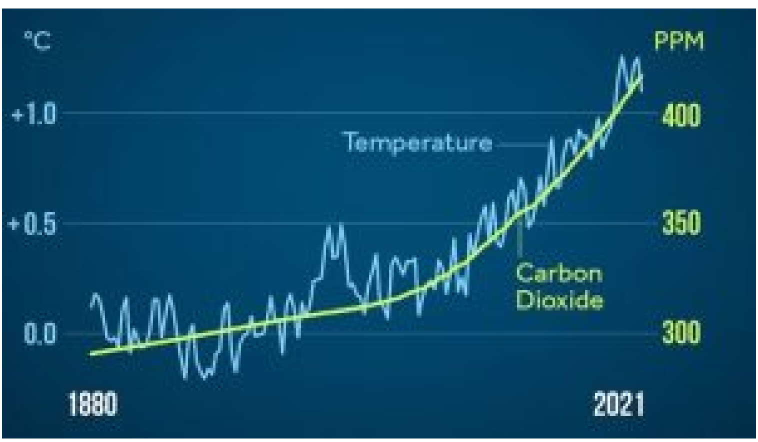


Figure 14.22.b. *The relationship between atmospheric CO_2 concentration and global temperature since 1880. US NASA, Public Domain.*

Question #3: If the answer to #1 is yes, then what would be the consequences? Continuing the list of questions started a while ago, we now ask: who cares if the temperature goes up by a few degrees? Does it really matter? Well, we noted earlier that the biggest risk of global warming is not that things will be noticeably warmer everywhere, but that climates will be affected in various and, in some cases, profound, ways. Since organisms are connected to and shaped by their environments (as explored multiple times in this book), all members of the biosphere, including humans, could be adversely affected by climate change.

Changes in precipitation patterns

Temperature influences the movement of water among Earth's reservoirs in two important ways. First, global warming heats the oceans and drives more **evaporation** into the atmosphere (as noted earlier, more water vapor leads to more warming and subsequent evaporation). Second, rates of condensation of water from air decline as temperature of the air goes up; in essence, warmer air carries more water than does cooler air. Researchers expect that these changes to the atmosphere will bring about higher rates of precipitation in some places while some areas will get drier during the next several decades. In the United States, for example, increased volumes and rates of rainfall have already been noted in northeastern regions, whereas droughts have gotten worse in some southwestern regions[17]. Organisms adapted to precipitation patterns in their current natural environments would undergo stress and possible extinction, but the consequences of shifting rainfall to humans could be profound as well. Among the worries

17. U.S. Environmental Protection Agency (EPA). Climate Change Indicators: US and Global Precipitation. www.epa.gov/climate-indicators/climate-change-indicators-us-and-global-precipitation#:~:text=Since%201901%2C%20global%20precipitation%20has,increases%20in%20precipitation%20than%20others..

facing many of the world's peoples are: flooding and other extreme events would cause loss of life, property, and agricultural production. Scarcity of resources like clean water, habitable land, and food could then lead to increased conflicts among states and nations.

Rising sea level

Water expands as it is heated. So, even without additions, water in the oceans will simply take up more space—that is, **thermal expansion** will cause it to rise and encroach further onto land—as global temperatures go up. More water *is* introduced as well, though, when glacial melt (more below) flows into the oceans. Taken together, these two phenomena have driven sea level to increase by approximately 22 cm, or 8 inches, since 1880[18]. Worth noting, though, is the way the rate of increase has accelerated in recent decades. For most of the 20th Century, it was about 1.4 millimeters per year but rose to 3.6 mm per year between 2006 and 2015. Scientists project that the surface of the world's **oceans will continue to rise by at least one meter throughout the rest of this century** (estimates range from 0.2 to more than 2.0 m). Natural ecosystems, particularly those at delicate water-land interfaces, are sensitive to such rapid and dramatic change, and **extinctions** of species living in specialized habitats are possible. Since roughly one third of the world's peoples live within 100 km of a coastline[19], even a modest rise would also disrupt human structures, systems, and

18. National Oceanographic and Atmospheric Administration (NOAA). Climate Change: Global Sea Level. www.climate.gov/news-features/understanding-climate/climate-change-global-sea-level#:~:text=April%2019%2C%202022-,Highlights,3.8%20inches)%20above%201993%20levels.

19. United Nations (UN). The Ocean Conference. www.un.org/sustainabledevelopment/wp-content/uploads/2017/05/Ocean-fact-sheet-package.pdf

lives. Cities such as New York, Washington, Miami, New Orleans, Los Angeles (all U.S.A.), Tokyo (Japan), Venice (Italy), Mumbai (India), and Shanghai (China,) as well as areas within many nations in Asia, Africa, and South America are at risk. The Maldives (Indian Ocean) is a small, low-lying island nation made famous in part because it could be completely inundated by 2100 if sea level continues to rise (Figure 14.23).

Figure 14.23. *One of the islands of the low-lying nation of the Maldives. It is vulnerable to flooding as sea levels rise and could be uninhabitable by 2100. Gzzz, CC BY-SA.*

The relevance of severe flooding surely is obvious: death, damage, and destruction are all likely. However, you should keep in mind that many people find themselves increasingly plagued by low-impact events. Among other trouble caused by this nuisance flooding is heightened vulnerability to extreme storms. The frequent presence of excess water can make losses due to, say, a hurricane, even worse[20]. Also note that stagnant flood waters

20. National Oceanographic and Atmospheric Administration (NOAA).

are ideal environments for reproduction of disease-carrying mosquitoes.

Changes in ocean circulation

Water mixes and moves among Earth's oceans via a complex mechanism called the **global conveyor belt** (Figure 14.24).

Climate Change: Global Sea Level. www.climate.gov/news-features/understanding-climate/climate-change-global-sea-level#:~:text=April%2019%2C%202022-,Highlights,3.8%20inches)%20above%201993%20levels.

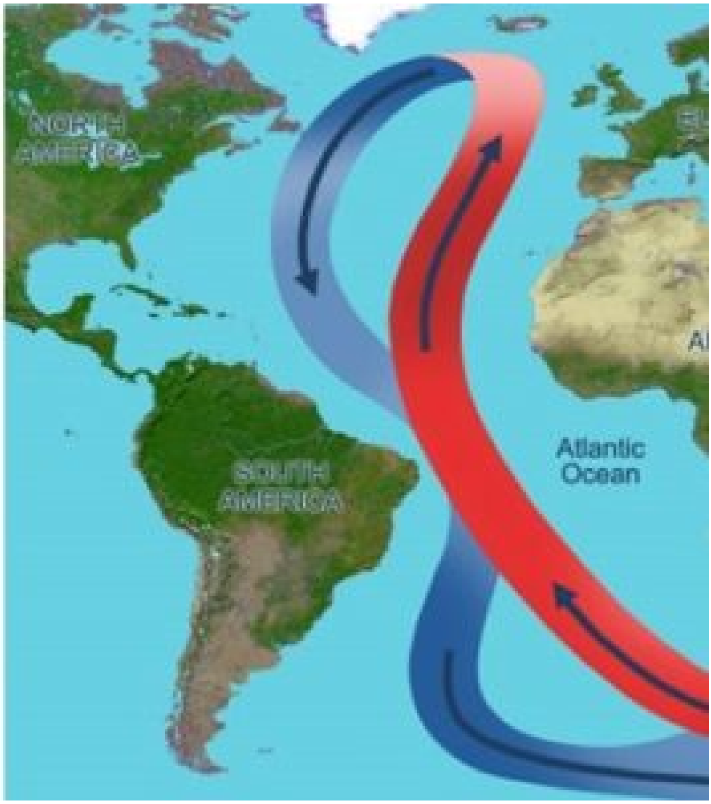


Figure 14.24. *A diagram of a portion of the global conveyor belt. As noted in the main text, changes in ocean chemistry could weaken the movement of warm equatorial water to the United Kingdom and other places that are far north (in the figure, the red arrow pointing up from the southern Atlantic Ocean). US NASA, Public Domain.*

Briefly, water sinks or rises due to temperature and **salinity** variations, and these vertical displacements ultimately drive the lateral circulation of massive volumes of water from cold to warm regions (and back again). Among other important outcomes it produces, the conveyor belt moves heated equatorial water to the northern Atlantic Ocean, making The United Kingdom and other parts of Europe much warmer than they would be otherwise. A combination of changing precipitation patterns and melting glaciers would introduce more freshwater into the oceans, disrupt the current cycle of rising and falling water, and ultimately slow circulation. A weakened conveyor would cool affected areas—perhaps by a great deal—providing one example of the way general warming of the Earth would bring about a range of location-specific effects.

Melting of glaciers

Like the many other systems we have considered, the relationship between inputs and outputs—in this case, the freezing of water and the melting of ice, respectively—controls the amount of material inside the **glacier** at any time. As average global temperatures have risen, rates of melting have outpaced rates of freezing, and that imbalance seems to be increasing. Ground-based photography has shown that many glaciers around the world have gotten appreciably smaller during the past century (Figure 14.25), and recent data collected from satellite images indicate about 400 billion metric tons in annual net ice losses[21] from those in Greenland and Antarctica—which together contain about 99% of Earth's ice—since 2002[22,23].

21. It's hard to picture how much that is! 1 Gt of water occupies 1 cubic kilometer of space—a cube that is 1 km or 0.6 miles long in each dimension.

22. NASA. Global Climate Change. Vital Signs. climate.nasa.gov/vital-signs/ice-sheets/

23. US Geological Survey (USGS). Where are Earth's glaciers located? www.usgs.gov/faqs/where-are-earths-glaciers-located

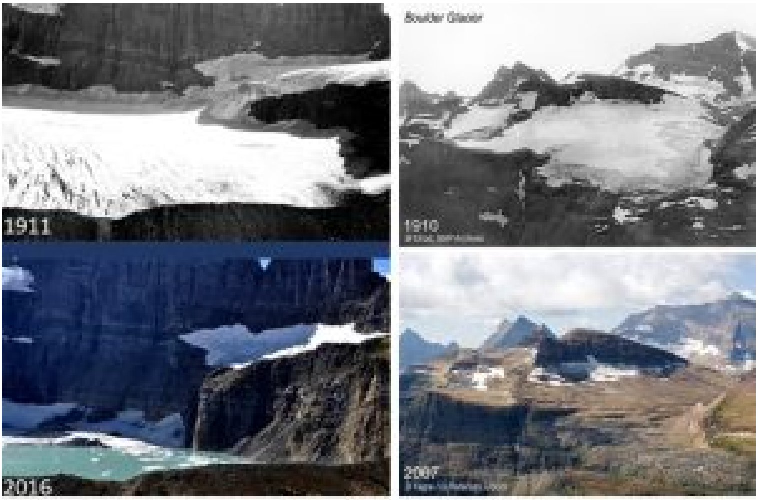


Figure 14.25. *Photographic evidence shows retreat of many glaciers around the world, including the two here. Left (a): Grinnell Glacier, Montana (top is older than bottom); right (b): Boulder Glacier, Washington (top is older than bottom). USGS, Public Domain (a and b).*

Glacial melting is important for at least three reasons. First, the water it releases contributes to rising sea levels. In fact, contributions from meltwaters relative to that of the thermal expansion of water (described above) is thought to be greater now than it was just a decade ago because of accelerating rates of melting[24] (note that melting *sea* ice does not contribute to these rises—the water is already in the ocean, albeit in solid form). Second, it slows the global conveyor belt (above) because it

24. NASA. Global Climate Changes. Vital Signs. Greenland, Antarctica Melting Six Times Faster Than in the 1990s. climate.nasa.gov/news/2958/greenland-antarctica-melting-six-times-faster-than-in-the-1990s/#:~:text=Andrew%20Shepherd%20at%20the%20University,the%202010s—a%20sixfold%20increase.

introduces freshwater into oceans and reduces salinity. Third, areas that depend on water released from melting glaciers to **recharge** depleted **groundwater** and other reservoirs (Chapter 11) face water shortages as the amount of water stored in ice dwindles each year.

Ocean acidification

Although not directly linked to global warming, this outcome is related because it is caused by increased CO_2 in the atmosphere. Put simply, excess carbon dioxide accumulates and alters atmospheric chemistry as we have noted. Some of it moves into the hydrosphere, though. Once in the oceans, it combines with water and a naturally occurring ion called carbonate (CO_3^{2-}) to produce bicarbonate (HCO_3^-). This reaction has lowered ocean pH by 0.1 unit (review Box 14.1, above), or, in more meaningful terms, increased acidity by 30%. Marine organisms needing carbonate to build shells (e.g., clams, corals, sea urchins) struggle to succeed in their changed environment[25]. If the rise of atmospheric carbon dioxide continues its current trajectory, **acidity will increase by 150% by the year 2100**[26][/footnote]. Shifts in **ecosystem structure** are likely as organisms adapted to current pH levels would face increasingly hostile conditions. People who rely on marine resources for food and income would also be adversely affected.

Changes to global primary production?

Climate change could reduce the number of marine **phytoplankton**, the largest source of fixed C and gaseous O_2 on Earth. Recent data indicate declines in both the size and success of

25. National Oceanographic and Atmospheric Administration (NOAA). 2018. What is Ocean Acidification? PMEL carbon program. www.pmel.noaa.gov/co2/story/What+is+Ocean+Acidification%3F
26. [27]National Oceanographic and Atmospheric Administration (NOAA). 2018. What is Ocean Acidification? PMEL carbon program. www.pmel.noaa.gov/co2/story/What+is+Ocean+Acidification%3F
27.

the communities of these critical organisms. If this trend continues, not only would ocean ecosystems suffer damage, all **aerobic** organisms (including humans) would be imperiled [28].

Altered disease distribution?

We know that changes in the physical, chemical, and biological properties of an environment can influence all biological entities present in an ecosystem, including those that cause diseases (i.e., **pathogens**). So, higher average temperatures have the potential to change the distribution of human parasites. Malaria, for example, is generally confined to tropical and subtropical regions because both the **protozoan** that causes it and the mosquito that transmits the protozoan to people require relatively warm and wet conditions. If temperatures rise, though, the ranges of this and other important diseases could increase and affect northern countries. Additional research is needed before definitive conclusions can be drawn.

An increase in both the intensity and number of storms and fires

Many climate scientists have suggested that since hurricanes gather energy from the ocean, they will strengthen as water temperature rises. Furthermore, some already dry areas—like those in western portions of the U.S.—will become increasingly subject to wildfires as droughts worsen. Further research is needed to sort out the validities of these prediction, but they are plausible and even seem to be supported by recent events (i.e., unusually large and devastating wildfires).

Rapid, dramatic changes

We should make a final note about these and any other potential adverse consequences of warming and climate change. Given the

28. NASA. Earth Observatory. Phytoplankton Productivity Down in Gulf of Maine. earthobservatory.nasa.gov/images/149915/phytoplankton-productivity-down-in-gulf-of-maine#:~:text=Research%20published%20in%202021%20showed,warming%20has%20affected%20the%20phytoplankton

myriad interactions and complexities in the system, a growing number of researchers express concern that the rate of some changes could greatly accelerate if Earth reaches what is often called a **tipping point**, a critical level at which a cascade of events is initiated. So, although atmospheric carbon dioxide might continue to rise at a constant rate for an extended period, it could reach a concentration that causes sudden and unexpected shifts in temperature, precipitation, ecosystem distribution, or biodiversity. Once change occurs, it would then be very difficult to return an affected system to its former conditions. It is somewhat akin to people shuffling slowly and steadily toward the edge of a cliff. At the edge, just a few centimeters of horizontal movement results in hundreds of meters of vertical travel as they suddenly plummet to a low elevation.

Question #4: If the answer to #1 is yes, then what can we do about it? Finally, we get to our last question. Since evidence increasingly points to human activity, notably greenhouse gas emissions, as the primary cause of global warming, we have the power to affect it. Simply put, if we can bring rates of inputs and outputs of GHGs back into balance, we could stop further warming and climate change and perhaps even reverse what has already occurred. Some possible solutions are listed here. Keep in mind that no single strategy would solve the problem, rather a combination of ideas and approaches will very likely be needed. Here are some ideas likely to be most effective.

Reduce GHG emissions

Carbon dioxide. Since this is the single most important anthropogenic greenhouse gas, it will receive the bulk of our attention. Of the many steps people could take, we will consider three having the greatest potential to slow the accumulation of CO_2 in the atmosphere. First, decreasing the amount of coal, oil, and natural gas burned to generate electricity and power transportation—that is, accomplishing the same amount of work with less combustion—would lower carbon emissions. Improved efficiency of air conditioners, refrigerators, and other appliances, along with better fuel economy for motor vehicles, are some of the ways reduced fossil fuel usage is accomplished. For instance, the tightening of federally mandated fuel efficiency standards for automobiles sold in the United States has lowered the use of gasoline to a degree. Ultimately, though, a substantial switch away from fossil fuels to energy sources like wind, solar, and hydroelectric that do not lead to emissions of CO_2 will likely be needed to effectively combat the problem of global warming (we will return to this point below). Exhaustive coverage of the economic, cultural, and scientific issues that limit the ambition, rigor, and effectiveness of such measures is not possible in this space. Suffice it to say that enacting relevant laws is enormously difficult because of multiple disagreements and competing priorities among interested parties. Second, managing expectations of ever-higher standards of living—that is, encouraging less consumption of goods, services, and transportation that rely heavily on fossil fuels—would reduce emissions of carbon dioxide. Finally, more use of agricultural practices that minimize decomposition of **soil organic matter** would reduce an important source of CO_2.

Methane and nitrous oxide. Better agricultural practices and management of waste, as we saw earlier in this chapter, limit the amount of these gases reaching the atmosphere.

Others. Stricter regulations and monitoring of the powerful synthetic GHGs noted earlier (fluorinated and chlorinated gases) through recycling, the use of alternative compounds, and careful control to limit releases, all lessen their ability to increase Earth's greenhouse effect.

Sequester carbon dioxide

A combination of technology may be used to pull CO_2 from smokestacks and other sources before it moves into the atmosphere. After its capture, the carbon dioxide can be stored in the lithosphere or other places (a practice called carbon capture and storage, CCS) thus reducing the rate of inputs of the gas to the atmosphere. A related strategy, **carbon capture and utilization** (CCU), makes use of the carbon dioxide in the manufacturing of plastics and other products (under the proper conditions, CO_2 can be combined with other chemicals to build useable solid materials). Although CCU currently operates on a small scale, it has the potential to dramatically reduce the net amount of carbon in the atmosphere and reduce the problems of plastic pollution. Research on both CCS and CCU continues[29].

Slow or reverse deforestation

Reducing rates of **deforestation** would affect both inputs and outputs of atmospheric CO_2. The burning of trees—a typical approach when large, forested areas are cleared—is an important GHG source because it converts stored organic carbon into carbon dioxide. The sink side of the equation is influenced in that fewer trees and other photosynthesizing organisms would draw carbon dioxide out of the atmosphere. In short, land clearing for agriculture, housing, or forestry increases the movement of carbon from the biosphere to the atmosphere while simultaneously decreasing the movement in the opposite direction. Actively increasing the number of trees is another piece of this solution, although **afforestation** would conflict with other land uses to support a growing population.

> *Recall from Chapter 12 that larger trees sequester far more carbon than do smaller trees. Thus, protecting existing forests has a much bigger impact on atmospheric* CO_2 *than does planting new (i.e., smaller) trees.*

Adapt to a changed world

29. International Energy Agency (IEA). Carbon Capture, Utlisation and Storage. www.iea.org/fuels-and-technologies/carbon-capture-utilisation-and-storage

What if we do not take the steps needed to address global warming? Some people already have given up, concluding that nothing can be done in any case, are in denial that change will occur, or believe it is natural and inevitable. How might we diminish the adverse consequences and live with climate change?

Build protective structures. As oceans rise, flood risks in low-lying areas increase. Venice (Italy), for example, has contended with sinking land for many centuries, but it could become uninhabitable within a hundred years if sea level changes as predicted. Many other coastal zones are threatened by rising water, and the problem is likely to get worse in coming decades. What can be done? In addition to relocating millions of people to higher ground, bigger, stronger, and more expensive levees, sea walls, and dams could be built to protect the most vulnerable areas. These structures would need to be designed to hold back ever-more-powerful storm surges, one of the possible outcomes of global warming. In other words, big investments would be required to avoid the kind of damage seen in 2005 when Hurricane Katrina overwhelmed inadequate levees intended to protect New Orleans, Louisiana (U.S.A.), a city that already sits below sea level[30].

Restore beaches. As coastal residents know, sandy beaches are far from static. Even in a relatively uneventful year, they move and change shape due to multiple natural forces. Resort communities that rely on summer tourism therefore invest considerable amounts of money to replace lost sand and maintain attractive beaches indefinitely. If oceans rise and hurricanes become more intense, beach maintenance and restoration will become increasingly difficult and expensive.

Relocate affected people. Even if we can build structures or otherwise modify environments to protect them from the effects of climate change, certain places are likely to undergo irreparable

30. weather.gov/mob/katrina

damage. Simply put, some people will be forced to permanently move elsewhere. The number affected is hard to predict with certainty, but since about a third of the world's population lives near a coastline[31], many hundreds of millions could potentially be displaced. It is unclear where they would go and how the enormous cost associated with their relocation would be met.

Manage food shortages, diseases, and conflicts. We will not reiterate what was said above about these consequences, but they all should be added to the list of necessary adjustments if people elect inaction and adaptation in favor of prevention and reversal of climate change.

Coordinating a global response

The solutions proposed here have been difficult to implement for many reasons, including the fact that climate change is a global-level concern that requires global-level strategies to be mitigated. In other words, peoples and countries with very different levels of influence, wealth, and attitudes need to coordinate their efforts. Understating matters by a fair amount, it has not been easy for the world's nations to come up with an effective plan palatable to everybody. Two groups trying to manage scientific research as well as appropriate action are very briefly described here.

Intergovernmental Panel on Climate Change (IPCC). This is an international group working through the U.N. to monitor and summarize the current state of the science of climate change research. It has issued five formal statements during its roughly three-decade existence (one every six or so years) that provide predictions and recommended actions based on the available data. It has been a very important and influential organization and continues to gather, interpret, and disseminate the evidence related

31. United Nations (UN). The Ocean Conference. www.un.org/sustainabledevelopment/wp-content/uploads/2017/05/Ocean-fact-sheet-package.pdf

to human-caused climate change. You are encouraged to explore the work of the IPCC, including its latest report online[32].

United Nations Framework Convention on Climate Change (UNFCCC). Established in 1992, this organization is designed to help member countries study and reduce anthropogenic contributions to climate change (consult its website for more information[33]. The Kyoto Protocol (1994) and the Paris Agreement (2016) are two major initiatives of the UNFCCC to limit greenhouse gas emissions. Efforts to reach their stated goals have been hampered by several obstacles, however. Notably, **developing** and **developed nations** are motivated by different priorities and notions of fairness and justice. Countries like the United States, *currently* responsible for about 25% of the CO_2 put into the atmosphere, disagree with countries such as China, which are likely to emit far more of this and other greenhouse gases in the *future*, on what is the most equitable way to move forward. A nation like The Maldives (described above), argues that, although it contributes relatively little to the problem, it bears a disproportionate amount of the burden of climate change. Arguments about the economic consequences of any proposals also tend to limit their scope and effectiveness. Moreover, the fact that some people do not accept the science of climate change, a topic we will revisit shortly, adds to the difficulties of implementing appropriate responses.

Evidence-based predictions

Careful analysis of data enables scientists to describe the *history* of Earth's climate during the past few centuries. Now, though, we

32. The Intergovernmental Panel on Climate Change. http://www.ipcc.ch

33. U.N Climate Change. http://unfccc.int

must ask: what about the *future*? This critical question is the subject of continued study.

Using computer models. Computer models are employed to help predict what will occur in coming years. In simple terms, data about climate such as current temperatures, ice cover, cloud cover, and greenhouse gas emissions, are inputs to a computer program and forecasts about temperatures, rainfall distribution, sea level, and other concerns are outputs. These programs themselves are designed to account for as many variables as possible as well as the multiple and complex interactions among all the relevant parts of the system. Although models differ in the specific outputs they produce, most point to a warmer planet overall—as stated previously, on average, the increase will likely be 1.1 – 5.4 °C by the end of this century. Models also quantify how much sea level will rise (0.2 – 2.0 m) and weigh in on the other consequences described above.

Challenges to modelling the future. A model is only as good at predicting the future as is the validity of its inputs and assumptions. Thus, as our understanding of Earth's climate systems improves, so does the quality of our models. Several challenges still limit the reliability of models.

The role of feedback

Recall from Chapter 2 that **feedback** can either accentuate or attenuate a system response, leading to change or stability, respectively. Some unanswered questions about the role feedback will play in climate change hamper, somewhat, the validity of models.

Positive. Here, past increases in greenhouse gas concentrations, higher temperatures, shifting precipitation, melting glaciers, and sea level rises get amplified by current and future system responses; in other words, the rate of climate change increases. Two relevant examples illustrate this type of feedback. First, as glaciers melt and get smaller, they shrink faster and faster. When they are very large, their light-colored ice and snow reflects most of the incoming light energy, that is, they tend to raise Earth's albedo. As temperatures go up, the melting ice reveals the dark-colored rocks and soil that were beneath, leading to lower albedo. More energy absorption, in turn, increases temperature and accelerates future melting (Figure 14.26; compare to Figure 2.10).

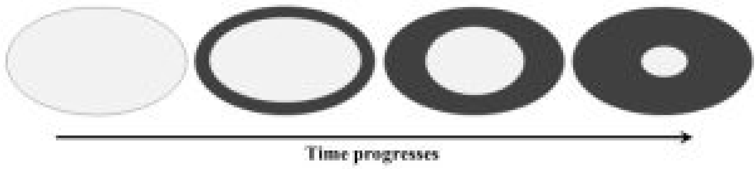


Figure 14.26. *A diagram of the way positive feedback affects glacial melting (map view). Moving left to right: as ice disappears, more and more dark soil is revealed, leading to more absorption of energy and rapid temperature increase in the vicinity.* Kelsey, CC BY-NC-SA.

Second, since evaporation is driven by temperature, more and more water is converted from liquid to gas as oceans warm. The addition of water vapor to the atmosphere then increases the amount of infrared energy emitted back to the surface, leading to even higher temperatures and rates of evaporation in the future. As we have seen repeatedly: current change accelerates future change. Models give varying weight to positive feedback as our understanding of it evolves.

Negative. Past outputs are essentially cancelled out by current outputs in this case, and temperature and climate remain stable. Again, we refer to two important examples. The first involves the response of primary producers to increased levels of carbon dioxide in the atmosphere. We might expect that **carbon fixers** like plants and algae would be stimulated to grow as more C is added to their environments, and this increase in **primary production** would then draw more CO_2 out of the atmosphere. So, Earth's system would remain stable: it would respond to more of this important GHG by increasing the size of the living C sink. Unfortunately, this proposed feedback mechanism probably contributes little to stability. Keep in mind that carbon is generally not the **limiting factor** for primary producers, rather, lack of nutrients such as nitrogen, phosphorous, or iron tends to slow their growth before carbon is exhausted; environmental changes like loss of **topsoil** are also more important (see Chapter 9 for more about limiting factors). Furthermore, some research even suggests that excess CO_2 might slow the growth of plants, therefore reducing rates of fixation. Our second example of negative feedback returns to the way increased temperature leads to more evaporation. Unlike its role in positive feedback noted above, though, water vapor driven from heated oceans would adhere to dust particles in the atmosphere and form more clouds. This increase in cloud cover would block more incoming light energy—that is, **albedo** would be higher—and Earth's average temperature would ultimately decrease. Many climate models incorporate changes in cloud cover, but the extent to which this variable will affect future trends is not entirely understood.

Which type dominates? Each of the four examples of feedback just described, along with several others, is surely active to some extent. However, given the observed continuing measurable increases in atmospheric CO_2, average temperature, and sea level (as well as the higher rates at which they have changed in recent decades) positive feedback seems to be dominant. Negative feedback may be *slowing* the rate of climate change but its power to bring stability is clearly limited.

The complexity of Earth's systems

The system under consideration—which includes the atmosphere, hydrosphere, lithosphere, and biosphere—is characterized by enormous size and complexity. Although we have learned much, our understanding of the interactions and factors affecting climate is certainly incomplete and continues to evolve.

Unknowns about future human behavior

The predictions about temperature and sea level noted above are expressed as ranges in large part because we simply cannot predict with certainty what humans will be doing during the coming decades. Will greenhouse gas emissions increase? Decrease? Remain the same? Arguments can be made for each of those possibilities, and they depend on future population growth, standards of living, dominant agricultural techniques employed, fossil vs. non-carbon fuel usage, and extent and rates of deforestation. Only with time will we know the answers for sure.

Measurement uncertainty

Scientists use multiple tools and instruments to measure, among other properties, temperature and chemical composition of Earth's complex atmosphere. Many readings need to be taken at different locations, altitudes, and times, instruments must be carefully maintained and calibrated, and scientists must be consistent and well trained. So, as in all research, measurement **uncertainty** is acknowledged, quantified, and considered when predictions are made.

Testing the reliability of models. How do we know if a computer model can be trusted? A climate model can be tested, or **validated**, if we input data from the past, say 300 years ago, and then see how well the output predicts what we know to be the current state of the climate (i.e., according to the computer, what will happen in the future). Since we assume the rules governing the behavior of Earth's climate systems do not change with time, a model that accurately describes what happened during the previous centuries will likely be equally successful predicting the century that is yet to come. It turns out that some models are better than others, but none is 100% reliable. Interestingly, a common problem is that evolving models *underestimate* how much warmer things will be: attempts at validation indicate that temperature should be lower than it really is, indicating that changes could be *more severe* than we expect.

Climate change doubters and deniers: who and why?

Even as evidence for its causes and consequences mounts, some people deny the reality of climate change.

Degrees of denial. It is fair to say that climate-change denial is generally not a yes-or-no question but that most people find themselves somewhere on a spectrum from complete denial to complete acceptance. Some people refuse to believe that any warming or other changes have even occurred, despite evidence to the contrary. Some recognize that data about increases in greenhouse gases and temperature are convincing but maintain there is insufficient evidence to link the observed changes to human actions. Importantly, they tend to argue that we should wait to take any steps intended to combat climate change until the science is stronger. Still others accept that climate change is caused at least in part by humans but that the processes cannot be reversed. In other words, they see our only option is adaptation to a changed world. Finally, some people, including most scientists, are persuaded by the evidence and accept that human-induced climate change is a reality that can be at least partially mitigated. They often hold the view that, although more needs to be learned about the science of the problem, the data are sufficiently compelling to justify making proactive changes immediately, even some that might be very inconvenient.

Who are the deniers?. Just who denies and who accepts the science of climate change? Lumping together all the doubters mentioned above into one heterogenous group, numerous estimates suggest that just 3% of experts—that is, climate scientists—do *not* accept that human emissions of GHGs are affecting the greenhouse effect and driving climate change[34]. A somewhat larger percentage of people *without* expertise in science (the number likely ranges between 10 and 50%) doubt at least some of the science of climate change. Put another way, knowledge of the scientific method and data analysis tends to be an important factor determining where one lies on the denial spectrum.

34. NASA. Do scientists agree on climate change? climate.nasa.gov/faq/17/do-scientists-agree-on-climate-change/Climate.NASA.gov

Sources of doubt. The reasons people cite for their doubt and denial are numerous. Some draw upon the science of climate change—often misrepresenting or misunderstanding the meaning of the data—but many are rooted in values, opinions, or other non-scientific arguments. Although strong and objective data generally move scientists, many other people, including lawmakers and members of the public, are not so easily swayed by even the most compelling evidence. Convincing skeptics is challenging, but not impossible: see the essay in Box 14.3 by guest writer and former member of the U.S. House of Representatives Robert Inglis (Republican from South Carolina) for one example of the power of science to change minds.

Box 14.3. **Evidence changed my mind: a note from Robert Inglis, former member of U.S. House of Representatives**

During my first six years in the U.S. Congress (1993–1999) I said that climate change was nonsense, a figment of Al Gore's imagination. I represented a very conservative district in South Carolina, and I spouted the party line. Then, after being out of Congress for six years, I had the opportunity to run again for the same seat in 2004. My son challenged me to care about the environment. I got re-elected and went to Antarctica with the House Science Committee and saw the evidence in the ice core drillings. On another Science Committee trip at the Great Barrier Reef I was inspired by the faith of an Aussie climate scientist named Scott Heron. I wanted to be like Scott, loving God and loving people, so I came home and introduced into Congress a revenue-neutral, border-adjustable carbon tax. My political timing wasn't so good as the Great Recession was on. I got tossed out of Congress in a Republican primary in 2010 for various heresies against the temporary Republican orthodoxy, the most enduring of which was my willingness to continue to say that climate change is real. Ever since I've been out to convince fellow conservatives that it's actually quite conservative to care about climate change. I founded and direct a grassroots educational campaign that we've branded as republicEn.org.

My advice to students and members of the public at large is this: when factual observations overtake our shaky ideologies, it's better to be overtaken.

–Bob Inglis, 2019

Here is a short list of some common sources of doubt, along with brief responses and clarifications.

Natural sources of carbon dioxide outweigh anthropogenic ones

We learned earlier that anthropogenic sources contribute about 10% of the CO_2 that enters the atmosphere annually. Some people have used these data to suggest that humans play an insignificant role in the cycling of this important greenhouse gas. However, that relatively small amount has thrown off the equilibrium between sinks and sources that was in place for many centuries prior to 1880. The yearly surplus from fossil fuel combustion and other activities accumulates, driving the atmospheric concentration of carbon dioxide higher and higher (review Figure 14.19).

But….water vapor comes from natural sources!

Water vapor is the most important greenhouse gas, accounting for most of Earth's greenhouse warming. It is also largely produced by natural processes. So, the argument often goes, observed changes in temperature cannot be linked to human emissions of carbon dioxide—they are simply trivial in comparison to this natural GHG. As is stated in the above point, though, human contributions have shifted a long-standing balance between sources and sinks of several greenhouse gases. Yes, water is more abundant and more important, but that hardly means increasing levels of CO_2 are therefore irrelevant: as we know, temperatures started to rise only after the widespread burning of fossil fuels by humans. We need to also remember that warming caused by carbon dioxide and other anthropogenic GHGs increases the rate of evaporation and the amount of gaseous water moving into the atmosphere. In other words, human activity does indeed affect the water cycle.

Fluctuations in temperature and other climatic factors are normal and natural

This statement is supported by scientific evidence: Earth's long history has been characterized by many changes to its average temperature and climate. The rapid pace at which current changes are occurring, though—including the high rate of GHG accumulation in the atmosphere—appears to be unprecedented.

Lower sunspot activity will cool Earth in the coming decades

Some people—including a small number of scientists—believe the Earth is actually cooling and that any upward pressure on temperature will be cancelled out by larger climate forces in the near future. An important piece of evidence used to support this hypothesis is the prolonged period of low sunspot activity and coincident cooling that occurred during the 17th Century. The so-called Little Ice Age (review Figure 14.14, above) was characterized by lower average temperatures for several decades, but most scientists hold that the reduced solar activity during the same period was not the only factor at work and would have been insufficient to account for all the observed climate change in any case. Major volcanic eruptions likely had a significant impact, for example. Other problems limit the plausibility of this cooling effect. First and arguably foremost, we have not learned how to predict future solar output, only record what has already occurred.

In short, we really have no idea what will happen to sunspot activity in the short or long run. Secondly, the fact that Earth's temperature has risen steadily since 1880, despite several changes in energy output from the sun in the past century, suggests other forces must make a larger contribution to global warming.

If we can't forecast tomorrow's weather, how can we predict next century's climate[35]?

Frustration with inaccurate weather forecasts is used by many to dismiss the science of climate change as altogether unreliable. It turns out, though, that "weather" and "climate" refer to distinct phenomena, with the former being harder to predict than the latter. **Weather** is made up of short-term conditions such as temperature, humidity, rainfall, and wind and is highly variable from day to day. **Climate**, on the other hand, includes long-term characteristics prevalent in a region such as the range of temperature and total rainfall in a given year. In short, the weather conditions at any time are a function of the climate of an area. January snow in a place like upstate New York (U.S.A.), for example, is hardly surprising because its climate is appropriate for such events. So, what does this have to do with the reliability—or lack thereof—of predictions? Weather forecasting is particularly difficult because it tries to describe details about the timing and severity of properties like upcoming precipitation and temperature. "Will it rain Tuesday?" is the kind of question a weather forecaster tries to answer by providing probabilities ("70% chance of showers", and the like). Since so many forces and interactions influence specific weather over the course of just a few days, we do not always get it right. Alternatively, climate projections aim to describe trends that will happen in the long run. We know, for example, that regular shifts in the tilt of this planet bring spring and melting of the snow

35. Written with contributions from Jeff Fetzer, NASA scientist. 2010. ClimateNASA.gov

in upstate New York and cause winter to return by November or so. Changes unfolding during many years can only be detected with long-range, persistent monitoring, and might be temporarily obscured somewhat by today's weather events. As computer models become increasingly accurate in their ability to take past data and predict current conditions (i.e., they can be validated as described above), our confidence in predictions of future climates grows. But you should still consider packing an umbrella just in case tomorrow's forecast for "10% chance of showers" turns out to underestimate the likelihood of rain.

How can global warming be real when this winter was so cold?

Our succinct answer to arguments like this one is: climate change refers to long-term global trends, not what occurs in a location during a single month, season, or year. So, an unusual snowstorm in May, for instance, is not evidence of anything in particular except that weather can vary from year to year. It certainly does not invalidate more than a century of data on increased temperature and sea level rise. People's misconceptions about the meaning and relevance of climate change are not surprising, though, because "global warming" seems like it ought to lead to warmer weather everywhere. And, deniers tend to seize upon isolated facts to support their world views—the same people who claim that a warm February debunks the science of climate change are generally silent on the possible significance of an unusually hot October (again, neither anomaly should be cited to support *any* hypothesis).

The economic consequences of change would be too severe

This common view reflects some uneasiness about the steps we need to take, but it in no way invalidates the science of climate change. That is, although the consequences of reducing GHG emissions may indeed be inconvenient and difficult, it is illogical to conclude that global warming is therefore something we can dismiss as fiction. It is also by no means certain that changing from non-renewable fuels like coal and gasoline to renewable sources such as solar and wind would lead to economic hardship. Yes, a conversion to a non-carbon economy, with all the necessary modifications to infrastructure, transportation, agriculture, and so on, would not be easy to manage and would likely cause some short-term pain. However, the development of new renewable energy technology would offer new investment opportunities, encourage innovation and entrepreneurship, and provide long-lasting economic, military, and public-health security.

We do not know enough to justify making changes

Briefly, those that worry about the social and economic costs of reducing greenhouse gas emissions often argue that the science is still too uncertain to justify plunging the world into the dramatic changes that would result from a move away from fossil fuels. Again, financial disaster would not necessarily be the outcome in any case—like environmental science, economic forecasting is itself plagued by uncertainty. In somewhat oversimplified terms, it comes down to a conflict between those who think we should wait for more science before doing anything and those who do not want to risk waiting any longer to respond to the threat of climate change despite lingering uncertainties and some unanswered scientific questions.

We cannot do anything about it anyway

There are some who either think global climate change is entirely the result of natural forces beyond our control or is a problem that has already gone too far for us to remedy. We have addressed the first argument once before. As to the second, current climate forecasting considers various scenarios based on different anthropogenic GHG emissions during the coming decades and concludes that we still have an ability to slow warming if we take steps now. Most climate scientists acknowledge that some additional warming is inevitable at this point—something like 2 °C within a hundred years—but predict it will be much more severe—up to 6 °C during the same timeframe—if we do nothing.

Other, non-scientific, objections

Many other arguments having nothing to do with objective science, those based in values, beliefs, opinions, and attitudes, also underlie denial and skepticism about the reality of climate change. These are often deeply entrenched and difficult to overcome. For example, it is very hard to convince people that the data and evidence supporting conclusions about climate change should be taken seriously if they believe scientists are inherently unethical or assume the whole story is a hoax perpetrated by the news media.

Denial matters. The solutions to human-caused climate change have been very difficult to implement for many reasons. Although the steps needed are seemingly straight forward, that is, greatly reduce GHG emissions, getting to an economy based on non-carbon fuels is of course not at all easy and requires people to make some tough choices. Vocal and influential deniers of the evidence-based conclusions of scientists make the necessary policy changes that much harder to bring about because they cast doubt on the notion that the short-term challenges are truly worth the effort.

Now what?

Acknowledging information gaps and uncertainty, data strongly indicate human activities have increased Earth's average temperature and brought about climate change. Whether the resultant adverse consequences listed above come to pass tomorrow depend on choices made by us today.

14.2.4. Ozone depletion

Our second global-level concern does not get as much attention as global warming these days, but it is still quite relevant. Here we briefly explore the causes, consequences, and remediation of this nuanced and complex problem.

First things first: "ozone depletion" and "global warming" are not synonymous!

Despite what is widely believed, these two global-level air pollution issues are not the same. As we just learned, warming is caused by the accumulation of greenhouse gases and is linked to climate change, rising sea levels, and many other consequences. This second problem involves the loss of an important atmospheric gas, ozone. Depletion of ozone results from different mechanisms than does climate change, and the adverse effects of these two phenomena are distinct. Most importantly, ozone depletion should *not* be understood to be the primary cause of global warming.

A second clarification: it is not really a hole.

We should also address a second misconception before we get to the details of this topic. Although the term "**ozone hole**" is often used to describe ozone depletion, it is not the best way to visualize the situation. As we will learn shortly, depletion refers to a *decline* in the amount of stratospheric ozone over a given region of the Earth's surface, not an actual opening in which *zero* ozone is present. This clarification should not be read as an attempt to minimize the urgency of this problem, because a reduction in ozone levels can lead to serious, in some cases even deadly, outcomes.

Ozone gas: the essentials

Our discussion of ozone depletion requires a little chemistry background.

Chemistry. Ozone has the formula O_3 and is made up of three oxygen atoms as shown in Figure 14.27.

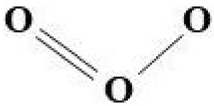


Figure 14.27. The chemical structure of ozone.

It is very reactive, which can be understood to mean that it has the tendency to readily interact with and alter things it contacts.

Where is it? We saw how it is a ground-level (i.e., tropospheric) secondary pollutant in the discussion of smog, but the focus of the current section is **stratospheric ozone** found high above the surface. The O_3 molecules located between altitudes of about 10 and 50 kilometers[36] make up what is known as the **ozone layer**, although the name is a bit misleading because the gas is spread out within a large vertical column of the atmosphere. If consolidated at the surface, all the ozone molecules dispersed vertically within that 40-km column of atmosphere could be compressed to a width of just a few millimeters.

Cycle of formation and destruction. Existing O_2 molecules produced from photosynthesis move to the stratosphere and are split into two single oxygen atoms. Each O then binds with another O_2 molecule to form a total of two new O_3. Ozone molecules can then be struck by ultraviolet radiation (UV) from the sun and split into a dioxygen molecule and a single O (a crucial reaction, as we will see). The cycle repeats indefinitely (Reaction 14.3.a and 14.3.b).

(Reaction 14.3.a) $O_2 \rightarrow O + O$ ***then, for each* O,**

(Reaction 14.3.b) $O + O_2 \rightarrow O_3 \rightarrow O_2 + O$

Uneven distribution. Complex global circulation patterns in the atmosphere lead to a dynamic and heterogeneous distribution of ozone. Research has revealed two important factors affecting O_3 levels in the stratosphere.

Latitude

The amount of ozone in the stratosphere is not uniform around

36. National Oceanographic and Atmospheric Administration (NOAA). Stratospheric Ozone, Monitoring and Research in NOAA. www.ozonelayer.noaa.gov

Earth. Simply put, movement due to prevailing winds leads to the highest O_3 concentrations near the poles and lowest near the equator (Figure 14.28). It is a complicated story, though, because concentrations also vary with time (next paragraph).

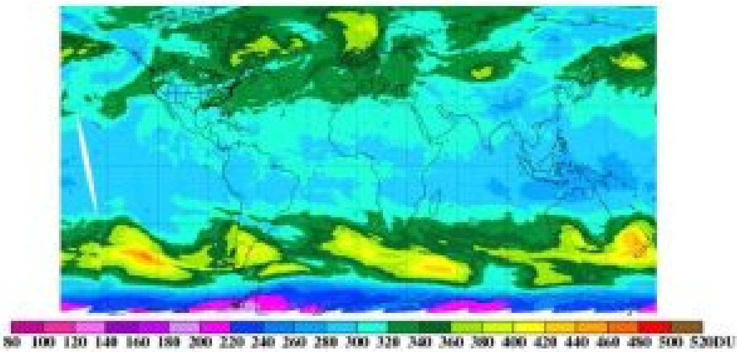


Figure 14.28. *The global distribution of stratospheric ozone in August 2018. Values are in DU, Dobson Units. US NOAA, Public Domain.*

The Dobson Unit is used to measure ozone concentration in the atmosphere: 1 DU is equal to the number of molecules needed to make a layer of ozone that would be 0.01 mm thick at the Earth's surface[37].

Time

In the short term, ozone levels fluctuate with season. Near Antarctica, for example, changes in temperature and other atmospheric properties combine to reduce O_3 concentrations to their lowest points during September, October, and November, and then they rebound in December. The magnitude of these swings is dramatic, as levels rise or fall by one third or more from season to season. Changes are less pronounced at lower latitudes. In the long term, the 11-year sunspot cycle we encountered in the section on

37. esrl.noaa.gov/csd/assessments/ozone/2006/chapters/Q9.pdf

global warming contributes to variations in ozone concentration. Since O_3 production is linked to UV-induced destruction of O_2 molecules, average annual ozone levels tend to be at their maxima during times of highest solar energy output[38]. The challenge, then, is to detect changes in O_3 concentration that exceed normal variations (i.e., are above **background fluctuations**).

Protective action of stratospheric ozone. Ozone in the stratosphere performs a critical service for the **biosphere**: it absorbs harmful incoming **ultraviolet** rays and prevents them from reaching Earth's surface (Figure 14.29).

Note that when one of the O atoms is knocked off an ozone, yielding O_2 and O, the UV energy that drove the reaction is converted to thermal energy. In other words, dangerous ultraviolet energy from the sun never makes it to ground level and instead heats the stratosphere. Why does this process matter? Without the ozone shield, life as we know it could not survive in terrestrial environments: intense UV energy would pass through the atmosphere and destroy the living cells it strikes (aquatic organisms living sufficiently deep would survive because water also absorbs UV). See Box 14.4 for more about the history and relevance of ozone in the stratosphere.

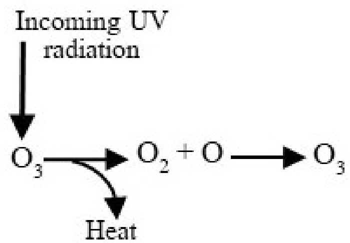


Figure 14.29. *A simplified diagram of how O_3 absorbs harmful UV radiation. Note the incoming radiation is converted to heat as an O_3 is destroyed.* Kelsey, CC BY-NC-SA.

38. National Oceanographic and Atmospheric Administration (NOAA). Twenty Questions. esrl.noaa.gov/csd/assessments/ozone/2006/chapters/Q9.pdf

Box 14.4. Enjoying your ozone? (or: Box 1.1 reprised)

Remember Box 1.1? In a nutshell, it describes how our oxygenated atmosphere formed and why the earliest organisms were necessarily **anaerobic**. Here we can add that they were restricted to **aquatic** habitats because the lack of an ozone layer required those ancient creatures to live under water. Once **photosynthesis** developed, though, everything changed. We already know that the presence of dioxygen gas allowed for the evolution of **aerobic** organisms, but you should also realize that it led to the production and accumulation of O_3 (via the processes described in Reaction 14.3). With the formation of the ozone layer, organisms did not need water to protect them from damaging UV and began to live on land. Once again we see the intimate connection between organisms and their environments!

Ozone depletion was noted decades ago.

The first hints were seen in the late 1950s and 1960s[39]. Using ground-based monitoring systems, researchers first recorded lower-than-expected levels of the gas near the South Pole in the winter of 1957. But the data were plagued by uncertainty and misunderstanding and deemed inconclusive at the time. Only in retrospect has the relevance of those early findings been

39. NASA Ozone Watch. esrl.noaa.gov/csd/assessments/ozone/2006/chapters/Q9.pdf

recognized. During the 1960s and early 1970s, a few researchers hypothesized that certain anthropogenic compounds detected in the atmosphere, today broadly classified as **ozone-depleting substances (ODS)**, could bring on the destruction of O_3 molecules (more about these chemicals and their actions will be presented shortly).

Initial reactions by government, industry, and the public were hostile. Although based on models, limited sampling, and laboratory experiments, the early science was compelling enough to capture the attention of many with an interest in ODS. Manufacturers of the suspect compounds reacted antagonistically to the science while many people who used the compounds started to question if they were damaging Earth's environments and endangering future generations. Political conflicts, advocacy on all sides, and worries about change, economic consequences, and inconvenience intensified[40]. The United States, home to a large fraction of producers and users of the controversial products, began to restrict certain applications of the compounds by the late 1970s even as research continued.

Science confirmed ozone loses in the 1980s[41]. During the 1980s, direct measurements over Antarctica verified what many already feared: the amount of ozone gas in the stratosphere was indeed decreasing. Even accounting for the background fluctuations described above, minimum concentrations of O_3 dropped below 220 DU, the lowest level recorded prior to 1979 (and one assumed to result from anthropogenic influence on an otherwise natural cycle). Subsequent measurements revealed a continued decline into the early 2000s (Figure 14.30, top). At the same time, the size of the area

40. Perhaps this story sounds familiar? Does it remind you of the current debate about and denial of climate change science?

41. NASA Ozone Watch. esrl.noaa.gov/csd/assessments/ozone/2006/chapters/Q9.pdf

of reduced O_3 levels, the so-called ozone hole, got bigger and bigger (Figure 14.30, bottom).

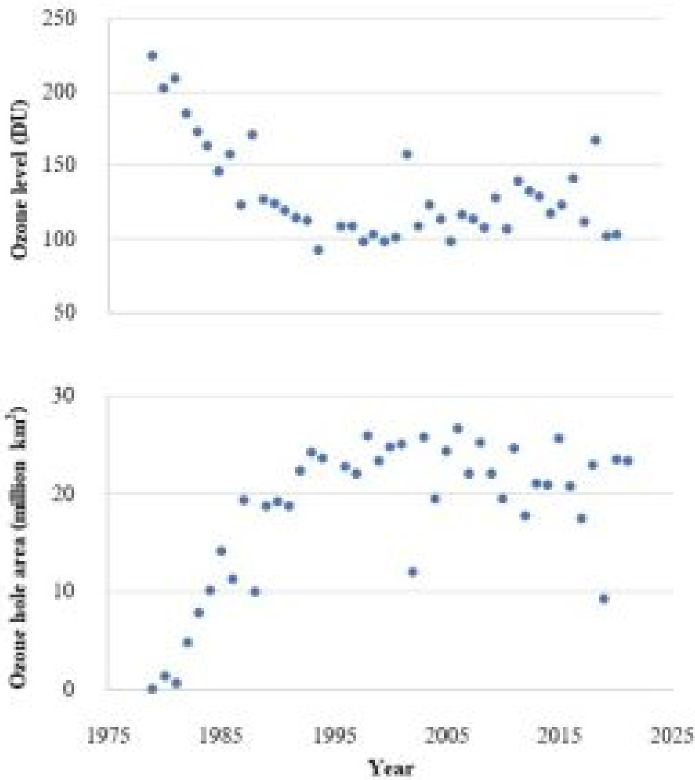


Figure 14.30. *Monitoring ozone over Antarctica. Top: decline in stratospheric ozone concentration since the 1970s. Bottom: increase in the area of the stratosphere affected by reduced ozone since about 1980 (expressed as square km). Data from US NASA, Public Domain, figure by Kelsey,* CC BY-NC-SA.

More about ozone-depleting substances

Although natural emissions—such as those from volcanoes—do play a minor role, research conducted in the 1980s strongly and clearly indicated that certain anthropogenic substances, two of which we will explore here, were responsible for most of the ozone depletion observed since the 1950s[42,43].

Chlorofluorocarbons (CFCs). This large class of ODS will receive most of our attention due to the enormous role it plays in the depletion of ozone from the stratosphere. In an unrelated but important point, we saw compounds in this group previously because they are very powerful greenhouse gases.

Chemistry

Chlorofluorocarbons are built of various combinations of chlorine (Cl), fluorine (F), and carbon (C) atoms. Figure 14.31 shows two examples.

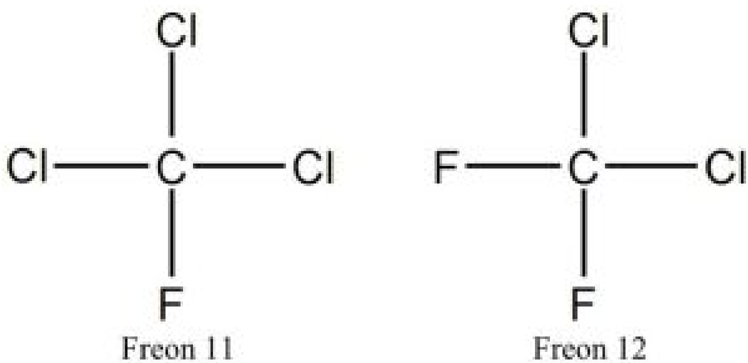


Figure 14.31. *The chemical structures of two common CFCs. Stable in the troposphere, they break apart in the stratosphere and release Cl (see the main text). Public Domain.*

42. NASA Ozone Watch. esrl.noaa.gov/csd/assessments/ozone/2006/chapters/Q9.pdf

43. NASA Earth Observatory. earthobservatory.nasa.gov/world-of-change/Ozone

Applications

CFCs were once widely used as refrigerants in air conditioners and freezers, insulators, lubricants in engines, and aerosol propellants in cosmetics, medical devices, and paints[44]. The reason for this pervasiveness is straight forward: CFCs are both effective and non-reactive compounds. In other words, they not only perform their assigned tasks very well, they are unlikely to ignite or explode. This stability is clearly desirable in an appliance powered by electricity—for example, your refrigerator will not blow up if a spark accidently contacts the coolant gas. As we will see, new compounds have taken over the roles formerly carried out by CFCs, but historic emissions continue to influence stratospheric ozone.

How CFCs disrupt the natural equilibrium and contribute to ozone loss

Chemical stability. The same property that makes them so useful is also a big problem once they are released into the air. Unfortunately, they tend to persist in the troposphere because the conditions in the lower atmosphere are simply not conducive to their destruction. Residence time of individual CFCs varies, but some can remain in the troposphere for many decades.

Mobility, then breakdown. Complex circulation patterns in the atmosphere can carry CFCs up to the stratosphere. The intense ultraviolet radiation in this zone (unlike the conditions in the troposphere) separates them into their component chemicals, including extremely reactive chlorine atoms (the first "C" in the name "CFCs"). These freed Cl atoms, then, are responsible for the

44. National Oceanographic and Atmospheric Administration (NOAA). Global Monitoring Laboratory. Chlorofluorocarbons (CFCs). gml.noaa.gov/hats/publictn/elkins/cfcs.html#:~:text=After%20World%20War%20II%2C%20CFCs,%2C%20homes%2C%20and%20office%20buildings.

breakdown of O_3 molecules. Reaction 14.4 shows one example of Cl-induced destruction of ozone.

(Reaction 14.4) $2O_3 + Cl \rightarrow 3O_2 + Cl$

> *CFCs are not directly responsible for ozone depletion, rather, they act as a source of chlorine atoms that ultimately catalyze ozone-destroying reactions.*

Note the Cl atoms are called **catalysts** because they are not altered in the reaction—they only push ozone destruction forward and then are released, free to catalyze thousands of additional reactions until they are removed from the stratosphere by one of several slow, natural processes. The central idea here is that Cl atoms bring on a net loss of O_3 molecules and a net increase in O_2 within the stratosphere. Thus, the long-term natural equilibrium has been changed by human releases of chlorine-carrying gases such that ozone destruction outpaces ozone production (Figure 14.32)[45].

Figure 14.32. *The rate of output (destruction) of ozone from the stratosphere exceeded the rate of input (formation) to it, so the amount of the gas there declined during the past several decades.* Kelsey, CC BY-NC-SA.

45. National Oceanographic and Atmospheric Administration (NOAA). Twenty Questions. esrl.noaa.gov/csd/assessments/ozone/2006/chapters/Q9.pdf

Bromine-containing gases. The element bromine (Br) is like chlorine in many important ways—both are **halogens**, a group of five elements with similar properties. Through a mechanism resembling the one described for chlorine, Br moves to the stratosphere and is released from the larger compounds in which it is a component. It then can catalyze the destruction of O_3 molecules. For example, $CBrClF_2$, a compound used in older fire extinguishers, once was an important source of reactive bromine atoms to the stratosphere. The pesticide methyl bromide (CH_3Br), primarily used in agriculture, was another. Note that the previous sentences are written in the past tense because those compounds, like CFCs, have been largely phased out and replaced with substances that perform the same tasks but cause less ozone depletion (we will return to the topic of substitutes for ODS shortly).

Others. The chlorine and bromine gases described above, known as **halogen source gases,** make the largest anthropogenic contributions to O_3 depletion. Swimming-pool disinfection, fossil fuel combustion, industrial processes, and other human activities can emit small amounts of depleting chemicals as well. Again, to minimize ozone damage, most of these processes now use different compounds.

The threat of ozone depletion

A net loss of O_3 molecules from the stratosphere allows more harmful ultraviolet radiation to move to Earth's surface and damage biological entities. Some models predict that, without appropriate corrective action, levels of UV would have reached high-enough levels to profoundly change the biosphere within a few decades. Here we consider some of the most important potential consequences.

Harm to human health.

Skin cancer

Skin cancers can be caused by the type of ultraviolet radiation absorbed by ozone[46]. Because more UV makes it to the surface with increasing destruction of stratospheric O_3 molecules, higher rates of melanoma and other conditions are a concern. According to the United States Centers for Disease Control and Prevention (U.S. CDC), the rate of appearance of new skin cancers among Americans has increased in the past two decades, with the total number of new cases doubling between 1999 and 2019[47]. Now, many factors likely contribute to this observed trend, including changes in lifestyles, expectations, beach behavior, and so on, but loss of protective ozone during the same period is thought by many to play an important role.

Cataracts

This sight-limiting condition can be brought on by many agents, including exposure to ultraviolet radiation.

Immune disorders

Some evidence suggests UV radiation can alter immune function, increasing susceptibility to diseases. Further research must be done before definitive conclusions can be made about this consequence.

Harm to natural ecosystems.

Non-human organisms are also susceptible to UV-induced damage.

Effects on primary producers

Research has shown that terrestrial plants, even agricultural crops, can be adversely affected by ultraviolet radiation. As a result, organisms feeding directly or indirectly on these producers could also be harmed. Phytoplankton, aquatic primary producers, are

46. U.S. Environmental Protection Agency. Ozone Layer Protection. www.epa.gov/ozone-layer-protection

47. U.S. Centers for Disease Control and Prevention (CDC). Melanoma of the Skin Statistics. www.cdc.gov/cancer/skin/statistics/index.htm

similarly sensitive to increased UV levels because they live so close to the surface of water (and, therefore, are not deep enough to be protected). As we saw in our discussion of climate change, a loss in phytoplankton would devastate terrestrial and marine ecosystems.

Additional ecological effects

Organisms vary in their ability to tolerate exposure to ultraviolet radiation. High amounts of hair, fur, melanin, and various other structures provide some protection, so we would expect those having the appropriate adaptations to survive and reproduce more effectively than those that lack them. As we have seen throughout this textbook, **ecosystem structure** depends on environmental conditions, and we should expect that increasing amounts of UV reaching the surface to be like temperature, moisture, and all the others: it will influence the identity of dominant organisms throughout the biosphere. Eventually, if ozone levels were to drop sufficiently, terrestrial systems could become too hostile for any life forms—just as they were in the early days of this planet.

We have taken steps to restore ozone

Decades of research indicate that stratospheric-O_3 concentration declined between the late 1960s and early 2000s. Moreover, although a few critics argued otherwise, the link between this loss and anthropogenic activity is strong. As emissions of ODS increased, the rate at which ozone was depleted became increasingly high relative to the rate at which it was produced. The chemical mechanism of ozone depletion was also understood. Clearly, this was a real phenomenon caused by humans, and corrective action was needed.

The Montreal Protocol: an international response. To combat the threat, 24 nations (plus the European Union) signed The Montreal Protocol on Substances that Deplete the Ozone Layer in 1987 (eventually, all UN members joined). Notably, signatories pledged to stop using CFCs and other O_3-depleting compounds.

The agreement evolved with time (and science), and many amendments were added during subsequent decades. The initial timeline for developed countries to phase out certain CFCs by the year 2000 (2010 for developing countries) was pushed forward to 1996, with some participants, including the United States, beating the deadlines.

Given the utility and widespread use of some of the ODS, adoption of the rules of the Protocol required a substantial commitment. Careful recycling and other steps were taken to reduce the release of existing chlorofluorocarbons into the air. Compounds that could substitute for CFCs without damaging ozone were also developed and slowly introduced into use (some of these are considered transitional because they only slow, rather than halt, O_3 destruction and will themselves need to be replaced in the future). In the United States, for example, laws requiring non-ODS coolants in cars, refrigerators, and air conditioners were enacted during the 1990s. Other countries have taken similar steps.

Antarctic ozone levels have started to recover. It appears that the phasing out of ODS has allowed stratospheric ozone levels to slowly rebound. Recent measurements indicate O_3 concentrations over Antarctica are rising. If the trends continue, the layer will recover to 1980s levels by the year 2065[48,49]. We should proceed with caution, however. It is early in the process, O_3 is still below its

48. United Nations Environment Programme. Scientific Assessment of Ozone Depletion: 2022. Montreal Protocol on Ozone Depleting Substances quadrennial assessment report. ozone.unep.org/system/files/documents/Scientific-Assessment-of-Ozone-Depletion-2022-Executive-Summary.pdf

49. National Oceanographic and Atmospheric Administration (NOAA). Path to recovery of ozone layer passes significant milestone. research.noaa.gov/article/ArtMID/587/ArticleID/2900/Path-to-recovery-of-ozone-layer-passes-a-significant-milestone

pre-1970 levels, and the long-lived ODS already in the troposphere will likely contribute to ozone destruction for decades. Also of concern is the possibility that CFCs or other problematic compounds could be released in the future by nations that either did not sign the Montreal Protocol or choose not to adhere to it. Be all that as it may, many people see the international response to the shared concern of ozone depletion, notably the cooperation among scientists, policy makers, governments, and members of the public, as a model that could be followed for future agreements about greenhouse gas emissions, global climate change, and other threats to the world.

14.2.5. More about reduction and regulation of outdoor air pollution

We end this section with a brief look at some of the strategies used to control air pollution.

Ways to limit emissions

Approaches to reduce the release of air pollutants generally fall into one of three categories: those that reduce emissions by increasing efficiency of energy-using technology, those that clean up pollutants before they move into the atmosphere, and those that rely on low- or no-pollution energy sources.

Increase efficiency. In this case, machines, vehicles, and appliances are designed to require less power to perform the same services we have always gotten from them. For example, televisions, refrigerators, and lights that provide entertainment, cold storage, and illumination, respectively, with less electricity than older models of the same devices lead to less burning of coal (a commonly used source of energy; see Chapter 10) and therefore lower levels of emissions than previously seen. Similarly, less gasoline is burned—meaning fewer pollutants are released—if more kilometers can be traveled per liter of fuel (i.e., miles per gallon).

Clean up emissions. Instead of improving efficiency, air pollutants can be broken down or otherwise captured before they reach the atmosphere. A common example of this strategy is the use of **catalytic converters** in automobiles. In short, exhaust passes through one of these devices—essentially, a small metal box—as it travels from the engine to the tailpipe (Figure 14.33).

The converter contains metals like platinum and rhodium that speed up the conversion of certain pollutants produced from gasoline combustion, such as carbon monoxide and nitric oxide, into water and carbon dioxide (that is, the metals act as catalysts[50]). As part of the Federal Clean Air Act, the United States has mandated the use of catalytic converters on new cars for nearly 40 years, and many other countries have enacted similar laws. Arguably, this strategy has been successful in reducing certain types of local and regional air pollution, but it came with a very tangible price: converters add to the cost of a car. Emissions from coal-fired power plants can be reduced with analogous strategies, although in some cases the costs to install pollution controls are so high that affected companies choose to cease their use of coal in favor of cleaner fuels.

Figure 14.33. *A simplified diagram of a car showing the position of the catalytic converter between the engine and the tailpipe. Kelsey,* CC BY-NC-SA.

Use alternative energy sources. The combustion of non-renewable fossil fuels releases many more toxic products than does the use of renewables such as solar and wind power (Chapter 10).

50. US Energy Information Administration. Gasoline explained. eia.gov

So, a third strategy is to replace conventional power generation technology, say, coal-burning electricity facilities, with something like wind farms. Phasing out carbon-based fuels in general would vastly reduce most of the adverse consequences of air pollution we have explored in this chapter. Chapter 10 describes several other benefits that a switch away from the dominant energy sources would bring, as well as the obstacles that have impeded wide-spread changes to alternative fuels.

How do we get polluters to make necessary changes?

Standards and expectations regarding air quality differ among countries, states, and localities. Effects of air pollution also vary, depending on local climate, environmental conditions, population density, and of course the specific substances of concern. As a result, regulations on anthropogenic air pollutants range in their scope, rigor, and underlying assumptions. Here we take a quick look at some commonly used approaches.

Direct regulation. In this case, governments pass laws that set emissions limits on all entities releasing a target pollutant. The catalytic converters we saw above fall into this category: they are required on every new car sold in the United States. Other sources, both mobile and stationary, can be controlled with direct regulation. In most cases, those that do not comply are punished with a fine.

Cap and trade. A second approach establishes a limit on the total amount of a pollutant released—that is, the **cap**—by a group of sources in a region. Then, individual members of this group buy the right or **allowance** to emit a certain amount of that substance each year. The total cap is lowered with passing years, compelling the entire class of polluters to reduce its emissions by whatever means it chooses (for example, using one or more of the three ways noted just above). Individual companies are incentivized to reduce their emissions as soon and as much as possible because, as less pollutant is released, fewer allowances are required. Moreover, those that have taken steps to innovate and clean up their emissions can sell their unneeded allowances to those that have not done as much to reduce pollutants. This market-based style of control provides far more flexibility than does direct regulation and is supported by many. Cap and trade was used to decrease the amount of sulfur dioxide released by coal-powered electricity generators in the United States during the 1990s, thereby reducing acid precipitation. The success of this effort to cut regional air pollution has inspired a movement to employ a similar model to lower carbon dioxide levels in the troposphere and combat global climate change. Not everybody is so ready to embrace such a policy, though, and debate about it has gone on for many years.

14.3. INDOOR AIR POLLUTION

Outdoor environments are not our only concern. People working and living in offices, schools, factories, laboratories, homes, and other interior areas can also be exposed to dangerous levels of airborne toxins. We close Chapter 14 with a few words about this serious public health problem.

14.3.1. A uniquely susceptible environment

Enclosed spaces are particularly vulnerable to air pollution by design (if not intention). To limit fluctuations in temperature and humidity without an undue financial cost, circulation with outside air is minimized. This situation can be understood through a familiar series of events: when air conditioning or heating is turned on, all doors and windows are tightly closed. So, any toxins present can become highly concentrated.

14.3.2. Many potential pollutants

As we saw with outdoor environments, the list of toxins here is long and diverse and includes chemical, physical, and biological entities. The possible adverse health outcomes caused by indoor air pollution also are quite numerous and varied.

Chemical

Chemical toxins come from many sources: solvents and adhesives, ozone from electrical appliances, fumes from newly painted walls or installed carpets, lead from old and peeling paint, insecticides used to control ants and the like, bathroom cleaners, cigarette smoke, perfumes, and many others. Depending upon the nature and degree of the odors these toxins emit, people may or may not know they are being exposed until symptoms appear. Responses vary, but burning of skin, eyes, or lungs, shortness of breath, nausea, blurred vision, light headedness, allergic reaction, asphyxiation, heart failure, and other outcomes are possible.

Physical

This category includes particulates like sawdust, asbestos fibers, fine mists, glass shards, and similar materials. These substances generally induce different responses than do chemical pollutants because they directly injure tissues. Asbestos, for example, can become embedded deep inside lungs and ultimately lead to cancer. Dusts and shards can cause serious abrasions throughout the respiratory system.

Biological

Viruses, pathogenic bacteria, and mold spores are the principle concerns here. As we learned in Chapter 3, the first two entities on the list can cause human diseases. If they become airborne and are adapted to survive long enough outside a human body, they can move among people and spread diseases like the common cold, COVID-19, influenza, pneumonia, and many others. Perhaps the most vivid example of this phenomenon is seen on commercial airplanes. If one passenger boards with a head cold, the virus that caused it may spread to anyone else inside the cabin who happens to breathe the same air (i.e., everybody). Not all present will necessarily get sick, but it is likely many will carry more than just their luggage when then exit the plane. Mold spores that cause allergic reactions, the final item on the list, are relevant in spaces that are dark and damp.

14.3.3. Mitigating indoor air pollution

Steps may be taken to reduce indoor air pollution. First, treatment and filtration systems installed inside air handlers can minimize the challenges caused by limited air circulation. Given the diverse nature of potential pollutants, though, coming up with remedies for all the potential problems can be complicated and costly. Second, biological entities can be at least partially controlled by changing the environmental conditions of spaces to discourage these unwanted organisms. Better lighting, repair of water leaks, installation of dehumidifiers, and application of chemical disinfectants are among the strategies used.

14.3.4. Some final considerations

Not all indoor spaces are affected by the same level of air pollution. A short list of the variables that account for the differences is presented here.

Year of construction

Since rules have changed with time, generally getting increasingly strict during the past several decades, the age of a building is a very important factor. For example, leaded paint and asbestos, two materials banned in the United States because of the adverse effects they have on human health, were both widely used until the 1970s. In other words, older buildings will likely face different problems than newer ones.

Location, design, and materials used

Regardless of age, certain buildings are just more prone to indoor air pollution than others. Certainly, damp and dark regions will likely be characterized by more mold problems than dry and sunny ones. Also very important, though, are the decisions made before, during, and after construction. If costs are cut on quality of materials, expertise of contractors, maintenance, and appropriateness of infrastructure, including air handling systems, users and residents of buildings could experience long-term exposure to toxic substances in their homes or workplaces. In some cases, illnesses affect multiple people occupying the same space, for example employees housed in a single complex. This phenomenon, often called **sick-building syndrome**, can create a great deal of protracted suffering and conflict because owners / operators of suspect spaces are generally loathe to admit deficiencies in air quality (and subsequently pay the high costs required to fix the problem), often dismissing claims that workplace exposure is responsible for observed adverse consequences. Indeed, it can be very difficult for scientists and regulators to establish a causative relationship between symptoms and careless, irresponsible, or even criminal actions made by those responsible for a building. We will more fully explore the effects of several common toxins on human health in the final chapter in this book.

THE CHAPTER ESSENCE IN BRIEF [51]

Many human activities release pollutants into the atmosphere. Fossil fuel combustion is of particular concern, as it contributes to multiple local-, regional-, and global-scale effects. Much scientific evidence has linked human activity to one critical effect in particular, the warming of Earth's troposphere and consequent climate change. Stopping or even reversing the observed changes is difficult for multiple reasons, but reticence about switching away from carbon-dioxide-emitting power generation is among the biggest obstacles.

---

51. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 14 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[52]

In what ways are air pollution and water pollution similar? Different?

Can the same source of pollution contribute to local-, regional-, and global-scale concerns? How?

Carbon dioxide is a natural product (review Chapter 4, for example). Is it appropriate to categorize it as an air pollutant? Why or why not?

Drawing on the principle of environmental unity once again, connect the dots between melting arctic ice and low fuel efficiency in passenger cars (you might want to consult Chapter 10 as you ponder your answer). Here is another one: think about the way increasing standards of living worldwide (Chapter 8) could increase the amount of material in landfills…and ultimately lead to higher average global temperatures.

Are forests sources or sinks of carbon dioxide? Think back to Chapter 12 as you ponder your answer.

52. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

Why are the data-based connections between human activity and climate change so hard for some people to acknowledge? Arguably, denial of evidence has become more adamant in recent years, even as the science has become increasingly compelling and clear. What is going on? Can you think of any way to reach those folks who cannot accept the reality of human-driven climate change?

How can ozone gas be classified as an air pollutant in one instance and a necessary component of the atmosphere in another?

15. Human-Health and Environmental Toxicology

JASON KELSEY

Toxicology is a wide-ranging and important science that profoundly affects our daily lives. Toxicologists, the scientists who work in the field, draw upon knowledge from multiple areas, including biology, physiology, chemistry, mathematics, physics, soil science, and ecology to study the effects of poisons on biological systems.

Before proceeding any further into this final chapter, we will first take a moment to reflect on some of the ideas we have considered thus far. Clearly a few sentences cannot fully summarize 14 chapters of a textbook, but some common themes and trends stand out. First, living and non-living systems are inextricably bound together: as we know well, they influence each other in multiple, complex ways. Second, **anthropogenic** activity, particularly that of the industrialized era, has profoundly affected Earth's environments. Third, people depend on natural processes for survival as well as health, prosperity, happiness, and other hard-to-quantify benefits.

Why the sudden walk down memory lane? Well, the science of toxicology touches on and grows from many of those ideas we have studied already. For instance, as humans seek to fulfill our expanding demands for living space, long life, transportation, sustenance, and power we introduce many foreign substances, some of which are harmful, into our bodies, as well as to soil, air, and water. Put another way, the chemicals that are so integral to our modern lives also have the potential to adversely affect the health of both human and natural ecological systems.

Key concepts

After reading Chapter 15, you should understand the following:

- The long history and continuing relevance of the science of toxicology
- The fundamental principles of toxicology, including dose-response and selective toxicity
- How toxicological data are collected and expressed
- How poisons enter and move within organisms, are metabolized, and are excreted
- The multiple potential consequences of exposure to poisons
- The principles of environmental toxicology, including how toxic substances can affect ecosystems
- The fundamental principles of risk assessment and how they can be used to improve human safety

Chapter contents

15.1. TOXICOLOGY HAS ANCIENT ROOTS

Humans surely have been interested in poisons for as long as they have understood cause and effect. The first people to connect a red, itchy skin rash with rubbing up against something like poison ivy or stomach cramps with ingesting wild berries were, arguably, the earliest toxicologists. Later and more formally, the ancient Egyptians, Greeks, and Romans methodically studied and then employed poisons in warfare, assassinations, and executions. Others, such as early indigenous North Americans, harnessed the power of natural substances like peyote not to cause harm, exactly, but to alter their mental state and consciousness. In these and other cases, people first learned how a wide range of substances can affect humans and then proceeded to make use of their accumulated knowledge. A full accounting of the history of toxicology is neither necessary nor practicable here. Suffice it to say, toxicology is among the oldest human sciences. See Box 15.1 for one quirky historical example of inspired toxicology gone spectacularly wrong.

Box 15.1. Warrior, king…toxicologist?

We take a moment now to consider the fascinating and grisly tale of a somewhat obscure historical figure known as Mithridates VI. This remarkable man was the leader of ancient Pontus, a country located on the Black Sea in the region of modern-day Turkey. He died during the last century B.C., but not before dramatically adding to our knowledge of the science of poisons as well as making quite a bit of trouble for the Romans (getting right to the point, so much so that they very much wanted to parade him through the streets of Rome before subjecting him to a horrible death). Mithridates, like many of the rulers that came before and after him, feared he would be assassinated by poisoning (his father seems to have died by that fate—at the hands of his mother). To thwart would-be usurpers, Mithridates developed a concoction of small amounts of common poisons that he ingested **prophylactically** for years. His intentional low-dose exposure effectively ramped up his natural defenses and resistance to a range of toxic substances, allowing him to survive, rule his people, and annoy Rome for decades. It really was a brilliant feat, one that influences toxicologists and physicians to this day. However, a downside to this act of inspired genius eventually reared its head. If we fast forward to his final hours, we find the mighty scientist under siege and trapped by a Roman army ready and excited to finally take him prisoner. Knowing the fate awaiting him upon capture, Mithridates decided to kill himself before it was too late.

The end of this story nearly writes itself: the superhuman resistance that worked so fantastically well for so long also protected him against any of the available poisons that could have killed him gently when he actually wanted to die. Sadly, he had to opt for the far more painful method of falling on his servant's sword to get the job done.

15.2. SOURCES OF POISONS

Hyperbolic as it may sound, poisons are all around us. We can encounter toxic substances whenever we inhale, eat, drink, or touch anything. Anthropogenic and natural processes, products, and activities, many of which we saw in the previous chapters of this book, are important sources of potentially harmful substances. Here we examine some noteworthy examples. First, however, a word of caution is in order: the source of a substance is no guarantee of safety or risk of it. Many people assume that "synthetic" and "natural" are synonyms for "dangerous" and "safe", respectively. The world is far too complex for such simple conclusions, as some artificially generated substances are life saving whereas, as we will see below, some natural substances are extraordinarily harmful to humans.

15.2.1. Anthropogenic

Agriculture

Recall from Chapter 9 that fertilizers and pesticides are used extensively in conventional agriculture. Yes, such chemicals generally enhance food production, but they also can harm living systems.

Fertilizers. Farmers often add nitrogen, phosphorous, and other nutrients to soil. In addition to stimulating crop growth, their movement away from farms can initiate **eutrophication** and pollute drinking water. Fertilizers provide an important and generalizable lesson, namely, the same chemical compounds that are so helpful in one context can be deadly in another.

Pesticides. This broad group of substances includes many specific categories, two of which are used widely in agriculture:

insecticides (they target fruit flies, locusts, and other insects) and herbicides (weeds). Others, including rodenticides (mice, rats, and the like), fungicides (mold), and molluskicides (snails and slugs), are also used to varying extents. Pesticides can be very effective in killing organisms that reduce crop yields, but they also can adversely affect non-target organisms, that is, those we want to protect form harm. For example, a compound deadly to corn-eating caterpillars may also damage pollinating bees, lady beetles, birds, fish, and humans.

Food preparation, packaging, and consumption

The many steps used to transport, modify, prepare, and sell farm products can introduce a variety of biological, chemical, and physical poisons into food. Some, like preservatives, colors, and flavors are intentionally added in attempts to enhance food in one or more ways; these are all classified as **additives**. On the other hand, metals, cleaning agents, microbial toxins, and so forth that end up in food inadvertently are **contaminants**. What about the pesticides we noted above? One could argue they have attributes of both additives and contaminants in that we knowingly add them to food crops, but we do not intend for humans to ingest them. However you wish to categorize them, detectable pesticide residues can and do end up on fruits and vegetables, although they tend to be present in small quantities and so pose a minimal risk (as noted in Chapter 9, farm workers exposed on the job are far more likely to experience pesticide-induced health effects than are consumers—see Figure 15.1 as well). The importance of the amount of exposure to, or more correctly, dose, of a poison, is a central idea in toxicology that we will explore in depth shortly.

Figure 15.1. *This farmer is exposed to a high dose of poison while spraying a pesticide (the white cloud) onto a field.* US EPA, *Public Domain.*

Drugs

This group includes the enormous number of natural and synthetic substances used by humans for **therapeutic** (i.e., to remedy a debilitating condition or illness) or **recreational** (i.e., they play no formal medicinal role but provide pleasure of some kind) purposes. Since a full exploration of the story of drugs would entail a great deal of time and space and divert us from our primary path, we will need to settle for a few brief comments on the topic. First, with very rare exceptions, people who take drugs do so to derive some benefit. For example, you might swallow one kind of pill to relieve a toothache and another to combat a bacterial infection in your ear, administer a few drops of liquid into your eyes to reduce redness or apply a cream to help heal damaged skin, inhale a mist to mitigate the symptoms of respiratory allergies, or inject insulin to treat diabetes. Of course, some people swallow, inhale, or inject amphetamines, cannabis, or heroin, respectively, to gratify other desires. Second, drugs like those just listed have primary, intended effects, but they also have the capacity to trigger **side effects**, outcomes that are at best inconvenient and at worst debilitating or even lethal. A toxicologist would point out that we intentionally expose ourselves to toxic substances with any drug we take, therefore, the risk of potential adverse reactions must be weighed against the potential benefits (review the concept of cost-benefit analyses in Chapter 1).

Cosmetics

***Figure* 15.2.** *Red lipstick can get its color from heavy metals that are easily absorbed through thin skin. Breakingpic, Public Domain.*

Cosmetics are commonly used substances intentionally applied to exterior portions of human bodies. Like drugs, these products carry risks—they can contain toxic substances that are readily absorbed through skin (things like metals and solvents)—and benefits, hard-to-objectively-quantify changes to one's appearance, odor, and so forth (Figure 15.2 shows a familiar example).

Power generation and usage

Many of the strategies we use to meet our demand for power, most notably the burning of fossil fuels and harnessing of nuclear fission, can release poisons into our air, water, and soil. The **smog** tied to coal combustion, for example, contains chemical and physical substances that damage human and environmental health. **Radiation** from nuclear power plants, although generally contained, also has the potential to cause widespread and severe adverse effects. Review Chapter 10 and Chapter 14 for more about the risks associated with power generation.

Waste production and management

Industrial, agricultural, military, medical, academic, and domestic activity all generate waste products. Steps are taken to minimize pollution by these materials (for example, **sewage treatment** facilities and **landfills**). Still, human and natural systems are exposed to chemical, biological, and physical poisons from accidental releases and improper handling of waste. Review Chapter 13 for more about waste management.

15.2.2. Natural

As we noted in Chapter 7, modern humans are not alone in their capacity to create hazards and threaten the well-being of living things. Volcanic eruptions (including the emission of toxic gases), earthquakes, mass wasting, flooding, and diseases are among the natural phenomena and forces that can devastate human and natural systems.

Figure 15.3. *The red tree frog (about the size of a thumb nail) releases a toxin to its skin that can damage animal nervous systems. Splette, CC BY-SA.*

Our list can quickly grow if we add poison-producing animals (think rattlesnakes, black widow spiders, Komodo dragons, and tree frogs, Figure 15.3), heavy metals like arsenic and mercury in rocks, soils, and waters, radon gas, and **ultra-violet radiation** (UV, see Chapter 14) striking us from the sun. Finally, we should give credit to toxins produced by **microorganisms**. Consider just one of them, botulinum toxin. This exquisitely poisonous substance is produced by a naturally occurring **bacterial** species and is more potent than any other known substance. It ought to be clear that Earth offered plenty of risks before we appeared on the scene and continues to do so.

15.3. FUNDAMENTALS OF MODERN TOXICOLOGY

Of course toxicologists are bound by the **scientific method** that we have seen throughout this book. Additionally, though, they are informed by several concepts that are unique to their science. Here we briefly explore the terms, assumptions, and principles needed to understand toxicology.

15.3.1. Speaking like a toxicologist: essential terms

Xenobiotic ('zeno bi ä dik')

Toxicologists frequently use this term to describe any substance that enters an organism from an external source. That is, xenobiotics *do not originate within the body of an affected organism.* Keep in mind that xenobiotics can bring about a range of reactions, from benefit to harm (some of these we will see shortly).

Poison

This word refers to a substance that causes harm to a biological entity. **Poison**, and another word, **toxin**, are commonly used interchangeably, though most toxicologists recognize subtle differences in their meaning. The latter, toxin, is more specific, including compounds that are produced by organisms (i.e., animals, plants, or microorganisms), and the former includes a broader range of substances from any source (i.e., they are natural or synthetic) that are harmful. Importantly, both terms suggest substances that induce adverse consequences in exposed organisms.

Dose

The concentration of a xenobiotic within an organism is known as the **dose**. It typically is expressed as the amount of the substance in an organism divided by the mass of the organism, i.e.,

amount of xenobiotic (mg) / mass of organism (kg).

Body burden may be used to refer to the total amount of a xenobiotic in an organism.

Response

Response is broadly defined as the consequence(s) caused by a xenobiotic. We also use the term **endpoint** to indicate a reaction to exposure. Keep in mind that a xenobiotic can initiate a harmful or beneficial response (in fact, as we will see in 15.3.2., the same compound may cause *both* detrimental and helpful outcomes). Among the many possible endpoints are paralysis, reproductive failure, skin irritation, loss of bone mass, blindness, and death (some details about response are presented in section 15.4). One more caveat is appropriate before we proceed: you should re-read the first sentence in this paragraph and reflect on the word “caused”. Toxicologists establish a mechanistic link between a xenobiotic and whatever response is observed (review the difference between correlation and causation in Chapter 2).

Potency

The dose required to bring about a response is a function of the potency of a substance. This property ranges enormously among xenobiotics. Certain forms of mercury, for example, have **high potency** in that l**ow concentrations** in the body can cause measurable and serious adverse effects. On the other hand, many vitamins and synthetic drugs have **low potency**, meaning very **high doses** are required to cause damage. Importantly, accidental overdoses from low-potency compounds are unlikely (although not impossible, as we will see).

Routes of exposure

Exposure refers generally to the critical event during which a person comes into contact with a xenobiotic. Why is exposure deemed "critical"? Because even the most potent toxin known poses no threat unless it is encountered. But we take things a step further than a simple question of whether there was exposure because the *specific way* a xenobiotic gets into an organism influences how much of and how quickly a potentially harmful substance moves to sensitive tissues like the brain, heart, liver, and others. As suggested in the earlier paragraph about drugs, there are four distinct routes by which a xenobiotic can move into an organism: via **gastrointestinal** ingestion (GI, you swallow it), **dermal** contact (through your skin), **inhalation** (you breathe it in), and **intravenously** (direct injection into your bloodstream). You should realize that none of the routes is necessarily more or less dangerous than the others. Whether dermal exposure, for example, poses less of a threat than does GI exposure depends on the specific properties of the xenobiotic in question. You will see a related term throughout this chapter, **uptake**, which is often used nonspecifically to refer to the entry of a xenobiotic into an organism. We will learn more about the relevance of, as well as the factors affecting, exposure shortly.

Target system

A drug or other potential poison has an impact on an organism *not* because it interacts with all systems at once, rather, xenobiotics tend to preferentially affect one (or possibly a few) sites within the body that toxicologists refer to as **targets**. Damage at targets then can lead to a variety of consequences (or responses, as noted above) ranging from mild to severe. There are many potential targets within the human body, including the central nervous system (CNS, i.e., the brain and nerves), the respiratory system (lungs and associated structure), the digestive system (stomach, intestines), the cardiovascular system (heart and blood distribution) eyes, ears, skin, bones, and many others. Certain forms of mercury, for instance, cause damage by affecting their target, the CNS. As a result of disruptions to the nervous system, a body may suffer from seizures, paralysis, and potentially, death. Excess fluoride, on the other hand, may cause more localized damage to teeth only (you read that right: we will see more about fluoride below).

15.3.2. Two underlying principles

Dose determines response

We begin this section with the **principle of dose-response**, an idea first described in the 16^{th} Century by a Swiss scientist named Philippus Aureolus Theophrastus Bombastus von Hohenheim (thankfully, known as **Paracelsus** to his friends as well as we modern toxicologists). Put simply, Paracelsus noted how the effects caused by a xenobiotic depend on the amount and concentration of that xenobiotic in a body. The story quickly gets nuanced from here because response tends to run the gamut through increasing benefit to increasing harm as dose goes up. For example, a pain killer like acetaminophen will provide no relief if it is taken at a very low dose but will dull your headache once the concentration of it in your blood reaches the appropriate therapeutic level (i.e., after you ingest enough of it). At higher and higher doses of that same drug, though, harm is done to multiple systems until, at the **lethal dose**, death from liver damage occurs. The same phenomenon is seen with other over-the-counter and prescription medications, even as the specific therapeutic and toxic doses will vary considerably (in case it does not go without saying: you should always read and follow the dosing instructions for any medication!). See Figure 15.4 for a graphical representation of the dose-response principle.

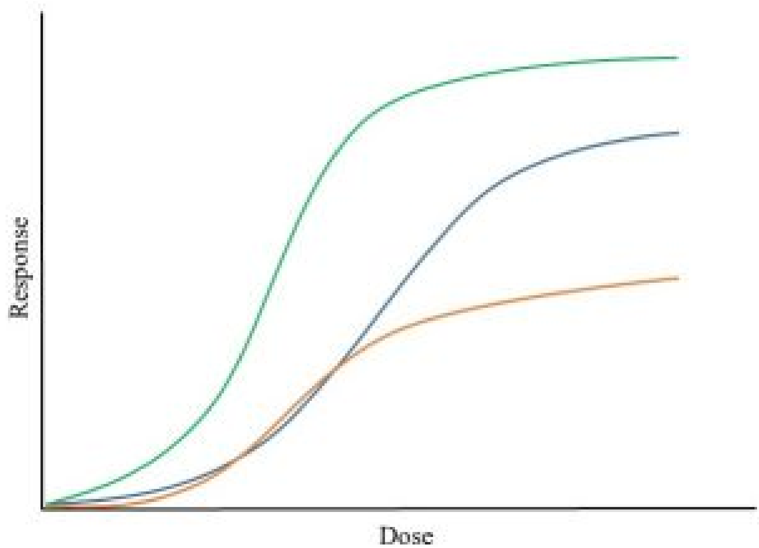


Figure 15.4*. The effect of dose on response for three different hypothetical compounds. Kelsey, CC BY-NC-SA.*

Now, things get even more complicated than we saw in Figure 15.4 because in many cases adverse response increases at *both* very low *and* very high doses. How can this be? Consider the relationship between dose and response for an essential nutrient like selenium. A deficiency of this metal—doses approaching 0 mg / kg—is harmful, meaning that likelihood of damage to the organism goes up as dose *goes down*. However, as selenium levels continue to increase beyond the therapeutic level adverse responses are seen again (see Figure 15.5 for the kind of dose-response curve we would see here).

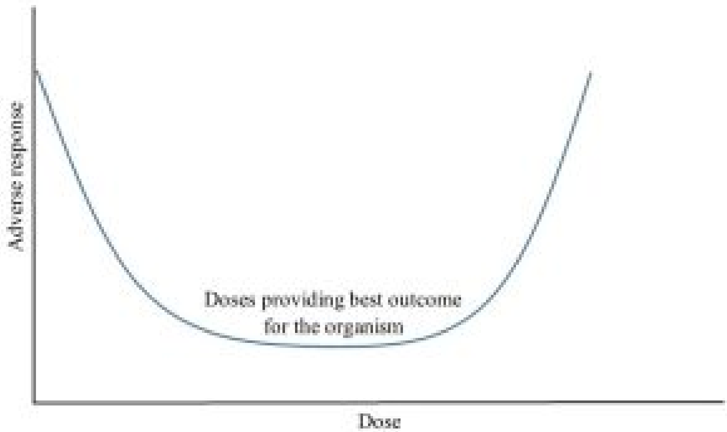


Figure 15.5. *Hypothetical dose-response curve showing how both low and high doses of a compound is harmful, the phenomenon known as hormesis. Kelsey,*

Similar U-shaped curves characterize the dose-response behavior of fluoride (too little and too much can lead to tooth decay) and ethanol (low doses may provide protection against heart disease whereas high doses increase risk from cancer and heart disease). In these and many other cases, there is a range of doses that is best, neither too high nor too low. This phenomenon, known as **hormesis**, is somewhat controversial within the scientific community for a number of reasons, including the fact that it is difficult to see how it could apply to everything. For instance, lead, mercury, and gamma radiation appear to be harmful at all doses; there is no range within which they provide benefit. Research in this area is ongoing.

Another wrinkle in the dose-response relationship merits a brief mention. We assume that relevant responses occur at all doses greater than 0 mg / kg, that is, even extremely low doses cause responses. Sometimes, however, responses are not evident until a dose appreciably above 0 is reached. Put another way, some minimum dose, one higher than 0, is required before any response is seen. Such xenobiotics are said to display a **threshold**, formally defined as the dose below which no response is observable (Figure 15.6).

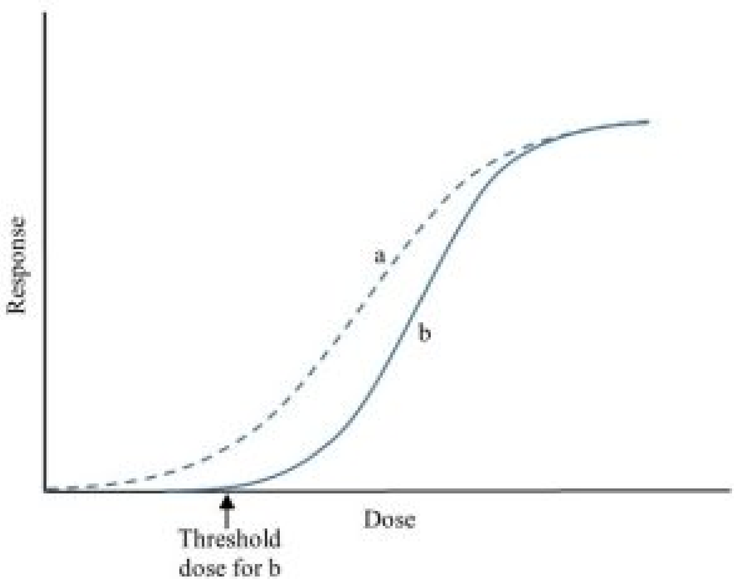


Figure 15.6. *Dose-response curves for two hypothetical compounds, a and b. Note that b displays a threshold whereas a does not.* Kelsey, CC BY-NC-SA.

The notion of a threshold may seem reasonable enough, but like hormesis, not all toxicologists accept it. Several objections have been raised, including a question that is embedded in the basic definition of the phenomenon: the dose at which no response is *observable*. What if the response is too faint to be detected with existing scientific instruments (i.e., we lack sufficient **sensitivity** to measure it)? In such a case, there would be no real threshold dose, just a difficult-to-see response to a low dose.

How do we determine the relationship between dose and response? The answer requires a great deal of money, work, scientific experimentation, and time (a decade or more is not out of the question). Every new drug (and cosmetic and food additive, for that matter) undergoes a series of tests before it can be sold legally to the public. These tests are conducted hierarchically, meaning we start with relatively simple analyses, including whether the test compound adversely affects cells in petri dishes, and then move on to additional, more complex tests, if initial data warrant such a continuation. Typically, a candidate that appears promising after early stages will next undergo testing in laboratory animals like insects, rats, and mice. In these latter and more complicated tests, groups of organisms from the same species are each given different doses of the xenobiotic and the number of subjects affected is recorded (tests on larger animals, and ultimately humans, will ultimately follow if and when the compound is deemed safe enough for such a step). From these data, the relationship between dose and response within whole populations can be predicted.

Finally, we address one additional, possibly surprising, implication of the dose-response principle. Not only does dose determine whether a substance will serve as a therapy or poison (e.g., acetaminophen and selenium, as we noted earlier), *every xenobiotic* will be deadly at some dose. Potency clearly plays an important role in whether a compound is more likely to help or hurt, but even the seemingly friendliest and most familiar substances have the capacity to harm. Consider water. We all know it is fundamental to survival, but few realize it can kill if it reaches a high enough concentration in the body (yes, water is poisonous in the range of many liters consumed within a short time period). Toxicologists sum it up this way: **the dose makes the poison**.

Responses vary among species

Known as the **principle of selective toxicity**, this second concept is critical to an understanding and application of the science of toxicology.

Explanations and examples. Put simply, different **species** receiving the same dose of a particular poison will not respond in the same way. Acetaminophen, a compound we have discussed already, provides a vivid example of this phenomenon. Due to some differences in their biochemical properties, humans can tolerate much higher doses of this common drug than can certain house pets; in fact, a very low dose of acetaminophen is likely to kill dogs and cats. Many other xenobiotics and organisms demonstrate the principle, including **herbicides** (far more lethal to plants than animals), antibiotics (kill bacteria, not animals), preservatives (inhibit microorganisms that can spoil food, but, in principle, do not affect human consumers). In some cases, the reasons for selective toxicity are straight forward: an herbicide targets a cell component that is absent in animals, for example. At times, though, differences in response are surprising and harder to explain, say how a notorious drug like thalidomide causes birth defects in humans but not mice. We will learn more about thalidomide shortly.

Applications. The fact that humans respond differently to certain poisons than do other organisms comes in handy. Many of the substances we encountered above are useful because of their selective toxicity. Think about a drug like penicillin. We take it to treat some illnesses because it kills the responsible bacteria without appreciably affecting the health of most people (review Box 5.6 for more about antibiotics). Why is penicillin selectively toxic? Because it targets the construction of **cell walls**, structures that humans (and all animals) lack. To varying extents, weed killers, insecticides, fungicides, preservatives, and many others cause less harm to humans than they do to their target organisms. Of course, despite the benefits we derive from selective toxicity, it also complicates the drug-development process described earlier in this chapter because the responses of test organisms may or may not be similar to responses in humans. We will return to this problem in section 15.6.

It also is seen within a species. We would be remiss if we left out this final point about selective toxicity. In addition to differences between species (e.g., humans vs. cats vs. mice vs. bacteria vs. plants…), members of the *same* species also tend to respond differently to xenobiotics. Genetic variation among humans, for instance, leads to differences in storage, metabolism, and excretion (three protective strategies we will explore in the next section), in other words, intraspecies selective toxicity. Now you may better appreciate the presence of disclaimers that sound something like "individual results may vary" in drug advertisements.

15.4. CONSEQUENCES OF EXPOSURE: A DEEPER DIVE

15.4.1. Xenobiotics may harm organisms

As we noted above, a poison can disrupt one or more sensitive targets. Changes to affected systems may then lead to local and / or whole-organism consequences. We have already listed some examples of target systems and endpoints, but here we will take a closer look at some of particular concern to toxicologists.

Mutations and cancer

As we saw back in Chapter 6, changes to **DNA** can alter the physical traits of an organism. **Mutations** also can lead to the induction of abnormal cell growth and the development of rapidly expanding masses known as a tumors, that, left unchecked and untreated, will cause the death of the affected organism. Among the many causes of mutation is exposure to xenobiotics that specifically target DNA. The list of potential **carcinogens**, cancer-causing agents, is lengthy and includes components of cigarette smoke, certain pesticides, food contaminants and additives, asbestos, industrial products, and many others.

Birth defects

Prenatal exposure to xenobiotics may disrupt the growth of an offspring, despite the extra layers of protection provided by the placenta. **Teratogenic** compounds (those with the capacity to induce birth defects) can affect one or more crucial stages of development and bring on abnormalities ranging from minor to catastrophic. One of the most important historical teratogens was a drug called **thalidomide**. Used to combat morning sickness in the 1950s and 1960s in several countries, it killed high numbers of cells in a fetus and led to several thousand babies born with missing or severely shortened limbs (it was quickly taken off the market in the early 1960s).

Ethyl alcohol is a second noteworthy teratogen. It has been linked to teratogenesis for centuries and continues to pose a risk if consumed during pregnancy. Unlike thalidomide, it tends to induce subtle, cognitive deficits, although it may cause abnormal physical features as well. Cleft palate, for example, has been linked to alcohol consumption by pregnant women (Figure 15.7).

Figure 15.7. *Cleft palate may be induced by prenatal exposure to alcohol. US CDC, Public Domain.*

Nerve damage and paralysis

Several xenobiotics affect the nervous system. Certain pesticides, for example a widely used compound known as malathion, are used because of the way they paralyze target insects. Humans exposed to such compounds can also experience similar life-threatening consequences.

Immune system disruptions

Our immune system is a complex assemblage of cells and tissues that helps protect us from the enormous variety of potentially deadly agents (e.g., bacteria, viruses, venoms, and other toxins) that enter us from outside sources. Damage to it can reduce its effectiveness, an endpoint known as **immunosuppression**, or ramp up its activity to unhealthy levels: both **hypersensitivity** (unnecessary reactions to non-threatening agents, essentially allergy) and **autoimmunity** (inappropriate destruction of one's own cells) are possible[1]. As with the other endpoints, many xenobiotics cause immunotoxicity. For example, members of the class of compounds known as per- and polyfluoroalky substances (PFAS) found in clothing, cookware, carpets, and industrial chemicals, have been linked to immunosuppression (as well as cancer, birth defects, and liver problems)[2]. Polychlorinated biphenyls (PCBs), another important class of industrial compounds, also act as immunotoxins[3].

Endocrine disruption

The endocrine system uses hormones to control key functions such as reproduction, growth, and metabolism. Certain xenobiotics can induce a range of adverse consequences, including the development of unwanted masculine traits in females (and vice versa), infertility, stunted growth, and others. The list of potential endocrine-disrupting chemicals is long and diverse, and includes the PFASs and PCBs we just noted as well as some plasticizers (e.g., BPA and others found in food storage containers and toys), and naturally occurring

1. For more information, see: U.S. Food and Drug Administration. 1999. Immunotoxicity testing guidance. www.fda.gov

2. U.S. Department of Health and Human Services, Agency for Toxic Substances Disease Registry. 2022. Per- and Polyfluoroalkyl Substances (PFAS) and Your Health. atsdr.cdc.gov

3. U.S. Department of Health and Human Services, Agency for Toxic Substances Disease Registry. 2011. Toxic Substances Portal. atsdr.cdc.gov

phytoestrogens (i.e., estrogen-like chemicals produced by plants such as soybeans)[4].

Organ damage

Our final, broad category includes what its name suggests: disruptions to the normal function of the liver, kidneys, respiratory system, skin, cardiovascular and other critical organs / systems in the body. Many metals, solvents, cleaners, fertilizers, and others have the potential to adversely affect one or more organs[5].

15.4.2. The body can protect itself from harm

At this point you might be wondering: do I have any defenses against xenobiotics, or am I done for? In brief, our bodies have developed many strategies to minimize (if not eliminate) the damage done by xenobiotics.

Barriers to uptake

Remember that a poison can do no harm until and unless it enters an organism. It is likely unsurprising, then, that we possess barriers that reduce uptake of xenobiotics into our bodies and critical body

4. National Institute of Environmental Health Sciences. 2022. Endocrine disruptors. www.niehs.nih.gov/
5. U.S. Department of Health and Human Services, Agency for Toxic Substances Disease Registry. 2015. Which Organ Systems are Affected by Toxic Exposures. atsdr.cdc.gov

systems. Here are a few of the major protective structures that serve us.

Membranes. Each of the trillions of cells that make up our bodies is enclosed by a **cell membrane**. Due to their physical and chemical properties, these essential barriers can control much of what goes in and out of cells, including nutrients, waste products and xenobiotics. Keeping things general, only chemical compounds with certain properties can easily cross membranes.

Extra protection for important systems. All internal systems are shielded by cell membranes as noted above, but some of the most critical and sensitive contain extra layers of protection. For example, because of the importance of the brain to overall health and survival, as well as its sensitivity to disruption, penetration of xenobiotics into it is further impeded by several features collectively known as the **blood brain barrier**. A developing fetus is similarly susceptible to toxic substances so is protected by the **placental barrier**. Neither system is fool proof and xenobiotics certainty do breach them, but the additional restrictions to penetration reduce risk from life-threatening damage.

Skin. This outer portion of our bodies consists of many layers of cells (with their membranes, of course) and serves as a first line of defense against dermal absorption (recall routes of exposure, above). Skin can inhibit the uptake of some xenobiotics, but many important poisons do cross this barrier and rapidly enter the blood stream.

Storage sites

Recalling the term "target", we know that systems in our bodies are not equally vulnerable to the effects of a particular xenobiotic. In fact, some tissues are insensitive enough that they can bond to certain compounds without meaningful adverse consequences. Storage sites such as bones, fats, and proteins carried in blood play important roles in binding and therefore reducing the concentration (i.e., the dose) of poisons at target sites. Too good to be true? Alas, yes, a few caveats are in order. First, not all xenobiotics can be effectively stored. Second, less than 100% of a xenobiotic is immobilized, meaning some still circulates throughout the body. Third and critically, storage is not a permanent condition: various processes can lead to the freeing of sequestered poisons and harm to an affected organism.

Metabolism

Xenobiotics that are not stored may undergo chemical transformations by natural metabolic processes within an organism (several body sites participate, although the bulk of this type of metabolism occurs in the **liver**). Some of these reactions appreciably reduce the danger posed by a xenobiotic, like the ones famously responsible for changing ethyl alcohol into a less-harmful product (thereby reducing the intoxicating effects of wine, beer, and spirits). On the other hand, other reactions actually *increase* the toxicity of a xenobiotic! What's with this counter-intuitive maneuver? Well, some of the compounds entering an organism do so in forms that are rather difficult to excrete (see the next defense mechanism on this list). Since eliminating toxic substances is often the best protective strategy, the body goes through a lot of effort to change the chemistry of a xenobiotic such that it is easily dissolved in water and flushed out. Unfortunately, the products of these reactions may be more toxic than the **parent xenobiotic** (the original compound that entered the body) and can damage some important targets before they exit. In short, metabolism can increase *or* decrease the likelihood that a xenobiotic will cause harm.

Elimination

The products of metabolism noted immediately above can be released from the body through one of several processes. In terms of the quantity removed, excretion via **urination** and **defecation** are the two most important elimination pathways. Other mechanisms of elimination include sweating, movement into hair, vomiting, exhalation, and lactation. Although none of these minor pathways has an appreciable effect on the total amount of a xenobiotic in an adult body, they are not irrelevant; the concentration of poisons in breast milk may be high enough to affect a nursing infant, for example.

15.5. ENVIRONMENTAL TOXICOLOGY

Now we turn our attention to the important branch of toxicology concerned with the sources, fates, effects, and management of pollutants (i.e., poisons) in **ecosystems**.

15.5.1. Sources of pollutants

As we saw earlier in this chapter, many anthropogenic activities generate and release poisons into soil, air, and water. Fertilizers, pesticides, sewage, drugs, cleaning products, trash, metals, solvents, and many other physical and chemical pollutants can end up in natural systems. Both the number and variety of potentially dangerous substances environmental toxicologists address are enormous.

15.5.2. Fates of pollutants

Whatever its source, a toxic substance can experience one or more **fates**. That is, pollutants rarely stay put, unchanged, after they are released, instead, they undergo transformations and displacements (i.e., they are moved) that are critical to the health of organisms and ecosystems.

Degradation

This first fate is somewhat analogous to the metabolism of poisons seen in section 15.4, above: some pollutants may be chemically altered by the actions of microorganisms like **bacteria** and **fungi** or by non-living phenomena. Often, such reactions transform, or degrade, a relatively large and complex parent compound (using "parent" like we did in section 15.4) into smaller, simple products. Biological processes, or **biodegradation**, are particularly important in soils because they may reduce or even eliminate the threat from a pollutant. Whether reactions *will* serve to clean up problematic contaminants depends on properties such as temperature, light and oxygen availability, and water content, though. In other words, if environmental conditions are not conducive to the success of, say, the bacteria that can carry out the necessary reactions, an otherwise degradable pollutant might persist in its parent form indefinitely (see Box 10.2 for a discussion of similar problems in polluted ocean waters). Furthermore, biodegradation does not affect all synthetic chemical compounds, as some are simply inherently difficult for organisms to break apart, nor are all microorganisms able to metabolize relevant poisons. Keeping its limitations in mind, degradation can diminish the likelihood that a pollutant will experience any of the remaining fates on this list. Before addressing them, though, we should revisit another point from the earlier section on metabolism, namely, that chemical changes often decrease the threat from a pollutant, but *an increase in toxicity is also possible.*

Movement to the atmosphere

In some cases, liquid environmental pollutants change into gaseous forms. This process, known as **volatilization**, can move highly toxic substances out of soil and into air. Whether pollutants travel to the atmosphere is a function of chemical properties (i.e., some are more likely to volatilize than others) and environmental conditions (e.g., higher temperature increases volatilization).

Movement to reservoirs of the hydrologic cycle

We know that water moves across Earth's surface via **runoff** and downward to **groundwater** via **infiltration**. Those pollutants that can be dissolved in water will travel out of soil in those same directions, potentially contaminating important reservoirs of the hydrologic cycle (review Chapter 11, including the discussion of **leaching**). As you probably expect to hear by now, a number of factors influence the extent to which poisons flow with water. First, as noted, only those that are soluble in water are likely to move very far. Second, degradation and volatilization could reduce the amount of pollutant in soil and therefore the amount in water.

Uptake by organisms

Toxicologists use three terms to describe **uptake**, the entry of environmental pollutants into the biosphere.

Bioconcentration. In this case, a compound moves from non-living soil, water, or air into the cells or tissues of an organism via non-dietary means (i.e., not due to feeding). For example, an earthworm in soil or a fish in water may absorb pollutants through their skin. If you are in a smoke-filled room, you similarly bioconcentrate toxins through your lungs. As usual, the extent to which this process occurs depends on some variables.

Properties of the pollutant

First, pollutants with very low water solubility, for example the insecticide DDT and the industrial compounds PCBs, readily cross cell membranes and then get stored in the fats of organisms. Second, pollutants that are inherently stable—they resist degradation—are more likely to persist long enough to enter organisms (and stay there) than are those with less resistance; DDT and PCBs meet this criterion as well.

Properties of the organism

Although membranes are similar across the biosphere, differences in physiology (i.e., structure of an organism), behavior, and habitat will influence how much of a pollutant will enter an organism. For example, you might imagine that plants, with their root systems, interact differently with soil and the contaminants in it than, say, earthworms, moles, or fungi.

Biomagnification. This second term refers to the movement of a pollutant into an organism by way of diet—that is, it enters an organism through what it eats. So, a small fish can ingest a poison if it eats contaminated algae. That poison will move into a larger fish that eats that smaller algae-eating fish, and so forth up to higher and higher **trophic levels**. We also describe the movement of a pollutant from prey to predator as **trophic transfer**.

The factors affecting bioconcentration are similar to those affecting biomagnification. Chemistry of the poison is particularly important because only a compound that persists inside a prey organism is likely to be present by the time a predator arrives on the scene. Furthermore, a compound that does not dissolve in water will move preferentially into fats where it can be stored until eaten. Properties of the prey are also critical. If an organism can metabolize and eliminate a poison, biomagnification by a predator is likely to be minimal. Fat content in the body of a prey organism will also influence how much of the pollutant it can store. To say it is a complicated matrix would be a bit of an understatement!

Bioaccumulation. In practical terms, it can be difficult to distinguish dietary from non-dietary uptake. Consider, for example, the fish from the previous paragraph. Assuming they are swimming in contaminated water, the pollutant can enter their bodies through their skin (bioconcentration) and through the food they eat (biomagnification). The term bioaccumulation simplifies our lives as it refers to all pathways by which a pollutant ends up in an organism.

A word of clarification might be helpful at this point: the fates described here are not mutually exclusive. So, after a pollutant is released into soil, some of it can be degraded, some can move to the atmosphere, some to water, and some into organisms. Figure 15.8 presents a whimsical overview of the way a hypothetical soil pollutant is subject to many fates simultaneously.

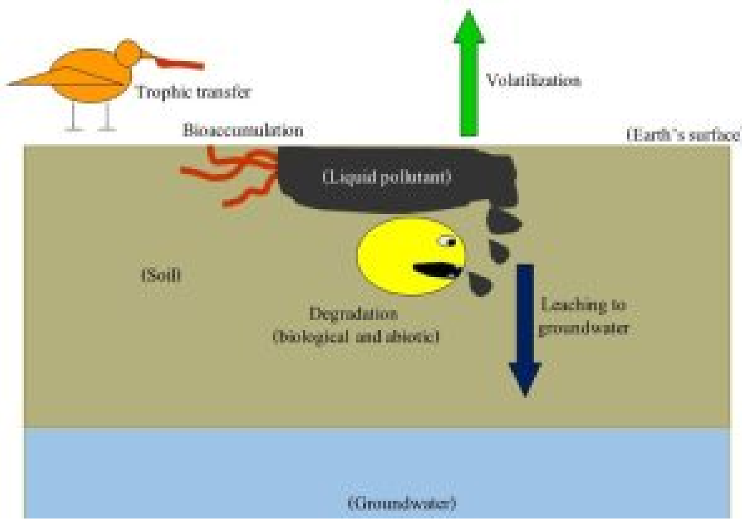


Figure 15.8. *Multiple fates can act simultaneously on a soil pollutant. In cross section, not to scale (e.g., bacteria, worm, and birds are not the same size; of course, real bacteria also lack teeth and eyes).* Kelsey, CC BY-NC-SA.

The relative importance of the possible fates, which can be represented in a **four-compartment model**, will depend on the chemistry of the compound as well as the properties of the environment into which it was dumped. Such a model visualizes each fate as a box, or compartment, in which some fraction of pollutant is contained. Note the distribution among the fates is likely to change with time as degradation and movement affect the amount and location of a pollutant in an area (Figure 15.9).

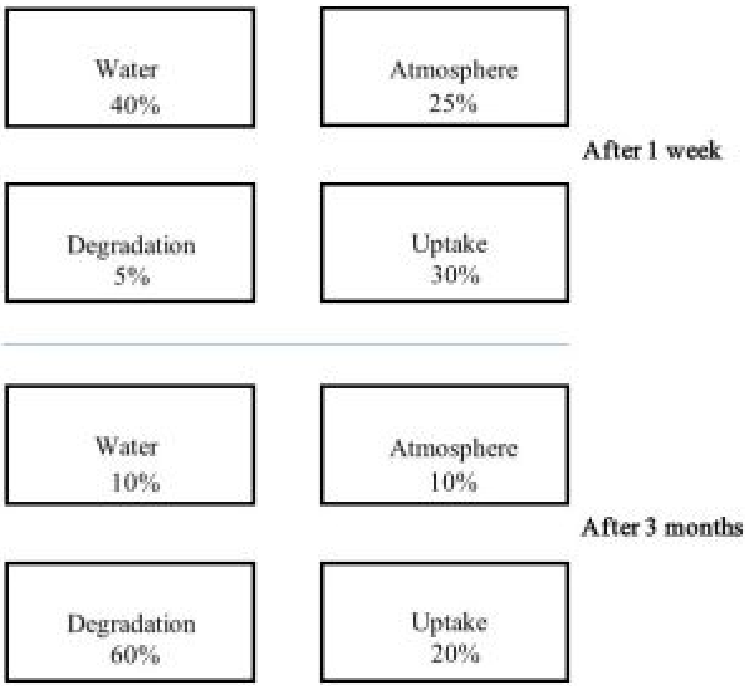


Figure 15.9. *The use of four-compartment model to show the relative importance of leaching, volatilization, degradation, and bioaccumulation one week and three months after a pollutant is released into soil. As suggested by the diagram, these four fates influence each other, and the distribution among them is likely to change with time. Note that the numbers shown are hypothetical and would be highly dependent on pollutant chemistry and environmental conditions. Kelsey,* CC BY-NC-SA.

15.5.3. Effects of pollutants

Finally, we reach the portion of the environmental toxicology section in which we answer a crucial question: who cares? Environmental releases of poisons can lead to multiple adverse outcomes, only some of which we will address here.

Water pollution

Contamination of water, a problem we have encountered several times before (e.g., **eutrophication**), should stand out to you as a major threat to the health of humans and natural ecosystems alike.

Changes to individuals and populations

Uptake of pollutants by organisms (through diet or otherwise) can bring on many of the consequences we saw in Section 15.4 (harm to individuals due to cancer, nerve damage, and so on). Bioaccumulation (described just above near the end of 15.5.2) is especially worrisome because it can lead to very high levels of poisons in the tissues of organisms at the highest trophic levels (i.e., **predators**), even if the original level of contamination in water or soil was low. How can this be? Remember from Chapter 5 (especially Figure 5.18) that the amount of **biomass** in an ecosystem declines precipitously from the lowest to highest trophic levels. If the total amount of a pollutant within that system does not change (which occurs if the compound in question resists metabolism), the compound will be concentrated in less and less tissue (Figure 15.10). See Box 15.2 to read about a rather famous example of the effects of pollutant bioaccumulation.

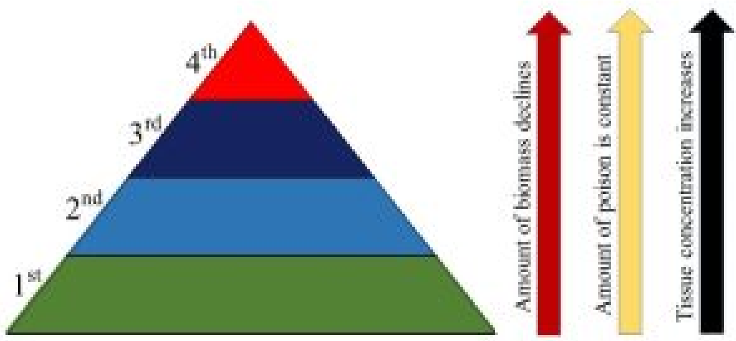


Figure 15.10. *A graphical model of bioaccumulation. The amount of biomass in a food web decreases with trophic level (in our hypothetical case, numbered 1st to 4th). If the total amount of a persistent pollutant within that web does not change, tissue concentration increases with trophic level. Compare to Figure 5.18 in Chapter 5.* Kelsey, CC BY-NC-SA.

Box 15.2. The power of bioaccumulation

The phenomenon illustrated in Figure 15.10 was responsible for dramatic and well-known declines in the numbers of predatory birds such as the bald eagle in the United States during the 1960s and 1970s. As we understand now, widespread application of the pesticide DDT in and near waterways was the culprit. Very briefly, the target of DDT, **malaria**-spreading mosquitoes, died and ended up as food for small fish. Since DDT resists metabolism (it undergoes very slight alteration to something known as DDE and then persists for decades) and is readily stored in fat, it was a great candidate for bioaccumulation. Thus, the eagles and other raptors that ate the largest fish in the aquatic system received a very high dose of DDE, an endocrine disruptor that specifically affects eggshell formation. Eaglets (baby eagles) inside faulty eggs did not survive, and population sizes plummeted. As noted in Chapter 1, DDT was banned due to concerns raised by scientists like Rachel Carsen as well as its effect on bald eagle populations.

Two additional points that are screaming to be made will close out this, our last box together.

1. In addition to serving as a vivid example of the potential of bioaccumulation to harm organisms, the mosquito-fish-bird story brilliantly demonstrates the principle of selective toxicity. Note the different responses among the species: immediate death to mosquitoes (due to CNS disruptions), no apparent effect in fish, even at high doses, and reproductive failure in predatory birds.

2. Here is yet another great demonstration of one of our favorite and fundamental concepts, the **principle of environmental unity**. Who knew that attempts to prevent the spread of malaria would very nearly lead to the extinction of bald eagles?

Change in ecosystem structure and function

Structure. Organisms that are well adapted to their environments will survive, reproduce, and thrive in their ecosystems (review Chapter 1 and Chapter 6). If conditions change due to the presence of a pollutant, though, a heretofore successful **population** may lose its advantage and dwindle in importance. A poison that diminishes efficiency of reproduction, resistance to disease, coordination and strength, or other vital adaptations, is likely to lead to a change in the identity of dominant species in an ecosystem. For example, the release of endocrine disruptors into aquatic ecosystems can alter the male-to-female ratio of certain amphibians and reduce their ability to reproduce effectively. Not only is that particular species affected—worldwide populations of many amphibian species are declining[6]—but organisms at trophic levels above and below that of the amphibian can be put at risk. Certain fish species have similarly been affected by endocrine-disrupting chemicals[7].

Function. Important functions such as levels of primary productivity are often measured as endpoints (review Chapter 5 for more about ecosystem functions). If pollutants affect populations of plants or algae, rates of carbon fixation will decline, meaning a loss of nutrients for **heterotrophs** as well as, potentially, increases in atmospheric **CO_2** concentrations (see Chapter 14, the portions of section 14.2 about global climate change, for why such increases matter).

Development of an ecological gradient

A quirky outcome of environmental pollution can be a pattern of organisms around or near a source of contamination. Due to

6. Hayes TB, Case P, Chui S, Chung D, Haeffele C, Haston K, Lee M, Mai VP, Marjuoa Y, Parker J, Tsui M. 2006. Pesticide mixtures, endocrine disruption, and amphibian declines: are we underestimating the impact? Environ Health Perspect. Apr;114 Suppl 1(Suppl 1):40-50. doi: 10.1289/ehp.8051. PMID: 16818245; PMCID: PMC1874187

7. U.S. Environmental Protection agency. 2022. Research on Detecting Endocrine Disrupting Chemicals from Animal Feeding Operations. epa.gov

selective toxicity, organisms that are most resistant to the adverse effects of the poison will dominate the high-dose areas, whereas organisms that are more sensitive will only appear in relatively clean areas. In terrestrial systems, a distinct gradient can be observed, often featuring **early stage successional species** like **lichen** and grasses (see Chapter 5 to review ecological succession) growing in the harshest conditions and later stage species like shrubs, small trees, and then large trees, appearing in sequence as a function of horizontal distance from the pollution source (Figure 15.11).

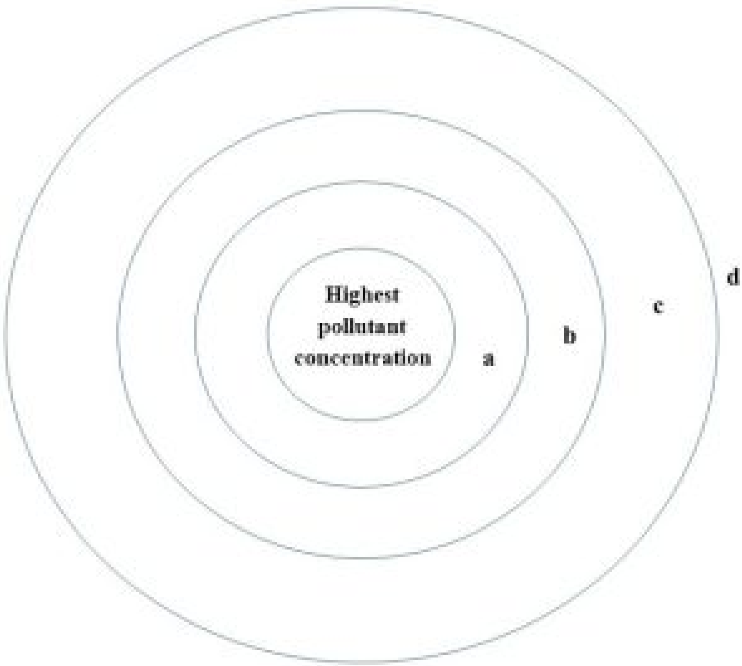


Figure 15.11. *A idealized diagram of an ecological gradient in map view. The pollutant concentration diminishes from the center outward. Lettered circles represent zones in which different organisms could be found in certain biomes: a, lichens, grasses; b, grasses and shrubs; c, shrubs and small trees (e.g., birch); d, large trees (e.g., maple and oak). The pattern at an actual site is likely to be less pronounced that the one suggested here.* Kelsey, CC BY-NC-SA.

Disruption of human food production

The success of many crops depends on the actions of pollinating insects (e.g., honeybees). Thus, a chemical pesticide that damages an important pollinator has the potential to adversely affect food supply.

Human health effects due to consumption of contaminated food

Humans are not isolated from the effects of environmental contamination. In addition to the risk of polluted water noted earlier, we also can consume high-level predators and suffer the effects of bioaccumulation. In one high-profile case, the release of PCBs into the Hudson River in New York State (USA) in the 1970s led to bans and restrictions on consumption of fish caught in contaminated waters because of potential health risks associated with this group of persistent chemicals[8].

15.5.4. Management of pollutants

It should be obvious by now that humans have good reasons to minimize releases of pollutants into natural systems. Many laws are in place to encourage proper handling of toxic substances and limit environmental contamination (see Chapter 1, section 1.2 for some examples). However, poisons still end up in soil, air, and water.

8. U.S. Environmental Protection agency. 2022. Hudson River Cleanup. epa.gov

Some of the strategies used to clean up, or **remediate**, polluted environments are briefly described below.

Bioremediation

Here, the activities of organisms are harnessed to remove poisons from soil or water.

Soil. In some cases, we manipulate environmental conditions to stimulate **indigenous** microorganisms possessing the ability to degrade a pollutant. Sometimes, the necessary bacteria or fungi are not present and so are added to a contaminated site. This latter strategy also often requires manipulations, including the addition of limiting nutrients or oxygen, for it to succeed. Another common approach, **phytoremediation**, takes advantage of the ability some plants have to extract pollutants from soil.

Water. Very often, contaminated water is removed from its reservoir and cleaned elsewhere. Groundwater, for instance, may be pumped to the surface and put into contact with bacteria able to degrade the pollutant in question. See Figure 15.12 for a diagram of such a pump-and-treat system.

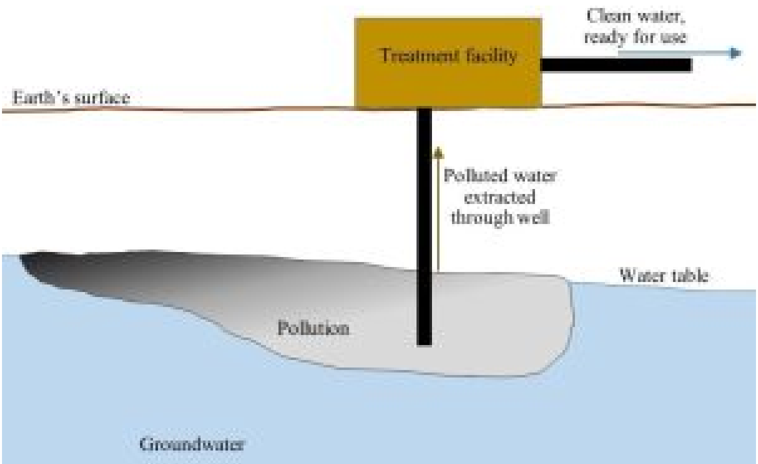


Figure 15.12. *A model showing a pump-and-treat system to clean up contaminated groundwater. Polluted water is moved to a treatment facility at the surface. In cross section. Kelsey,* CC BY-NC-SA.

Chemical remediation

Instead of, or to supplement, the use of organisms, one of several non-biological strategies may be used to treat pollutants in soil. Among the chemical approaches are some that enhance the movement of contaminants out of soil and others that attempt to immobilize pollutants within the soil matrix so they do not contact organisms or move into water.

Excavation

When biological and chemical approaches fail, a far more expensive and disruptive solution is employed: all the contaminated soil in a field is dug out, loaded onto trucks, and hauled away. The many tons of soil can be incinerated at high temperature or stored in a **landfill** designed to contain **hazardous waste**. As you might imagine, the party responsible for funding such a project will try to avoid excavation by whatever means, including litigation.

Isolation

This final strategy does not actually remove the pollutant, rather, a fence and other steps are taken to restrict access to a contaminated site. Clearly, we would expect such a move only to serve as a short-term solution (review Chapter 1, section 1.2 and Figure 1.2 for an example), one intended to protect human and natural systems until a proper clean-up strategy is designed. But fences have been known to stay up for years or decades.

15.6. RISK ASSESSMENT

Risk assessors work to determine the potential adverse effects of any hazard on human health. So, why do we end the toxicology chapter on this note? Because the science of risk assessment often makes use of data obtained by toxicologists.

15.6.1 Terms and concepts

Like so many other fields, risk assessment uses its own language. Here we define those ideas that are essential to our brief look at the subject.

Risk

Risk is the probability that an adverse outcome will result from contact with a substance or device or as the result of participation in an activity.

Hazard

This term has two related meanings. First, it is an intrinsic property of a substance, device, or activity. In short, hazard is the potential something has to do harm. A chemical like hydrochloric acid (HCl) is a good example of a hazard because it can damage your skin. Importantly, the inherent danger posed by hazards does not change. Risk, on the other hand, depends on exposure and can vary a great deal. Let's take another look at HCl to further distinguish between risk and hazard. If the acid is contained in a shatter-proof can that is stored behind a locked door that never gets opened, it poses no risk to you. Risk goes up only if your exposure to it goes up. Its hazard remains the same, though, whether or not you get near it as the acid always has the capacity to damage your skin. Second, the word "hazard" can refer to something that has an intrinsic, hazardous property and, thus, the potential to do you harm. This second meaning is extremely general, and includes everything from chemicals such as pesticides and drugs, to devices like artificial limbs and eye implants, to tools like step ladders and knives, to modes of transportation like cars and airplanes, to recreational activities like boating and amusement park rides, to athletic activities like skiing and baseball, to death-defying maneuvers like sky diving and bungie jumping. It is a very long list and quite a broad concept. Whatever the hazard, risk assessors calculate the probability of an adverse outcome as the result of exposure to it.

Safety

This is the inverse of risk, that is,

safety = 1 / risk.

If risk is the probability that you will be harmed from exposure to a hazard, safety is the probably that you will *not* be harmed by that same hazard. For instance, the *risk* of riding in a particular vehicle could be expressed as, say, 1 person out of every 20 ends up dead. Alternatively, that same ride could be said to be *safe* for 19 out of 20.

Adverse effect

Like response, an idea we encountered earlier in this chapter, adverse effect is defined as any undesirable / harmful outcome caused by exposure to a hazard. It is the endpoint measured by a risk assessor, the result of exposure to a hazard.

15.6.2. The process of risk assessment

Risk assessors seek answers to several questions as they analyze the likelihood that a hazard will cause harm.

Does the hazard cause the observed adverse effect?

Critically, we want to establish a mechanistic link between cause and effect. In the case of chemical compounds, data from humans exposed historically can provide some of the answer. Groups of people who worked with an industrial solvent for a decade or who smoked for 20 years, to provide just two of many possible examples, might be interviewed to connect observed health effects with past exposure. Often, tests using laboratory animals or individual cells also are used to help identify how a compound affects organisms.

What is the effect of dose on human response?

As above, human historical data as well as animal tests serve as good starting points here, but we will see in a moment that this question is not as easily—and reliably—answered as you might hope or expect.

What is the typical level of exposure to this hazard?

As we know, if the answer to this question is "none", there will be no risk. Most of the time, though, we are worried about xenobiotics to which exposure is likely.

What is a safe level of exposure to this hazard?

Once we get answers to the three questions posed above, we get to the last step, the **risk characterization**. For a chemical compound like a food additive, we would see a simple-looking answer that describes a safe dose in terms of mg of the xenobiotic per kg of body weight. For example, the answer to the question posed here could be something like:

a safe daily human dose is expected to be 1.0 mg / kg ± 0.2 mg.

Very important is an expression of the level of uncertainty in our risk characterization. Here, it is included as the error term "± 0.2 mg". Why is there so much uncertainty (20% in this case)? First, remember that all scientific measurements are expressed as a range due to **random** and **systematic** errors. But second, risk assessment adds its own sources of uncertainty that can produce rather large error terms. Critically, the principle of selective toxicity limits our ability to accurately predict individual human responses from animal dose-response data. To account for the differences in responses between and within species, risk characterizations tend to build in a lot of room for error, meaning that calculated safe human doses are commonly at least one hundred times lower than those seen in test organisms. The hypothetical recommendation of 1.0 mg / kg we saw above very likely would have been derived from an animal value of 100 mg / kg. The so-called safety factor, the amount by which the safe dose for animals is divided to establish a safe human dose (here it is 100), is used to better account for unavoidable uncertainty and reduce risk to humans.

The details of risk assessment of something other than a xenobiotic differ from the process shown here, but the guiding principles and goals are the same. For example, the likelihood of death due to riding in cars with and without seatbelts or the safety of a backyard trampoline could be calculated through analogous steps and analyses. We also can conduct **ecological risk assessments** to predict the potential adverse effects of a variety of human actions such as the construction of a new building, highway, or landfill on natural systems. Rather than endpoints related directly to human health, we would instead study loss of a **keystone species**, reduced **primary production**, or other ecological considerations described above.

15.6.3. Then what happens?

Results of studies conducted by risk assessors are provided to people who are at risk as well as those responsible for regulations on food, drugs, equipment, activities, and other hazards that could threaten human or environmental health. This latter group, known as **risk managers**, includes lawmakers who may accept the scientific data and craft legislation accordingly or ignore the recommendations due to multiple non-scientific factors including ideology, political expediency, and economic concerns.

THE CHAPTER ESSENCE IN BRIEF [9]

Myriad contemporary human activities produce potentially hazardous substances, although poisons have been a threat to organisms throughout history. Toxicology, the study of poisons and their possible effects, is founded on several important principles that influence the development, manufacturing, marketing, and use of drugs, pesticides, industrial chemicals, household cleaners and many other natural and synthetic materials.

---

9. As you will find throughout this book, here is very succinct summary of the major themes and conclusions of Chapter 15 distilled down to a few sentences and fit to be printed on a t-shirt or posted to social media.

Think about it some more...[10]

Is aspirin a poison or therapeutic drug? What about iron: harm or benefit? Ethanol? Cannabis?

How do antibiotics (Chapter 5) provide an example of selective toxicity?

Environmental scientists worry about the movement of poisons into the saturated zone. Briefly state WHY a compound that would otherwise break down in the unsaturated zone might persist in groundwater for an extended period nonetheless. As always, think like an environmental scientist as you ponder your answer (this sounds like another episode of connections between environments and organisms…).

Take another look at Figure 15.12. How do we know if the water exiting the treatment facility is clean enough? You might want to consider what we learned about water quality in Chapter 11 as you ponder your answer.

So, is everything toxic? Should you avoid contact with as

10. These questions should not be viewed as an exhaustive review of this chapter; rather, they are intended to provide a starting point for you to contemplate and synthesize some important ideas you have learned so far.

much stuff as possible? What should you do with the information provided by toxicologists and risk assessors?

Glossary

accumulation

In systems analysis, refers to an increase in the amount of material being measured. It results when rate of input of that material to a system exceeds rate of output from that system. See Chapter 2 for more. Accumulation can have another, distinct, meaning; see bioaccumulation for that usage (and Chapter 15).

accuracy

The correctness of a measurement in relation to a universally accepted standard. Contrast with precision and sensitivity. See Chapter 2 for details.

acellular

Biological entities that lack cells (e.g., viruses). See Chapter 3 for more.

acid mine drainage

An important environmental consequence of mining. Materials that were buried in low-oxygen environments are brought to the surface and exposed to oxygen and water. The resulting chemical reactions produce sulfuric acid that can flow to streams and other waterways. See Chapter 10 for more.

acid precipitation

Rain, snow, ice, etc., that has a pH lower than is natural. It can be caused by natural and human activity. See Chapter 10 for more.

acid precursor

A chemical substance released into the atmosphere that is converted into an acid. For example, sulfur dioxide gas emitted in coal combustion will react to form sulfuric acid. See Chapter 14 for more.

adapt

In response to environmental pressures, organisms develop strategies that allow them to survive and reproduce in a particular area. In ecology, a characteristic or trait that an organism develops in response to pressures placed on it by its environment is called an 'adaptation'. Well-adapted organisms posses the traits that allow them to be present in their current environment. Organisms with the best adaptations will be most successful in their environment (barring a random event that could disrupt them). Some simple examples of adaptations: fish have gills to extract oxygen from water, thus they can survive in water; cheetahs are very fast, thus they can catch prey in open landscapes; certain bacteria can withstand very acidic conditions, thus they can live inside the human stomach. See Chapters 1, 5, and 6 for more.

adaptation

In ecology, a characteristic or trait that an organism develops in response to pressures placed on it by its environment. Well-adapted organisms posses the traits that allow them to be present in their current environment. Organisms with the best adaptations will be most successful in their environment (barring a random event that could disrupt them). Some simple examples of adaptations: fish have gills to extract oxygen from water, thus they can survive in water; cheetahs are very fast, thus they can catch prey in open landscapes; certain bacteria can withstand very acidic conditions, thus they can live inside the human stomach. See Chapters 1, 5, and 6 for more.

adaptive radiation

A reason several species resemble each other physically. In this case, organisms from a single population took on somewhat different forms as they specialized within a small space. They look similar and are closely related genetically, but they are separate species occupying separate niches. Compare to convergent evolution and divergent evolution. See Chapter 6 for more.

additive

A substance intentionally put into food. Compare to contaminant. See Chapter 15 for more.

advantage

In ecology, refers to a trait that helps an organism survive and reproduce better than competitors. See Chapters 1, 5, and 6 for more.

aerobes

More formally, 'aerobic organisms', they require dioxygen gas for survival. See Chapter 1, especially Box 1.1, for more.

aerobic

An adjective referring to an organism that uses dioxygen in its metabolism. "Aerobe" is the noun, the organism itself. Contrast with anaerobic. See Chapter 1 for more.

aerosol

Aerosol is a general term referring to any fine particle or droplet suspended in air.

afforestation

The increase in forested area compared to the current situation. In other words, forests would encroach upon un-forested areas. See Chapter 12 for details.

agriculture

The human practice of cultivating crops and growing animals to produce food for human consumption. It involves extensive manipulations of natural environments and ecological processes. See Chapter 9 for details.

agroecosystem

A term used by environmental scientists to refer to a farm. It reflects the nature of agriculture: humans harness natural ecological processes to artificially produce more food than would otherwise grow in an area. See Chapter 9 for more.

albedo

The amount of solar radiation reflected off any planetary body. See Chapters 4 and 14 for details.

algae

A group of organisms that are eukaryotic, phototrophic, and autotrophic. Some are unicellular, but most are colonial. Note: some scientists include certain prokaryotic cells in this group, the so-called "blue-green algae"; others classify those prokaryotic organisms as bacteria (known as "cyanobacteria"). We will follow the latter convention in this book. See Chapter 3 for more.

alternative energy sources

A term often used to describe renewable energy sources such as solar, wind, and hydro power. See Chapter 10 for more.

amino acid

A chemical form of N that is biologically available to the biosphere. It is produced by fixation of nitrate and ammonium by plants, algae, and certain microorganisms. Amino acids are the building blocks of proteins. See Chapter 4 for more.

ammonium

An inorganic form of nitrogen with the formula $NH4+$. It is biologically available to plants, algae, and certain microorganisms. It is the product of the fixation of dinitrogen and can be further fixed to amino acids. See Chapter 4 for more.

anaerobes

More formally, 'anaerobic organisms.' They do not require dioxygen gas; in fact, many anaerobes are poisoned by $O2$. See Chapter 1, especially Box 1.1, for more.

anaerobic

An adjective referring to organisms that do not use dioxygen as part of their metabolism. Many such organisms die in the presence of oxygen gas. "Anaerobe" is the noun. Compare to aerobic. See Chapter 1 for more.

anoxic

Refers to an area that lacks dioxygen gas. Anaerobic organisms would be found in an anoxic environment. Compare to oxic and aerobic.

anthropogenic

An adjective that means of human origin. See Chapter 7 for details.

antibiotic-resistant bacteria

A population of bacteria that is no longer susceptible to the toxic effects of a particular antibiotic. Note it was once dominated by pesticide-sensitive individuals, but rapid evolution led to its ability to survive in the presence of the poison. See Chapters 5 and 6 for details.

antibiotics

Generally, a kind of chemical substance that kills living cells. In practice, 'antibiotic' refers to a drug designed to kill harmful bacteria without appreciably harming the host taking the drug. See Chapters 1 and 5 for more.

apex predator

In ecology, an organism that sits at the highest trophic level. In principle, no organism feeds on it. See Chapter 5 for details.

aquatic

An adjective that refers to organisms that live in water. Compare to terrestrial. See Chapter 4 for more.

aquifer

A geologic structure that can store and transmit usable quantities of water. See Chapters 4 and 11 for more.

Aral Sea

An inland freshwater sea in Central Asia. It is an example of how competing demands for water can lead to adverse consequences. In this case, diversions of the rivers that fed the Aral in the interest of agriculture led to profound shrinking of this formerly enormous body of water. See Chapter 11, especially Box 11.2, for more about conflicts over water.

Archaea, domain

One of the three domains of living things. It is made up of prokaryotes that tend to live in relatively extreme environments (e.g., temperature, acidity, pressure). Compare to Bacteria and Eukarya. See Chapter 6 for more.

artificial wetland

Constructed by humans to treat sewage. See Chapter 11 for more.

asexual reproduction

A type of reproduction in which a individual can reproduce without a mate. Bacteria and protozoa, for example, reproduce this way: a single cell divides into two identical cells. Contrast this strategy with sexual reproduction. See Chapter 6 for more.

assimilation

When an organism takes up a nutrient and incorporates it into its cells. See Chapter 4 for more.

asthenosphere

The thick layer of the Earth that lies beneath the lithosphere. It consists of much of the middle and lower mantle and behave plastically when stressed. See Chapter 3 for more.

atmosphere

One of Earth's four spheres, this is the envelope of gases that surrounds the Earth. See Chapter 4 for more.

atmospheric inversion

Also known as a thermal inversion. This is a natural phenomenon whereby cool air is trapped near Earth's surface by layers of warmer air. Inversions can exacerbate local air pollution problems like smog because pollutants released at the surface do not disperse. See Chapter 14 for more.

atom

A chemistry term, atoms are the building blocks of everything in the universe. They are the smallest units into which an element can be divided. Atoms consist of even smaller, subatomic particles, such as protons, electrons, and neutrons. See Chapter 4 for more.

autoimmunity

In toxicology, an endpoint caused by exposure to certain xenobiotics: the immune system inappropriately attacks the cells of its own body. See Chapter 15 for details.

autotroph

A type of organism that fix its own carbon, that is, it can convert unavailable carbon dioxide into useable glucose to meet its nutrient needs. Includes plants, algae, and some bacteria. Autotrophic an adjective. Compare to heterotroph. See Chapter 5 for more.

average life expectancy

In demography, the mean length of life, in years, of a population. See Chapter 8 for more.

background (as in background levels)

An important consideration when scientists make measurements. Background levels refer to the normal condition of whatever variable is under consideration. For example, normal human body temperature is 37 degrees C (98.6 F). Put another way, background is 37 (i.e., not 0); we use a thermometer to determine if temperature is ABOVE background (say 39) when a person is sick. Background levels can affect measurements and interpretations of experimental results. This idea is specifically referenced in Chapter 14 but is relevant for all scientific measurements and ideas described in the book.

bacteria

Refers to the many different kinds of microscopic organisms that are single celled and prokaryotic (i.e., do not possess distinct cellular organelles). Collectively, bacteria are extremely diverse, are the oldest organisms on Earth, and are critical to many environmental cycles and processes. See Chapter 3 for more.

Bacterial, domain

With a Capital B, refers to one of the three domains of living things. It is made up of prokaryotes that tend to be found in relatively moderate environments. Compare to Archaea and Eukarya. See Chapter 6 for more.

Bhopal

The site of a major accident at a Indian chemical manufacturing plant in 1984. Thousands of people were killed and hundreds of thousands were sickened. This event raised many questions about corporate responsibility and environmental contamination. See Chapter 1 for more.

bioaccumulation

Dietary and non-dietary uptake of pollutants by organisms; can be thought of as bioconcentration + biomagnification. See Chapter 15 for more.

bioavailability or biological availability

Refers to how usable an atom is. For example, some C atoms are not directly available for use (as in carbon dioxide) whereas others are directly available for use (as in glucose). See Chapter 4 for details.

biochemical oxygen demand (BOD)

An important water quality parameter. In simple terms, BOD is a measurement of the amount of aerobic metabolism occurring in a water sample, that is, the level of decomposition of organic materials by microorganisms. High BOD will drive down levels of dissolved oxygen in the water and, in some cases, lead to the deaths of fish and other aerobic animals. BOD can be thought of as an indirect way to measure water quality: when there is a lot of organic matter in the water, a lot of oxygen will be required by decomposing organisms to digest that organic matter. See Chapter 11 for more.

bioconcentration

Uptake of a substance (generally, some kind of pollutant) from a non-living medium like soil or water into an organism. See Chapter 15 for more.

biodegradation

A process analogous to decomposition of natural compounds, but this term is generally used to refer to chemical decomposition of synthetic products by organisms; fungi and bacteria are responsible for most of the relevant biodegradation seen in contaminated environments. See Chapter 15 for more.

biodiversity

A broad term that refers to the number of different organisms in an ecosystem or, often, the whole Earth. There are several specific ways to measure and express biodiversity. See Chapters 5 and 6 for details.

biogas

Methane-containing gases produced by anaerobic decomposition in landfills, sewage treatment plants, and farms. It can be burned as a fuel. See Chapter 10 for more.

biological evolution, theory of

Holds that populations change with time in response to a number of factors, including mutations, natural selection, and random chance. Biological evolution is the main scientific explanation for the development of life on Earth. See Chapter 6 for many more details.

biological variability

Differences in metabolism, physiology, behavior, etc., among different organisms. It adds uncertainty to scientific measurements. See Chapter 2 for details on uncertainty.

biomagnification

Uptake of a compound (generally, some kind of pollutant) into an organism via diet. See Chapter 15 for more.

biomass

A term referring to the total amount of living material in a place.

biome

In ecology, a type of ecosystem. Depending upon who you ask, Earth has between 5 and 15 biomes, and each is characterized and classified according to dominant organisms present and environmental conditions. See Chapter 5 for details.

biomedical waste

Produced by medical facilities like hospitals, this includes body fluids and tissues from humans. Due to its potential to spread disease, it is handled more carefully than is municipal solid waste. See Chapter 13 for more.

bioremediation

A general term that refers to the use of organisms to clean up polluted water or soil. See Chapter 15 for more.

biosphere

One of Earth's four spheres, refers to all living things. See Chapter 4 for more.

blood brain barrier

A set of physiological features that provides extra protection for the central nervous system against penetration by xenobiotics. It is not 100% effective, but it does prevent the entry of many potential poisons into the CNS. See Chapter 15 for details.

body burden

The total dose of a xenobiotic in a whole organism. See Chapter 15 for details.

bond

In chemistry, one of several attractive forces that hold atoms together. See Chapter 4 for more.

buffering capacity

A chemistry term related to acids and bases. It refers to how resistant a water-based solution is to pH changes following addition of strong acid or base. High buffering capacity indicates a solution that will resist changes--that is, its pH does not change appreciably--after acid or base is added. See Chapters 9 and 10 for more, including context and relevance.

calibration

A procedure to assess the accuracy of a scientific instrument. Accomplished by measuring a standard of known properties. For example, a balance could be calibrated through the use of a calibration weight, a carefully constructed object of known mass. A calibration weight with a mass = 1.00000 g could be used to assess the accuracy of a laboratory balance; if the mass of the object is determined to be 0.99880, we know the balance needs to be calibrated. Put informally, we need to remind the balance what 1.00000 g actually looks like! See also accuracy and Box 2.2.

canopy

In ecology, refers to the nearly continuous layer of leaves and branches in a mature forest that blocks much of the incoming sunlight that would otherwise strike the forest floor. See Chapter 5 for details.

cap and trade

A strategy used to reduce use and release of pollutants by a group of emitters. A limit on the total amount of the targeted substance is established and lowered each year. Individual polluters (i.e., companies) need to buy allowances to continue to emit the pollutant or may change their processes to eliminate missions all together. See Chapter 14 for more.

carbohydrates

A group of chemical compounds that include sugars like glucose as well as polymers like starch and cellulose. See Chapter 4 for more.

carbon

One of the known elements in the Universe, carbon (C) atoms consist of 6 protons and 6 electrons. Carbon is an important element for living things; organic C compounds make up the skeleton of biological systems. See Chapter 4 for details.

carbon capture and storage

A proposed strategy to reduce the amount of carbon dioxide moving to the atmosphere. In principle, the CO2 gas would be collected and moved to underground rock layers. See Chapter 14 for more.

carbon capture and utilization

A strategy that would collect carbon dioxide gas generated from fossil fuel combustion before it moves to the atmosphere. The captured carbon would then be used in the manufacturing of plastics. It is an idea that is in early stages of development. See Chapter 14 for more.

carbon dioxide or CO2

A gas made up of one carbon atom and two oxygen atoms, it is one of the many forms carbon can take. This gas is the product of aerobic respiration as well as combustion of fossil fuels, wood, and other materials containing organic carbon. It is not directly usable as a carbon source but can be fixed into useable forms of carbon, for example glucose, by autotrophic organisms. See Chapters 4 and 14 for more.

carbon fixation

The process of converting C from non-biologically available forms to forms that can be used by organisms. Typically, carbon dioxide molecules are converted to glucose. C fixation is carried out by organisms that possess the proper machinery to do so:

plants, algae, and some bacteria. Fixation is a general term that indicates an increase in the biological availability of any atom. See Chapter 4 for more details.

carcinogen

In toxicology, a xenobiotic that causes cancer. See Chapter 15 for details.

carnivore

An organism that eats the flesh of consumers. That is, carnivores usually do not eat primary producers. See Chapter 5 for more.

carrying capacity

Equal to the number of individuals of a species that can be sustainably supported by an ecosystem. See Chapter 1 for more.

Carson, Rachel (1907-1964)

Environmental scientist and activist, author of the book *Silent Spring* (1962). She was an important voice in the environmental movement and helped raise awareness about the potential dangers of indiscriminate use of pesticides like DDT. See Chapter 1 for more.

catalyst

A broad term referring to a substance that accelerates a chemical reaction without being affected by that chemical reaction.

catastrophism

An approach to studying Earth's history characterized by the belief that Earth was shaped by fast, dramatic upheavals, not slow processes. It is used to explain how a young Earth could contain so many seemingly ancient features. Importantly, no objective evidence supports catastrophism. Contrast with uniformitarianism. See Chapter 3 for more.

causation

Two events are linked mechanistically; one is influenced by the other. Compare to correlation. See Chapter 2 for more.

cell

The basic microscopic unit that makes up living things. Cells are systems made up of component parts and receive inputs and produce outputs. Different organisms are built of different types of cells, but cells have many things in common. All are enclosed by cell membranes, some have an extra rigid layer outside the membrane known as a cell wall. Some organisms consist of just one cell (e.g., bacteria) whereas others are made up of billions or trillions (e.g., animals and plants). See Chapter 3 for more.

cell membrane

The semi-permeable boundary that encloses all cells. Its chemical properties give it some control over what enters and exits a cell. See Chapter 3, especially Figure 3.9, for details.

cell wall

The rigid layer found outside the cell membrane in the cells of plants and many microorganisms (but not animals). See Chapter 3 for more.

central nervous system, CNS

Refers to the brain and nerves in the body. See Chapter 15 for details.

chain reaction

A process whereby the fission of a single unstable atom can lead to subsequent fission of more and more unstable atoms. See Chapter 10 for more.

Chernobyl

The site of a major nuclear accident in 1986. Located in Ukraine, this power plant exploded due to multiple user and design errors. The environmental effects of this accident persist to this day. See Chapter 1 for more.

chlorofluorocarbons (CFCs)

A class of chemical compounds built of C atoms bonded to some number of Cl and F atoms. CFCs are important greenhouse gases and also contribute to ozone depletion because they travel to the stratosphere and release Cl atoms; those Cl atoms are highly reactive and catalyze ozone destruction. See Chapter 14, including Figure 14.31, for more.

Clean Air Act

A law first passed in the U.S. in 1970, it was intended to reduce air pollution. Several subsequent amendments have strengthened it over the years. See Chapters 1 and 13 for more.

clear cutting

When all the trees in an area are cut down at one time. See Chapters 7 and 12 for more.

climate

The long-term, average atmospheric conditions that prevail in an area. Temperature (ranges and averages) and precipitation patterns are two important climatic characteristics. Compare to weather. See Chapter 14 for more.

climax community

In ecology, the group of species in ecological succession that are not replaced as a matter of course. See Chapter 5 for details.

CNS

See central nervous system.

coal

An important fossil fuel, coal is a term that refers to several types of soil rocks that were derived from ancient plants. See Chapter 10 for more.

coal seam

Where coal lies within the Earth. It is essentially a series of layers of coal that can be accessed through mines and then extracted. See Chapter 10 for more.

coal, grade

A way to categorize coal by sulfur content: high emits less sulfur upon burning than does does low. See Chapter 10 for more.

coal, rank

There are several different types of coal arranged according to their age and energy content, that is, rank. See Chapter 10 for more.

colonial

Refers to organisms that are built of multiple cells, but those cells do not specialize and communicate with each other to the same extent as cells in true multicellular organisms. Compare to unicellular and multicellular. See Chapter 3 for more.

colonizers

In ecology, organisms that can live on barren landscapes and begin the process of ecological succession. They are pioneer species. See Chapter 5 for more.

community

In ecology, a group of interacting populations. See Chapter 5 for details.

competition

An interaction among organisms characterized by rivalry for limited resources. All participants are harmed as a result. See Chapter 5 for details.

competitive exclusion principle

In ecology, states that two organisms with the same requirements cannot coexist in the same space and time indefinitely. Eventually, the organism that is best adapted to survive will dominate and force the weaker competitor out. See Chapter 5 for details.

composting

A practice that involves the accelerated decomposition of non-animal waste such as plant scraps, leaves, and related objects. Materials are collected in a structure known as a composter in which environmental conditions are maintained to stimulate aerobic decomposition. Compost, the product, is an organic- and nutrient-rich material that can be used as fertilizer. See Chapters 9 and 13 for more.

compound

In chemistry, a molecule that is built of more than one type of atom. Two oxygen atoms make up the molecule dioxygen, a pairing that is not a compound, whereas two hydrogen atoms bonded to one oxygen make up the molecule water, a group that is also called a compound. See Chapter 4 for more.

conclusion

A scientific notion derived from a limited amount of data. Conclusions are likely to be tested and profoundly modified with repeated experimentation
Compare to theory and law. See Chapter 2 for more.

condensation

A pathway of the hydrologic cycle. Water changes from gas to liquid. See Chapter 4 for more.

cone of depression

In the context of groundwater extraction by humans, the shape the water table takes around a well once pumping from it commences. See Chapter 11, especially Figure 11.5, for more.

conservation

In ecology, refers to an approach to wildlife and ecosystem protection. The goal of conservation is to manage wild living resources so they persist indefinitely and provide humans with specific benefits. Compare to preservation. See Chapter 6 for more.

consumption

A process whereby an organism ingests nutrients through diet. An organism that carries it out is a consumer. See Chapters 4 and 5 for details.

consumptive (water use)

A type of off-stream water use that does not immediately return the water to the reservoir from which it is removed. For example: human consumption. See Chapter 11 for more.

containment, in waste management

Refers to any one of several structures and strategies that restrict the movement of waste from its management facility (e.g., landfill) to water, air, or organisms. See Chapter 13 for more.

contamination/contaminant

Broadly, substances present in food, air, water, or soil that that do not belong where they are found (i.e., are out of place, should NOT be present). Compare with 'pollution', a word with a related, if distinct, meaning. In short, a contaminant can be thought of as any substance that does not belong where it is found (i.e., a metal dissolved in water). A pollutant is a contaminant that has reached a high enough concentration to bring on adverse effects (i.e., that metal noted in the previous sentence is at such a high level that it is killing fish; see Chapter 11 for more). Note that the terms are often used interchangeably by scientists and members of the public alike.

In the context of toxicology (Chapter 15) a contaminant is a substance that is inadvertently put into food. Compare to additive, something that is knowingly and intentionally added to food.

continental drift

In geology, the theory that Earth's continents have not always been in their current locations; rather, due to plate movements, continents have slowly changed position through history. See Chapter 3 for more.

control

Controls help establish causation. If controls and test subjects respond the same way, the factor tested likely is not responsible for the reactions noted. Essential, a control is deliberately left alone--it goes untreated--so a researcher can compare its reaction to that of subjects that ARE treated. See Chapter 2 for more.

control rods

Consisting of neutron-absorbing material, these are used to keep a nuclear reaction running at a more or less constant rate; they also can be lowered into a reactor to slow it in an emergency. See Chapter 10 for more.

control subjects

Controls or control subjects help establish causation. These are subjects that are untreated in an experiment; in other words, they provide the baseline against which test subjects are compared. Controls are essential to experimental science. If controls and test subjects respond the same way, the factor tested likely is not responsible for the reactions noted. See Chapter 2 for more.

convection cells

In geology, circular motion in the mantle that is thought to propel lithospheric plates. They are caused by the rising and falling of hot material. See Chapter 3 for more.

conventional oil drilling

The most widely used technique to recover petroleum from beneath Earth's surface. It involves drilling into reservoir rocks and then pumping the fluids to the surface. Compare to unconventional oil drilling. See Chapter 10 for more.

convergent boundary

In geology, a location where two lithospheric plates move toward each other or collide. See Chapter 3 for more.

convergent evolution

One reason members of different species can resemble each other physically: organisms that were not related genetically lived in the same environment and therefore developed the same physical adaptations. Compare with divergent evolution. See Chapter 6 for more.

core

A geology term referring to the innermost region of the Earth. It is partially solid and partially molten and made up of relatively dense elements. See Chapter 3 for more.

correlation

Two phenomena may occur at the same time but they may or may not have any influence over each other. Compare to causation. See Chapter 2 for more.

cost-benefit analysis

An evaluation of the advisability of a course of action that weighs the adverse consequences--the costs--of it against its potential advantages--the benefits. Inherent to a cost-benefit analysis is the assumption that any choice involves risks of some kind, and careful collection and interpretation of data can help us make a decision that poses the least risk. See Chapter 1 for more.

creep

A type of mass wasting event during which materials on Earth's surface move downhill extremely slowly. The movement is not perceptible in real time, but certain clues at the surface can indicate it is occurring. See Chapter 7 for more.

critical limits of tolerance

Both too little and too much of an essential nutrient can harm natural systems. See Chapters 4 and 15 for details.

crop yield

In agriculture, the amount of food produced from a field of a particular crop. See Chapter 9 for more.

crude oil

The raw material extracted from the Earth's crust by conventional or unconventional means. Except in very rare cases, crude is not immediately usable as an energy source; rather, it must be processed at a refinery to produce multiple fuels from it. See Chapter 10 for more.

crust

A geology term referring to the thin, brittle outer layer of the Earth. It is composed of relatively low-density elements. See Chapter 3 for more.

DDT

Dichlorodiphenyltrichloroethane, 1,1′-(2,2,2-trichloroethylidene) bis(4-chlorobenzene), or DDT is a pesticide designed to kill insects (i.e., it is an insecticide). It was used widely in the 1940s, 1950s, and 1960s but banned in the U.S. and largely phased out of use in the early 1970s. See Chapters 1 and 15 for more.

dead organic matter

Remains and waste products of organisms are grouped together as dead organic matter. It contains a lot of stored nutrients that can be used and recycled by decomposers. See Chapters 4 and 5 for more.

dead zone

Refers to a low-oxygen region of a body of water such as the Gulf of Mexico. Runoff from human activities such as agriculture introduces excess nutrients into the water; fish and other oxygen-requiring organisms die off in such a zone. See Chapters 1 and 4 for more.

decomposition

A general word that refers to the breaking down of relatively large, complex materials into relatively small, simple materials. Certain organisms, known as decomposers, break down the remains and waste products of other organisms and use those products as nutrient sources. Decomposers are critical to the health of ecosystems because they aid in the recycling of materials. Fungi and some bacteria are important decomposers. See Chapters 4 and 5 for more.

deductive reasoning

A strategy used by scientists in which specific cases are compared to previously developed and accepted principles. It is often used to categorize organisms or objects into preexisting groups. See Chapter 2 for more.

deep-well injection

A waste-management strategy whereby liquid waste is pumped underground into geologic layers deemed stable and appropriate for such materials. Site choice must be made carefully and waste must be stored below groundwater to minimize adverse consequences of this practice. See Chapter 13 for more.

deforestation

The removal of forest ecosystems. Multiple forces, anthropogenic and natural, can cause deforestation. See Chapter 12 for details.

degradation

A general term that refers to the breakdown of a relatively large, complex compound into smaller, simpler products. It involves the breaking of chemical bonds by biological or non-biological means. See Chapter 15 for details.

degradation of energy

Energy is converted from high to low grade forms, meaning it becomes less capable of doing work. See Chapter 4 for more.

demographic shift

Refers to a change in the relative size of different age cohorts within a population. For example, many human populations are experiencing a decline in the size of younger cohorts relative to the size of older cohorts. See Chapter 8 for more.

demography

The science of births, deaths, and other forces that affect population size. See Chapter 8 for more.

denitrifying bacteria

Bacteria that convert ammonia and/or nitrate to dinitrogen. In other words, they reduce the biological availability of N. See Chapter 4 for more.

density

Density is a term that essentially refers to how much solid material is packed into a given space, or it can be thought of as the relative amount of air space present in a particular substance. For example, a 1-liter bottle of foam packing peanuts will have a lower density than a 1-liter bottle of sugar which in turn will have a lower density than a 1-liter bottle of copper pennies. See Chapter 3 for more.

depletion

In systems analysis, refers to a decline in the amount of material being measured. It results when rate of output of that material from a system exceeds rate of input to that system. See Chapter 2 for more.

deposition

In geology, refers to processes that lay sediments onto the surface of the Earth. Typically, water or wind carries sediments until the velocity of the moving medium is insufficient to continue transport. See Chapter 3 for more.

dermal

An adjective that refers to skin. In toxicology, dermal exposure occurs when a xenobiotic comes into contact with skin. See Chapter 15 for more.

dermal absorption

In toxicology, movement of a xenobiotic into an organism through the skin. See Chapter 15 for details.

desalination

The removal of salt (by one of several methods) from water. See Chapter 11 for more.

detritivore

Small organisms that feed on the decaying remains of other organisms. See Chapter 5 for details.

developed world

Refers to nations that have relatively large amounts of technology, wealth, and infrastructure compared to the developing world. See Chapter 8 for details.

developing world

Refers to nations that have relatively small amounts of technology, wealth, and infrastructure compared to the developed world. See Chapter 8 for details.

diffusion

The movement of a material that is unequally spread throughout some medium (e.g., air or water) from areas of high concentration of the material to areas of low concentration of the material. See Chapter 4 for details.

dilution

The process by which a small amount of a relatively concentrated material is spread out within a large body or water or air.

dinitrogen or N2 gas

Two N atoms bonded together via a triple bond. It makes up nearly 80% of the gases in the atmosphere. Not biologically available, must be fixed before organisms can the N atoms in it. See Chapter 4 for more.

dioxygen gas

Two oxygen atoms are bonded to form O2, dioxygen. This is the gas required by aerobic organisms, including humans; it makes up approximately 20% of the gasses in the atmosphere. See Chapter 1, especially Box 1.1, and Chapters 3 and 4 for more.

disinfection

A broad term referring to a process that inhibits and/or kills unwanted microorganisms in water, food, or on surfaces. In the context of sewage treatment, it tends to be the last phase, one which reduces the number of pathogenic microorganisms in treated sewage before it is released into a natural water reservoir. See Chapter 11 for more.

divergent boundary

In geology, a location where two lithospheric plates move away from each other. See Chapter 3 for more.

divergent evolution

One reason different species resemble each other physically. In this case, individuals from the same species were separated from each other and moved into different environments. Since they developed in different environments, they developed different adaptations as time went by. Modern organisms that diverged into different species are still closely related genetically. Compare to convergent evolution. See Chapter 6 for more.

DNA

A biochemical molecule containing the code that makes every organism unique. See Chapter 6 for details.

domain

A term used to refer to one of the three major categories of living things based on genetic relatedness. See Chapter 6 for more.

dose

The concentration of a xenobiotic inside an organism, usually expressed as mg of compound per kg of body mass. See Chapter 15 for details.

drawdown

In the context of groundwater usage, the depth to which the water table is lowered by pumping. See Chapter 11, especially Figure 11.5, for more.

early-stage successional species

In ecology, organisms that resemble pioneer species more than climax-community species. They appear early in the process of succession. See Chapter 5 for more.

earthquake

Vertical shaking of the Earth's surface that results from movement along faults. See Chapter 7 for more.

ecological footprint

Refers to the amount of resources used per person. See Chapter 8 for more.

ecological gradient

A phenomenon whereby the dominant organisms in an area change with distance away from a pollution source. The most resistant to the adverse effects of the pollutant are found at highest-concentration areas. See Chapter 15 for more.

ecological risk assessment

Determines the likelihood that a planned activity will cause adverse effects to natural systems. See Chapter 15 for more.

ecological succession

In ecology, the sequential replacement of dominant species in an area; it is mechanism by which an ecological community comes into being. See Chapter 5 for details.

ecology

The science of interactions among organisms and between organisms and their environments. See Chapter 5 for more.

ecosystem

In ecology, all of the living and non-living entities present in a defined space. See Chapter 5 for more.

ecosystem function

One of the processes carried out by organisms in an ecosystem. Important functions include C fixation, decomposition, and consumption. See Chapter 5 for details.

ecosystem services

Essentially, the benefits we derive from ecosystems. They include carbon fixation, provision of habitat for valued species, enhancement of water supply, and others. See Chapters 5, 6 and 12 for details.

ecosystem structure

In ecology, refers to the specific identity of organisms present in an ecosystem. This property varies widely among ecosystems. See Chapter 5 for details.

electromagnetic spectrum

Describes many different types of energy; often shown on a plot where they are arranged along a continuum according to their wavelengths. See Chapter 14, especially Figure 14.17, for more.

electron

In chemistry, a negatively charged subatomic particle that orbits an atom's nucleus. Neutral atoms contain the same number of protons and electrons. See Chapter 4 for more.

element

In chemistry, refers to a substance that cannot be broken down into simpler units. Each element is made up of the atoms that have the same number of protons. For example, all atoms in the universe with 8 protons in their nuclei are grouped together as the element oxygen. See Chapter 4 for more.

elimination

In toxicology, refers to processes that expel unwanted waste products from the body, including the xenobiotics. See Chapter 15 for more.

endemic species

An organism that is native to its current habitat. See Chapter 5 for details.

endocrine system

One of the systems in the body, it uses hormones to regulate important processes and functions controlling mood, development, reproduction, and growth. See Chapter 15 for more.

endogenous

An adjective that refers to something that originates from within an organism. For example, endogenous errors in DNA are not caused by external agents but by some forces associated with the body. See Chapter 6 for a little more.

endosymbiotic theory

An explanation of the way modern eukaryotic cells possess organelles such as mitochondria. In short, it holds that ancient prokaryotes engulfed other, smaller, prokaryotes in a way that the two organisms ended up living together. See Chapter 6 for more.

endpoint

In toxicology, the outcome caused by a xenobiotic. It is often what is measured or assessed by a toxicologist. Endpoints include paralysis, blindness, death, and others. See also response, See Chapter 15 for details.

energy

Formally, energy is the capacity to perform work. It enables changes to be made to matter. For example, energy allows you to move a box from the floor to a table. See Chapter 4 for more.

energy pyramid

A graphical representation of the way usable energy dramatically declines from lowest to highest trophic levels. See Chapter 5 for more.

entropy

All energy conversions increase the amount of disorder, or entropy, in the universe. Entropy is related to the second law of thermodynamics. See Chapter 4 for details.

environment

A term that refers to all of the factors that surround, influence, and sustain an organism. Earth is made up of many different environments. See Chapter 1, for more.

Environmental Protection Agency (EPA)

U.S. federal agency tasked with protection of soil, air, water, and ecosystems. See Chapter 1 for more.

environmental science

Objective study of Earth's natural systems. Compare to environmentalism. See Chapter 1 for more.

environmental unity, principle of

A governing rule in environmental science that states, in short: all systems are connected to all other systems. The principle of environmental unity also implies that changes made to any system can bring about changes to other systems, even if those other systems are very distant from the initial change. Importantly, the resultant changes can be hard to predict and often are undesirable. See Chapter 1 for more.

environmentalism

Advocacy-based, subjective approach to protecting earth's natural systems; may or may not be informed by objective data or science. Compare to environmental science. See Chapter 1 for more.

erosion

Refers to processes that transport the products of erosion. Typically, sediments are carried in water, wind, or ice. Compare to weathering. See Chapter 3 for more.

ethanol

A kind of alcohol that is generally produced through the fermentation of plant materials like corn or sugar. Ethanol is an important ingredient in alcoholic beverages like wine and beer and can also be used as a liquid fuel. See Chapter 10 for more.

Eukarya, domain

One of the three domains of living things. It contains all eukaryotic organisms. Compare to Archaea and Bacteria. See Chapter 6 for more.

eukaryotic cell

A cell that consists of discreet, membrane-enclosed subcellular structures like nucleus, chloroplast, ribosomes, and mitochondria. Contrast with prokaryotic cell. Also known simply as 'eukaryote'. See Chapter 3 for more.

eutrophication

A process whereby a relatively clear aquatic system is transformed to first a murky then swampy then, potentially, a dry terrestrial system. It is caused by the accumulation of excess nutrients (nitrogen and/or phosphorus). See Chapters 4 and 11 for details.

evaporation

A pathway of the hydrologic cycle. Water changes from liquid to gas. See Chapter 4 for more.

evapotranspiration

A term that refers to the combination of nonbiological evaporation and transpiration. Together, these pathways move water from the lithosphere and hydrosphere into the atmosphere. See Chapter 4 for more.

evolution

A general term that refers to change with time. In ecology, it is used to describe the heritable changes of organisms. See biological evolution, theory of, for more; also see Chapter 6.

exotic organism

In ecology, an organism that is not native to an area in which it currently lives. See Chapters 5 and 6 for more.

exotic species

An organism that is not native to its currents habitat. If it outcompetes native organisms in its new space it is termed 'invasive'. See Chapter 5 for details.

experimental error

In science, the difference between the actual value and the measured value. For example, if the length of a line is 3.2 cm but we measure it as 3.4 cm, the experimental error = 0.2 cm. Experimental error can result from many sources and contributes to uncertainty. See Chapter 2 for more.

experimental science

Research that examines the effects of stimuli or stressors (phenomena that could be defined as causes) on test subjects (systems that react to or are influenced by stressors). See Chapter 2 for more.

exponential growth

A rapid increase in size of a property being measured due to positive feedback. It results from a constant rate of increase, i.e., a constant % of a growing number. See Chapter 2 for details.

exposure

In toxicology, when an organism comes into contact with a xenobiotic. Also see routes of exposure. See Chapter 15 for details.

extinction

A process that leads to the complete disappearance of a species from Earth. See Chapter 6 for more.

extraction, of petroleum

Refers to the removal or recovery of oil and natural gas from underground reservoirs. Pumping is used to extract these fluids. See Chapter 10 for more.

fall

A type of mass wasting event in which material moves downhill; in this case, the falling rocks, etc., do not contact the Earth's surface as they descend. See Chapter 7 for more.

fault

A surface along which there is displacement of rocks. The movement is generally sudden and results from tectonic stress. Movement along faults can cause earthquakes. See Chapter 7 for more.

feedback

In systems analysis, feedback is defined as a system's response to its own output. See negative and positive feedback. See Chapter 2 for more.

fermentation

A natural process carried out by certain microorganisms like yeast. In short, yeast can metabolize plant materials like corn into several products, including ethyl alcohol (i.e., ethanol). Fermentation requires anoxic (i.e., not dioxygen) conditions. See Chapter 10 for more information about and context for ethanol.

fertilizer

In agriculture, a term that refers generally to nutrients added to enhance the growth of crops. Fertilizers can be produced synthetically or can be natural products (e.g., manure). See Chapter 9 for more.

first law of thermodynamics

States that the amount of energy in the universe is conserved, that is, energy cannot be created or destroyed. Conversions among different forms of energy is allowed, though. See Chapter 4 for more.

fixation

A general term related to biogeochemical cycles. It refers to any process that increases the biological availability of an atom. See Chapter 4 for details.

flood plain

The largely flat areas adjacent to either side of a stream channel. Water moves into a flood plain during floods; flood plains are also caused and expanded by repeated flooding. See Chapter 7 for more.

food web

A graphical representation of feeding relationships in an ecosystem. See Chapter 5 for details.

fossil

The remains of an ancient organism. Such remains can take many forms, including impressions in rocks or chemical remnants that were transformed into fuels such as coal, oil, and gas. See Chapters 3 and 10 for more.

fossil fuels

Formed from the remains of ancient organisms. These include coal, oil, and natural gas. They contain a lot of organic carbon compounds that are converted to inorganic products like CO2 when the fuels are burned. See Chapters 3 and 10 for details.

fungi

Classed as eukaryotic microorganisms. Some are unicellular (e.g., yeast), many are multicellular (e.g., mushroom). They are are heterotrophic and tend to act as important decomposers in ecosystems. See Chapter 3 for more.

gastrointestinal

An adjective that indicates digestive system. In toxicology, gastrointestinal or GI exposure occurs when a xenobiotic is swallowed. See Chapter 15 for more.

gene pool

A term used to refer to the total amount of genetic information within the DNA of all members of a species. See chapter 6 for more.

generalist

In ecology, an organism that occupies a broad niche, that is, it can live under a wide range of conditions and meet its needs with varying strategies. Contrast with specialist. See Chapter 6 for more.

genetic diversity

How many different genetic codes are present in an area. See Chapter 6 for more.

genetically modified organism

Through the use of one of several techniques, the DNA sequence of an organism can be altered. For example, cross breeding of related species can produce offspring with desired traits; genetic engineering that involves direct manipulation of DNA sequences is a modern way to alter the genome of an organism. Genetically modified organisms are often used in agriculture to increase food production, but they are controversial. See Chapter 9 for more.

genome

Refers to all the information that can be used to copy cells. See Chapter 6 for more.

geologic actualism

An approach to understanding Earth's history. It is a modification of uniformitarianism and holds that Earth's past was like the present day in that most features arise from slow and steady processes but allows that occasional short-lived catastrophes can also play roles. See Chapter 3 for more.

geothermal energy

Associated with the Earth's molten core and responsible for the tectonic cycle. See Chapter 4 for more.

glacier

A large mass of ice that persists for many years (up to hundreds of thousands) despite changing seasons. It is formed through accumulation and compression of snow into ice. Glaciers can profoundly shape the surface as they slowly move, store a lot of freshwater, and contain clues about past climatic and atmospheric conditions. See Chapters 4 and 14 for more.

global climate change

One of the major consequences of global warming. As the name suggests, climates around Earth are altered in subtle and dramatic ways, e.g., amount of rainfall, average temperature, cloud cover. Notably, because of the complexity of Earth's systems, some local areas can become cooler even as the overall temperature of the planet increases. See Chapter 14 for more.

global conveyor belt

Refers to the way water circulates around the globe. Notably, warm waters moves northward from the equator to northern latitudes. See Chapter 14 for more.

global warming

An increase in the average temperature of the Earth. See Chapter 14 for more.

global warming potential

GWP. How much warming a gas can cause relative to that caused by carbon dioxide. So, carbon dioxide has a GWP = 1, methane has a GWP = 24, and nitrous oxide has a GWP =

3007, meaning one methane causes as much warming as 24 carbon dioxide molecules and one nitrous oxide causes as much warming as 3007 carbon dioxide molecules. See Chapter 14 for details.

glucose

A form of carbon that is biologically available to the whole biosphere. It is organic, having the formula C6H12O6. It is produced by fixation of carbon dioxide.

gravitational energy

Associated with the pull all objects in space have on each other, e.g., the way the moon and Earth are attracted. It is responsible for the tides. See Chapter 4 for more.

greenhouse effect

A natural atmospheric phenomenon. It leads to a warmer planet than Earth would be without it. Certain gases (known as greenhouse gases) absorb outgoing heat and then re-radiate much of that energy back to the surface; as a result, Earth's surface warms. Human activity increases the amount of greenhouse gases in the atmosphere and intensify the greenhouse effect. See Chapter 14, especially Figure 14.18, for more.

greenhouse gases

A group of gases found in Earth's troposphere. They absorb heat (i.e., infrared energy) that is released from Earth's surface and re-radiate some of that energy back to Earth. Greenhouse gases come from both synthetic and natural origins and include carbon dioxide, nitrous oxide, water vapor, and several industrial products. See Chapter 14 for more.

groundwater

A reservoir of the hydrologic cycle located below the Earth's surface. See Chapter 4 for details.

growth rate

In demography, the net change in the size of a population; it is generally expressed as a percentage. See Chapter 8 for more.

guano

Bird excrement that tends to be rich in phosphorus. See Chapter 4 for more.

GWP

See Global Warming Potential.

habitat

In ecology, refers to the physical location in which an organism lives. Compare to niche. See Chapter 5 for details.

habitat destruction

The physical environment in which an organism lives is rendered incapable of supporting that organism; the term generally refers to human activity. In the case of species with very specific needs, habitat destruction can lead to extinction. See Chapter 6 for more.

habitat fragmentation

Barriers to the movement of organisms can effectively cut up a large habitat into small sections. Highways, railways, and pipelines are among the structures that can fragment habitat. In some cases, it can endanger organisms. See Chapter 6 for details.

half-life

The length of time required for half of the unstable atoms in a space to decay to stable products. See Chapter 10 for more.

halogens

A group of elements that have similar chemical properties. Among others, the halogens contain environmentally important bromine (Br) and chlorine (Cl).

hazard

An intrinsic property of a substance or an activity related to how much harm it will cause. Also, certain substances or devices are referred to as hazards (e.g., a step ladder or laboratory solvent). See Chapter 15 for more.

hazardous waste

Refers to waste that poses more risk to human and environmental health than does domestic waste. Some waste from industry, hospitals, universities, and other facilities is classified as hazardous if it meets certain criteria. See Chapter 13 for details.

heat

A low-grade form of energy that is the result of degradation of high-grade energy. See Chapter 4 for details.

herbicide

A type of pesticide that specifically targets unwanted plants (i.e., weeds). See Chapters 9 and 15 for details.

herbivore

An organism that consumes plants. See Chapter 5 for details.

heritable

Passed from parents to offspring. See Chapter 6 for more.

heterotroph

A type of organism that relies on preformed, organic carbon to meet its nutrient needs. Includes animals, fungi, and many bacteria. Heterotrophic an adjective. Compare to autotroph. See Chapter 5 for more.

hormesis

In toxicology, the idea that there is an optimum dose of a xenobiotic; both low and high doses can cause harm. See Chapter 15 for details.

Hubbert Curve

Named after an American geologist who predicted a rise and then fall in petroleum production, also the notion of "peak oil". See Chapter 10, especially Box 10.3, for more.

hydraulic fracturing (fracking)

One widely used unconventional petroleum-recovery technique. It involves injection of fluids to free up fluids stored in pores of reservoir rocks. See Chapter 10, especially Figure 10.11, for more.

hydro power

A term used to refer to the ways humans can capture the energy of moving water to generate power. See Chapter 10 for more.

hydrocarbons

A general chemical term that refers to chemical compounds made up of hydrogen and carbon atoms bonded together. This term encompasses a large array of compounds with different numbers of C and H atoms connected in different ways. See Chapter 10 for details and context.

hydroelectric power plants

A facility that uses moving water to spin turbines and generate electricity. See Chapter 10 for more.

hydrogen

A chemical element made up of atoms that have 1 proton and 1 electron. It is an important building block for organisms and is the most abundant element in the universe.

hydrologic cycle

Describes the pathways and reservoirs of Earth's water. See Chapters 4 and 11 for more.

hydrosphere

One of Earth's four spheres, this one contains all of Earth's water. See Chapter 4 for more.

hypersensitivity

In toxicology, an endpoint caused by exposure to certain xenobiotics: the immune system has a more pronounced and inappropriate reaction to substances that enter the body. It is similar to allergic reaction. See Chapter 15 for details.

hypothesis

A proposed explanation for an observed phenomenon. See Chapter 2 for more.

igneous rocks

In geology, rocks formed directly from cooling magma or lava. See Chapter 3 for more.

immunosuppression

In toxicology, an endpoint that may be caused by xenobiotic exposure: the immune system does not work as well as it once did, meaning the affected organism is more susceptible to infectious agents and cancers. See Chapter 15 for details.

immunotoxicity

In toxicology, refers to damage done to the immune system by a xenobiotic. See Chapter 15 for details.

in-stream water use

When humans go to a reservoir and utilize the water in place (rather than removing the water from the reservoir to use it elsewhere). Examples include boating, swimming, fishing. See Chapter 11 for more.

incineration (as waste management strategy)

Combustion of waste at very high temperature in a structure known as an incinerator. These facilities are designed to reduce the volume of waste and release minimal air pollution. See Chapter 13 for more.

indigenous

An organism that is native to an area. An exotic organism would be the opposite, one that came into an area from an outside system. See Chapter 5 for more.

inductive reasoning

A strategy used by scientists in their study of the natural world. It involves the observation of specific, representative phenomena to develop general conclusions or rules that govern the universe. See Chapter 2 for more.

infant mortality

In demography, refers to the number of pregnancies that end in a stillborn baby; it is generally expressed as a number per 1000, e.g., 20 / 1000. See Chapter 8 for more.

infiltration

A pathway of the hydrologic cycle. Water flows vertically downward due to gravity. See Chapter 4 for more.

infrared radiation (IR)

Also known as heat, it is a type of energy that behaves differently than light and ultraviolet energy (and all the other energy types on the electromagnetic spectrum). See Chapter 14 for more.

inhalation

When an airborne substance enters a body through the respiratory system (i.e., nose/mouth). In toxicology, exposure via inhalation is of concern for gaseous and particulate xenobiotics. See Chapter 15 for more.

inorganic

In chemistry, refers to chemical compounds that lack at least one of the following: a C atom and at one C-H bond. Compare to organic. See Chapter 4 for details.

inorganic fraction

The portion of soil that was derived from rocks. See Chapter 9 for more.

integrated waste management (IWM)

The simultaneous use of multiple strategies to collect, handle, and minimize waste. See Chapter 13 for more.

intended use, of water

An important concept related to water quality. It refers to the way a given quantity of water will be utilized by humans. Different uses include consumption, irrigation, cleaning; different uses tend to require different levels of water quality. See Chapter 11 for more.

Intergovernmental Panel on Climate Change (IPCC)

A group of scientists working through the United Nations. It evaluates scientific data from studies of climate change to predict future trends and make recommendations. See Chapter 14 for more.

interspecific competition

Rivalry among different populations, that is, among organisms that are not from the same species. See Chapter 5 for more.

intraspecific competition

Rivalry among members of a species, that is, those in the same population. Compare to interspecific competition. See Chapter 5 for more.

intravenous

Refers to injection into a blood vessel. In toxicology, intravenous or IV exposure is of concern because a xenobiotic can quickly enter the bloodstream. See Chapter 15 for more.

invasive species

In ecology, exotic species that outcompete native organisms. Often, invasives can take over an area. See Chapters 5 and 6 for more.

invertebrate

An animal that lacks a backbone. This is a large group, and includes worms, jellyfish, and spiders. Compare to vertebrate.

ion

In chemistry, when a neutral atom gains or loses electrons, thereby gaining or losing negative charge, it becomes a charged ion. Loss of electrons produces a positive ion, gain of electrons produces a negative ion. See Chapter 4 for more.

irrigation

In agriculture, the application of water onto crop fields. See Chapter 9 for more.

isotopes

Atoms with the same number of protons but a different number of neutrons in the nuclei. For example, U235 and U239 are different isotopes of uranium. Each has 92 protons (as well as 92 electrons), but the former has 143 neutrons and the latter has 146 neutrons. See Chapter 4 for details and Chapter 10 for some relevance and context.

K strategy

In ecology, the reproductive strategy that involves producing a small number of offspring that are cared for by their parents. See Chapter 5, especially Box 5.2 for more.

kerogen

A waxy intermediate between dead organisms and petroleum. It forms from remains of buried marine microorganisms; it can undergo further transformations into petroleum. See Chapter 10 for more.

keystone species

In ecology, an organism that plays a uniquely critical role in an ecosystem. If such an organism disappears, the affected system is likely to undergo dramatic adverse changes. See Chapter 6 for details.

kinetic energy

Energy in motion. See Chapter 4 for details.

landfill

A structure designed to contain waste. See chapter 13 for details.

landslide

One type of mass wasting event: it involves the rapid movement of materials downhill; those materials slide along a slope. See Chapter 7 for more.

larvae

A larva (plural: larvae) is an immature form of an animal, one that can undergo a dramatic metamorphosis and develop into an adult. For example, caterpillars are butterfly larvae.

late-stage successional species

In ecology, organisms that are more like climax-community species than pioneer species. See Chapter 5 for more.

lava

In geology, molten rock that flows on top of Earth's surface. See Chapter 3 for more.

law, scientific

A theory that has persisted for many years and continues to persist in the face of attempts to discredit it may be elevated to the status of scientific law. Generally, laws can be expressed by a (relatively) simple mathematical formula. Few scientific notions are given this standing. The laws of thermodynamics and the law of gravitation are two examples. Although widely supported by data and history, a law may be subject to review or change as our understanding of the universe increases. Compare to theory and conclusion. See Chapter 2 for more.

leaching

A process whereby a pollutant is carried vertically downward in moving water. It is of concern because it can lead to groundwater pollution. See Chapters 11 and 15 for details.

lethal dose

In toxicology, the tissue concentration of a xenobiotic that causes death in an exposed organism. See Chapter 15 for details.

levee

A ridge of material that lies at the edge of a stream channel; levees run parallel to a stream. A levee forms via flooding and grows with each flood. See Chapter 7 for more.

lichen

An example of a very close mutualistic relationship involving algae and fungi. This association is so evolved and so close that lichens take on a form that is distinct from that of each partner. See Chapter 5 for details,

limiting factor

The one requirement for growth that is least abundant relative to demand for it. It is a concept and problem relevant for both natural and agricultural systems. See Chapter 9 for details.

lithification

In geology, the multiple processes that produce sedimentary rocks from unconsolidated sediments. See Chapter 3 for more.

lithosphere

One of Earth's four sphere. This term refers to the brittle, outer layer of the Earth made up of the entire crust and the upper portion of the mantle. It is nonliving and lies on top of the asthenosphere. See Chapter 3 for more.

liver

The organ in which most xenobiotic metabolism occurs. See Chapter 15 for details.

logistic growth

A type of population growth curve that is more or less S-shaped. It is characterized by an initial slow phase which is followed by a rapid, exponential phase, and then another slow, nearly flat phase. See Chapter 8, especially Figure 8.2, for more.

Love Canal

City in western New York State (USA) that became famous when hazardous waste buried beneath it came to the surface. National attention and outrage led to new environmental protection laws. See Chapter 1, especially Table 1.1, for more.

magma

In geology, molten rock that moves below the Earth's surface. See Chapter 3 for more.

malaria

A potentially deadly disease caused by a kind of protozoan; that microorganism is itself spread by a certain type of mosquito.

Malthus, T.R.

18th-Century economist and cleric who predicted that the human population would exceed carrying capacity and therefore suffer war, starvation, and other adverse consequences. He is credited with the 'Prophecy of Malthus.' See Chapter 8, especially Box 8.6, for more.

mantle

A geology term that refers to the middle zone of the Earth that sits between the crust and core. It is thick, mostly plastic in behavior, and has a higher density than does the crust. See Chapter 3 for more.

Marine Protection, Research, and Sanctuaries Act

A U.S. law passed in 1988 that, among other things, restricted the dumping of waste into the ocean. See Chapter 13 for more.

mass extinction event

Refers to a relatively short period of geologic time during which 75% or more of existing species ceased to exist on Earth. See Chapter 6 for more.

mass wasting

Generally, the gravity-driven downward movement of Earth's materials from high to low elevation. There are several different types of mass wasting events. See Chapter 7 for more.

material

Physical matter that takes up space and has mass. This term is very general and includes many things, for example, individual atoms, chemical compounds, water, soil, and fossil fuels.

matter

A formal chemical term that refers to anything that takes up space and has mass. See Chapter 4 for more.

maximum sustainable yield

In agriculture, the largest amount of a crop that can be sustainably grown in a field. See Chapter 9 for more.

mean

One way to represent multiple readings of the same phenomena. For example, five separate readings of the temperature of a container of water could yield five different readings. The mean is often chosen as the best representation of the answer in such a situation. Calculated by adding up the values of all measurements made and then dividing by the number of measurements; equivalent to average. See also standard deviation. See Chapter 2 for more.

metabolism

Two meanings are relevant to environmental scientists. 1: chemical breakdown of food so nutrients can be obtained by the organism ingesting that food. See Chapter 5 for details. 2: In toxicology, the chemical breakdown of a xenobiotic; usually done to increase the likelihood of elimination (i.e., excretion) of the xenobiotic. See Chapter 15 for details.

metamorphic rocks

In geology, one of the three rock types. These are formed from rocks subjected to solid-state changes. Increased temperature and/or pressure will alter existing rocks into new, metamorphic rocks. See Chapter 3 for more.

methane

A gaseous form of C consisting of 4 H atoms bonded to 1 C atom. See also natural gas. See Chapters 4, 10, and 13 for more.

microbe

Anything that is too small to be seen without the aid of a magnifier. Microorganisms are living microbes. See Chapter 3 for more.

micron

A metric unit of length that is equal to 0.0000001 meter. The abbreviation for this unit is μm. See Chapters 2 and 3 for more.

microorganism

Refers to organisms that are too small to be seen without the aid of a magnifier, including bacteria, protozoa, fungi, algae, and viruses. In the case of algae and fungi, some of the forms are large enough to be seen with the naked eye, but they are included here because some of their forms are microscopic. See Chapters 1, 3, and 4 for more.

microplastics

Very small plastic particles. These are classified as between 1 micron and 5 millimeter in size. They can be generated through the disintegration of larger plastic trash that is improperly disposed of and are also manufactured for inclusion in certain products. Microplastics pollute water and food. See Chapter 13 for more.

mid-stage successional species

In ecology, organisms that dominate during the intermediate stages of ecological succession. See Chapter 5 for more.

Mithridates VI

King of Pontus during the last century, BC. Important to modern-day toxicologists because of the way he built up his resistance to poisons by daily ingestion of low doses of all known toxic substances (to thwart would-be assassins). See Chapter 15, especially Box 15.1, for more.

mitochondria

An important subunit or organelle in eukaryotic cells that converts energy into usable forms for a cell. See Chapter 3 for more.

molecule

In chemistry, two or more atoms held together by bonds. See Chapter 4 for more.

monoculture

In agriculture, the practice of growing a single crop in one field. See Chapter 9 for more.

motile

A term that refers to an organism that can move under its own power. Motility is the noun. See Chapter 3 for more.

mountaintop removal

A type of surface mining of coal which involves the clearing of existing mountain peaks to expose a coal seam. See Chapter 10 for more.

multicellular

An adjective that refers to an organism made up of many cells that communicate with each other. Humans and plants, for example, are multicellular whereas bacteria are unicellular. See Chapter 3 for more.

municipal solid waste (MSW)

Refers to garbage produced by households and other domestic activities. See Chapter 13 for more.

mutation

Refers to a change from the normal or original genetic code. Mutations tend to damage or kill affected organisms but may impart new and helpful traits. See Chapter 6 for details.

mutualism

In ecology, a relationship among organisms in which all participants benefit. See Chapter 5 for details.

natural

In the context of environmental science, a phenomenon or process that occurs in the absence of human activity. See Chapter 7 for more.

natural gas

An important fossil fuel, it is mostly comprised of gaseous methane. See methane for more. See also Chapter 10 and 13.

natural selection

In evolutionary theory, individuals that are best adapted to survive in their environments will reproduce better than those are less well adapted. See Chapter 6 for details.

negative feedback

In a system, output attenuates or cancels previous output such that a system does not change appreciably from initial conditions. Compare to positive feedback. See Chapter 2 for more.

neutron

In chemistry, a subatomic particle found in the nucleus of an atom. These are uncharged. Different atoms of the same element may vary in the number of neutrons they possess, even as they must have the same number of protons. Atoms having the same number of protons but different numbers of neutrons are called isotopes. See Chapters 4 and 10 for more.

niche

The role an organism plays in its habitat. A niche can be narrow, that is, organisms occupying it have very specific needs and activities. Slight changes to the conditions in an ecosystem could devastate such organisms. On the other hand, a niche can be broad, meaning that organisms in it can use a wider variety of strategies to meet their needs. In this latter case, organisms are more likely to survive changes to environmental conditions. Compare to habitat. See Chapters 5 and 6 for more details.

nitrate

A form of nitrogen with the formula NO3-. It is inorganic and biologically available to plants, algae, and some microorganisms. It can cause water pollution, including eutrophication. See Chapter 4 for more.

nitrogen

A chemical element made up of atoms with 7 protons and 7 electrons. Certain forms of N are essential for organisms' survival. See Chapter 4 for more.

nitrogen-fixing bacteria

Specialized bacteria that can convert dinitrogen into ammonium. They are crucial in providing biologically available N to most of the biosphere. See Chapters 4, 5, and 9 for more.

non-consumptive (water use)

A type of off-stream use of water that quickly returns the removed water to the reservoir from which it was extracted. For example, water can be pumped from a river to cool an industrial process and then returned shortly thereafter. See Chapter 11 for more.

non-point pollution

Diffuse, non-discreet sources of water pollution. Unlike point sources, they are difficult to identify and remediate. See Chapter 11 for details.

non-point source of water pollution

Contaminants that enter water from multiple, hard-to-identify outlets. Runoff from agriculture is a good example of non-point source pollution. Compare to point source pollution. See Chapter 11 for more.

non-potable water

Non-potable water is not safe for human consumption (opposite: 'potable water'). See Chapter 11 for more.

non-renewable energy sources

Those that are finite in that they cannot be replaced fast enough to keep up with their rate of usage. Examples include coal, oil, natural gas, and nuclear energy. See Chapter 10 for more.

non-target organism

An organism that humans do not intend to kill with pesticides, antibiotics, or other poisons; generally, an organism that provides some benefit to humans. However, non-target organisms often are affected nonetheless. See Chapters 9 and 15 for more.

nuclear fission

The process that splits a single atom into several products; it releases an enormous amount of energy, some of which can be harnessed to generate electricity. See Chapter 10 for more.

nuclear fuel rods

Thin tubes packed with fissionable material such as uranium atoms. These rods undergo a chain reaction inside a nuclear reactor. See Chapter 10 for more.

nuclear meltdown

An accident at a nuclear power plant that involves the uncontrolled heating of the reactor core such that it starts to melt. See Chapter 10 for more.

objective

An approach to study used by scientists. It is characterized by observation of phenomena without imposing judgements or other biases on the results. Contrast with subjective. See Chapter 1 for more.

observational science

The status of a test subject at a moment in time is studied and reported. Observational science can be used to classify objects into previously established categories; for example, an unknown mineral sample could be evaluated and identified according to preexisting criteria and grouped accordingly. See Chapter 2 for more.

off-stream water use

When humans extract water from a reservoir to utilize it elsewhere. For example: irrigation, domestic consumption. See Chapter 11 for more.

oil seep

A natural opening in the seafloor through which small amounts of petroleum can enter ocean water. See Chapter 10 for more.

oil trap

An impermeable geologic structure against which migrating petroleum is concentrated. See Chapter 10 for more.

open burning

Once a widely used practice to manage waste, ihis involves the simple incineration of trash without air pollution controls. See Chapter 13 for more.

open dump

A strategy for waste management that was used widely in the past. Garbage is dumped onto Earth's surface without containment. Note that open dumps are not the same as sanitary landfills. See Chapter 13 for more.

organelle

Subcellular structure such as nucleus, mitochondria, chloroplast, etc. They are found in eukaryotic but not prokaryotic cells. See Chapter 3 for more.

organic

In chemistry, an adjective that refers to chemical compounds that consist of at least one C atom that is bonded to at least one H atom. Compare to inorganic. See Chapter 4 for details.

organic fraction

The portion of soil that was derived from organisms' waste and remains; synonymous with SOM (soil organic matter). It consists of many sub fractions with specific chemical and physical properties. See Chapter 9 for more.

Organization of Petroleum Exporting Countries (OPEC)

Comprised of about 15 nations (membership fluctuates) that control about 80% of Earth's oil reserves. See Chapter 10 for more.

overburden

The geologic material that lies on top of a coal seam; it must be either removed or penetrated by mines so the coal can be extracted. See Chapter 10 for more.

overdraft

In the context of water usage, when water is extracted by humans from the groundwater reservoir faster than it is replaced by natural processes. See Chapter 11 for more.

overgrazing

In agriculture, when the feeding by livestock exceeds the carrying capacity of grasses in a field. See Chapter 9 for more.

oxic

A space that contains sufficient dioxygen gas to support aerobic organisms. Compare to anoxic and anaerobic.

oxygen

A chemical element made up of atoms with 8 protons and 8 electrons. It is an important building block for organisms. Two O atoms bonded together makes up dioxygen, a gas which is essential to the survival of aerobic organisms.

ozone

A molecule made of e oxygen atoms bonded together. See Chapter 14 for details.

ozone hole

An area in the stratosphere in which there is less ozone than is typical. See Chapter 14 for more.

ozone layer

The ozone molecules located in the stratosphere. This layer absorbs much of the harmful ultra violet light that reaches the Earth from the sun. See Chapter 14 for more.

ozone-depleting substances (ODS)

A broad term given to chemical compounds that can lead to destruction of stratospheric ozone. See Chapter 14 for more.

Pangaea

In geology, an ancient landmass made up of all Earth's continents. It started to break up due to movement of tectonic plates about 200 million years ago. See Chapter 3 for more.

Paracelsus

16th-Century scientist whose work led to the dose-response principle in toxicology. See Chapter 15 for more.

parasitism

In ecology, an association in between organisms in which one is harmed and one derives benefit. Typically, a smaller organism, the parasite, lives on or in a larger organisms, the host. See Chapter 5 for details.

parent material

Existing substances that can be modified into soil via natural physical, chemical, and biological processes. Parents include rocks and the remains / waste products of organisms. See Chapter 9 for more.

parent xenobiotic

In toxicology, the form of the compound to which an organism is exposed, the form that enters an organism. The parent compound may be chemically altered within an organism. See Chapter 15 for details.

pathogen

A general term that includes any disease-causing organism. Some bacteria, fungi, protozoa, and algae are on this list, as are many viruses. Certain worms and other larger organisms can act as pathogens as well. See Chapters 3 and 5 for more.

pathways

Refers to natural processes that move materials among earth's natural reservoirs. For example, evaporation moves water from surface reservoirs like the ocean to the atmosphere. See Chapter 1 for more.

peat

A precursor to coal, formed from the decomposing remains of plants. See Chapter 10 for more.

penicillin

The first antibiotic discovered by humans. Produced by a certain fungal species, it inhibits cell wall synthesis in many bacteria. Thus, penicillin can cure an infected person of a disease caused by susceptible bacterial pathogens. See Chapters 5, 6, and 15 for details.

pesticides

Refers to a broad class of chemical compounds intended to kill unwanted organisms. Widely used in agriculture, they can kill pests such as weeds, certain insects, and others, but they also can harm organisms we wish to protect, including humans. See Chapters 1, 9, and 15 for more.

petri dish

A laboratory tool in which microorganisms such as bacteria can be grown. See Chapter 3 for context.

petroleum

A broad term that refers to fuels derived from ancient marine microorganisms. These fossil fuels include oil and, according to some, natural gas. See Chapter 10 for more.

pH

In chemistry, a way to express the acidity of a water-based solution. The pH scale runs from 0-14, with the lower numbers (those below 7) indicating acids and higher numbers (those above 7) indicating bases. A neutral solution is pH 7. See Chapters 9 and 10.

phosphorus

A chemical element, P, made up of atoms have 15 protons and 15 electrons. It is an essential nutrient for organisms, but its relatively scarcity on Earth can limit production and growth in ecosystems. Excess P can cause eutrophication in water. See Chapter 4 for more.

photochemical smog

Sometimes called Los Angeles smog. A local air pollution phenomenon caused by the specific conditions in LA, California (and other places): low wind, high sun, and high density of cars and trucks (which release hydrocarbons and nitrous oxide, two important primary pollutants). Compare to sulfurous smog. See Chapter 14 for more.

photosynthesis

A process carried out by certain organisms (plants, algae, and some bacteria) in which the energy of the sun is used to drive the fixation of carbon. See Chapters 4 and 5 for more details.

phototroph

An organism that uses the sun as its energy source. See Chapter 5 for more.

photovoltaic cells (PV cells)

Small constructed objects that contain specialized chemicals able to absorb and convert solar energy into electricity. Many cells can be combined to form a solar panel. See Chapter 10 for more.

phylogenetic tree

A graphical representation of relationships among organisms based on comparisons of DNA sequences. See Chapter 6 for more.

phytoplankton

Plant-like plankton that are producers; they are consumed by zooplankton. They live in surface waters where sunlight is available. They tend to be on the order of a few millimeters or so in size. See Chapter 5 for more.

phytoremediation

The use of plants to clean up contaminated soil. See Chapter 15 for more.

pioneer species

In ecology, the organisms that can colonize barren landscape and thus initiate ecological succession. See Chapter 5 for more.

placental barrier

Physiological features within the placenta that protect a developing offspring from many xenobiotics. See Chapter 15 for details.

plankton

Small organisms that float at or near the surface of bodies of water. They range in size from less than a millimeter to a centimeter. See phytoplankton and zooplankton.

plate tectonics, theory of

In geology, describes how the lithosphere of Earth is broken into several units called plates. These plates slowly move horizontally with respect to each other. Plate movements are responsible for continental drift, earthquakes, volcanoes, and other important phenomena. See Chapter 3 for details.

point sources of pollution

Discrete outlets that release contaminants into water reservoirs. For example, a pipe that drains from an industrial site is a point source of pollution. These are relatively easy to identify and remediate; compare to non-point sources of pollution. See Chapter 11 for more.

poison

A general term that includes substances that can cause harm to living systems. See Chapter 15 for details.

pollutant plume

The more or less triangular shape a contaminant in flowing water takes after it is released. Concentration is highest at the point at which the pollutant is introduced and diminishes with distance from the source. See Chapter 11, particularly Figure 11.7, for more.

pollution/pollutant

Terms referring to something that degrades the quality of air, water, or soil. The term 'pollutant' is used to refer to that material that degrades quality. Compare with 'contaminant', a word with a related, if distinct, meaning. In short, a contaminant can be thought of as any substance that does not belong where it is found (i.e., a metal dissolved in water). A pollutant is a contaminant that has reached a high enough concentration to cause adverse effects (i.e., that metal noted in the previous sentence is at such a high level that it is killing fish; see Chapter 11 for more about water pollution). Note that the terms are often used interchangeably by scientists and members of the public alike.

population

In ecology, a group of organisms that are all members of the same species. See Chapter 5 for details.

population risk

A cause of extinction of a species that stems from a low number of individuals of that species. See Chapter 6 for more.

positive feedback

In a system, current output accentuates or magnifies the effect of previous output. This type of feedback leads to change away from initial conditions. Compare to negative feedback. See Chapter 2 for more.

potency

In toxicology, refers to the dose of a xenobiotic required to bring on an observable response. Low-potency compounds require high doses and high-potency compounds require low doses. See Chapter 15 for details.

potential energy

Stored energy. See Chapter 4 for details.

power

A term related to energy and the capacity to do work. It can be derived mathematically: it is the amount of energy transferred per time unit. Conceptually, power allows a certain amount of work to be done in a certain period of time. It can be expressed with several different units, including the watt. A power generation plant receives thermal power from the burning of a fuel like coal and releases mechanical power in the form of electricity. See Chapter 10 for context.

power grid ("the grid")

A collection of generation facilities and power lines that produce and distribute electricity to a large number of users. See Chapter 10 for more.

precautionary principle

In short, it holds that new chemicals, technology (etc.) are assumed to be hazardous and risky until proven otherwise. Often drives regulations on new substances and devices introduced into commerce. See Chapter 1 for more.

precipitation

A pathway of the hydrologic cycle. Water that has condensed in the atmosphere falls to the surface as liquid or solid. See Chapter 4 for more.

precision

How well multiple readings, or replicates, of the same phenomena agree with each other. Contrast with accuracy and sensitivity. See Chapter 2 for details.

precursor phenomena

Observable events that signal an earthquake is imminent. See Chapter 7 for more.

predation

An interaction between organisms in which one is harmed and the other derives benefit. Specifically, a predator consumes a prey organism. See Chapter 5 for details.

predator-prey relationship

In ecology, an association between organisms in which one, the predator, lives at the expense of the other, the prey. Predators consume prey. See Chapter 5 for details.

primary energy sources

Their energy be converted directly into electricity or heat; oil, gasoline, coal, solar energy, and others are primary energy sources. See Chapter 10 for more.

primary pollutant

A substance that has not been altered from its original form, that is, the form that is released into a natural environment. For example, carbon monoxide emitted from the tailpipe of car. See Chapter 14 for more.

primary producers

Organisms that fix carbon for all members of an ecosystem;

they are autotrophs. They also tend to convert non-biologically available energy into biologically available forms. For example, plants both fix carbon and convert the sun's energy to chemical bond energy. Consumers live off of the C and energy stored in primary producers. See Chapters 4 and 5 for more.

primary succession

In ecology, ecological succession that follows a disturbance extreme enough to wipe out an existing community. Pioneer species can colonize a barren landscape in this case. Compare to secondary succession. See Chapter 5 for more.

primary treatment (of sewage)

One of the stages of sewage treatment, involves the physical separation of solids from liquids. See Chapter 11 for more.

Prince William Sound

Alaska (USA), the site of a major marine oil spill in 1989. This event raised many concerns about marine oil tankers, oil usage, and contamination of pristine environments. See Chapter 1 for more.

principle of dose-response

Foundational principle in toxicology that states the effects caused by a xenobiotic will depend on the concentration of that xenobiotic in a body or, more likely, at a target site within a body. See Chapter 15 for details.

prokaryotic cell

A type of organism that lacks any membrane-bound organelles such as a nucleus. These tend to also be small and single celled. Bacteria are important prokaryotes. Also known simply as 'prokaryote.' See Chapter 3 for more.

prophylactic

A general term indicating a measure taken to prevent something unwanted from happening. Note how prophylaxis is distinct from responding to something after it has occurred.

proton

Positively charged subatomic particle found in the nucleus of an atom. The number of protons varies by, and defines, atom type (i.e., atoms of specific elements are defined by the number of protons in their nuclei). See Chapter 4 for more.

protozoa

A group of organisms that are microscopic, eukaryotic, mobile (also known as "motile"), and usually heterotrophic. They are roughly ten times larger than bacteria; they also tend to prey on bacteria. See Chapter 3 for more.

proven reserves

The amount of any material that is known to exist and can be recovered (i.e., obtained) with existing technology. For example, "petroleum reserves" refers to the oil we can count on. See Chapter 10 for more.

r strategy

In ecology, a reproductive strategy that involves the release of high numbers of offspring but little or no parental care. See Chapter 5, especially Box 5.2 for more.

radiation

A kind of energy that can travel at the speed of light from some source through air or even a vacuum. There are many different types of radiation. Some high energy radiation such as gamma rays can penetrate deep into objects (living and non living) and cause chemical changes. This kind of ionizing radiation can damage living cells. See Chapters 10 and 15 for more.

radioactive decay

A process whereby an unstable atom emits energy and particles as it is transformed into a smaller atom. See Chapter 10 for more.

radioactive waste

Products of nuclear fission, some of this emits low levels of radiation and some emits high levels of radiation. In the latter case, long-term containment is necessary to minimize risk of exposure of humans and other organisms to dangerous radiation. See Chapter 10 for more.

random errors

Cause measurements of a phenomena to be inaccurate in an inconsistent way (i.e., either above or below the actual value). For example, five different scientists might read a meter stick from five different angles and report five different answers; some of the answers will be too high and some will be too low. Contributes to scientific uncertainty. Compare to systematic errors. See Chapter 2 for details.

reaction, chemical

In chemistry, occurs when two or more atoms are brought close together and interact. Reactions can lead to breaking of existing bonds and formation of new bonds. In some cases, separate smaller units are combined into larger units via reactions. Energy released during some reactions can be captured by organisms and used to do the work of living. See Chapter 4 for more.

recharge

Refers to replacement of water in the groundwater reservoir. Recharge is largely the result of precipitation followed by infiltration of water downward. Melting snow and ice at the surface can also contribute. See Chapter 4 for more.

recombination

A process that combines genetic information from two parents into a new offspring; changes in a population are a result. See Chapter 6 for more.

recreational

An adjectival word that refers to a substance that is not medically necessary; a person may ingest it for pleasure. Compare to therapeutic. See Chapter 15 for more.

recycling

A waste management strategy characterized by the collection and reprocessing of paper, glass, plastics, metals, and other materials (including hazardous waste) so they can be used again. See Chapter 13, especially Figures 13.12 and 13.13, for more.

refining, oil

The process by which several different fuels are derived from crude oil. It involves heating and high pressure to yield such products as gasoline, motor oil, kerosene, and others. This work is done at a facility known as an oil refinery. See Chapter 10 for more.

remediation

Clean up or repair of an existing problem. See Chapter 15 for more.

renewable energy sources

Those that can be replaced in a relevant period or time and/or are so abundant that human usage has no appreciate effect on their supply. Solar energy is a good example because of the vastness of its supply. See Chapter 10 for more.

replacement rate

In demography, refers to the number of births required to maintain a population indefinitely. See Chapter 8 for more.

replicates

A shorthand word often used by scientists to refer to multiple measurements of the same phenomena: each reading is a replicate. Put another way, we replicate or repeat a measurement multiple times. See also precision. See Chapter 2 for more.

reservation

In ecology, refers to an approach to wildlife and ecosystem protection. The goal of preservation is to protect wild living entities for their own sake, not to provide direct benefits to humans. Compare to conservation. See Chapter 6 for more.

reserves

The amount of a material that is known to exist and can be obtained with existing technology. The term is often applied to oil and natural gas. See Chapter 10 for details.

reservoirs

Refers to natural sites in which important materials such as water, petroleum, nitrogen, and many others, are stored. See Chapter 1 for more.

residence time

The length of time an average unit of a particular material remains inside a system. It can be calculated as shown in Box 2.4.

resources

The quantity of material that includes reserves as well as other, more tenuous pools of materials. Often used in reference to oil and natural gas. See Chapter 10 for details.

respiration

A process carried out by living things. The chemical bonds in relatively large molecules are broken with the following results: some energy is harnessed by the respiring organism and the large molecule is broken down into relatively small fragments. See Chapter 4 for more.

response

In toxicology, an outcome caused by a xenobiotic. See also endpoint. See Chapter 15 for details.

risk

The likelihood that a substance or activity will cause harm. See Chapter 15 for more.

risk characterization

The final step or product of a risk assessment. Typically, it is a short statement of the risk or safety of a substance or activity. For example, a risk assessment of a new food additive might conclude with a statement like "the safe daily dose is 0.1 mg / kg body weight +/- 0.5 mg." See Chapter 15 for more.

risk manager

People who regulate risky substances, devices, actions, and behaviors. They make laws (etc.) in response to risk assessment data; other inputs also contribute to their decision making, including politics and ideology. See Chapter 15 for more.

rocks

Defined by geologists as materials that are naturally occurring, solid, consolidated (i.e., grains glued together through some mechanism), and made up of some combination of inorganic minerals and remains of organisms. More about the three types of rocks and the rock cycle can be found in Chapter 3.

routes of exposure

In toxicology, the possible ways a xenobiotic can enter an organism, including via diet, dermal (skin) contact, inhalation, and injection. See Chapter 15 for details.

runoff

A pathway of the hydrologic cycle. Water flows across the surface of the Earth; it moves with gravity and follows the contours of the surface. See Chapter 4 for more.

safety

The likelihood that a substance or activity will not cause harm. It can be understood to be the inverse of risk. See Chapter 15 for more.

salinity

Refers to the salt content of water.

saltwater intrusion

A type of water pollution that affects coastal groundwater reservoirs. Overdraft of coastal aquifers can draw saltwater from a nearby ocean into an aquifer and well. See Chapter 11, especially Figure 11.11, for more.

sand dune succession

An example of ecological succession that transforms a barren sand dune into a small forest. Among other changes that occur with time: the dune is stabilized by grasses specifically adapted to grow in unstable sand. See Chapter 5 for details.

sanitary landfill

A widely used strategy to manage municipal solid waste (MSW). Large structures are excavated and divided into cells. Each cell is lined with impermeable plastic before trash is packed into it. Once full, a cell is sealed all around and the waste is contained indefinitely. See Chapter 13, especially Figures 13.7 and 13.8, for more.

saturated zone

A layer that can be found beneath the Earth's surface; it is below the water table. It consists of solids and pores, and the pores are completely filled with water. Groundwater is found in the saturated zone. See Chapter 4 for details.

scientific method

A systematic strategy used by scientists to answer questions about the natural world. It is designed to facilitate objectivity and repeatability. The specific steps used by scientists different somewhat among disciplines, but they generally include some variation on: observation, hypothesis, experimentation, conclusion, repeat / refine hypothesis, share information with the scientific community. Chapter 2 contains a detailed description.

second law of thermodynamics

In simple terms, describes how energy conversions must always go from higher to lower grade, i.e., conversions always lead to degradation of energy. See Chapter 4 for more.

secondary energy sources

Unlike primary energy sources, these are not sources of energy but instead can be used to carry power from a primary source to a user. Electricity is a good example of a secondary energy source. See Chapter 10 for more.

secondary pollutant

A product of reactions involving primary pollutants. For example, ozone gas, an important secondary air pollutant, can be produced from reactions involving hydrocarbons released from a car's tailpipe. See Chapter 14 for more.

secondary succession

In ecology, the recovery of an ecosystem after a relatively minor disturbance. Some remnants of previous communities are present. Compare to primary succession. See Chapter 5 for more.

secondary treatment

An important step in the sewage treatment process, it involves the use of microorganisms to digest some of the decomposable organic pollutants in sewage. See Chapter 11 for more.

sector

In the context of environmental science, refers to a major group of activities that are related to each other. Important sectors are industrial, agricultural, and domestic. For example, we can speak of water use by sector (as in Chapters 9 and 11).

secure landfills

Places at which hazardous waste is contained. See Chapter 13 for more.

sedimentary rocks

On geology, one of the three rocks types. These are formed from weathering products of existing rocks. See Chapter 3 for more.

selective toxicity

A fundamental concept in the science of toxicology. Different species, and even different individuals from the same species, will respond differently to the same poisonous substance. See Chapter 15 for more details.

sensitivity

Refers to the smallest change an instrument can detect; put another way, the smallest possible measurement that can be made by an instrument. Contrast with accuracy and precision, and see more details in Chapter 2.

septic system

One sewage treatment strategy used on a small scale. Typically, sewage from a single house is drained to a septic tank and then leach field as described in Chapter 11 (especially Figure 11.12).

sewage

Refers to the combination of biological, physical, and chemical waste that is flushed down toilets and washed down drains.

Sewage is a major source of water pollution and can be cleaned up by sewage treatment. See Chapter 11 for more.

sewage outfall

A pipe from which treated sewage is released into a natural reservoir of the hydrologic cycle. See Chapter 11 for more.

sewage treatment

A general term that refers to clean up of water polluted with human excrement and other products flushed down toilets and other drains. Treatment involves a complex series of steps to minimize the presence of chemical, physical, and biological pollutants in water before it is released back to natural systems. See Chapter 11 for details.

sewage treatment plant

A large-scale approach to the clean up of human sewage. Sewage treatment plants are typically operated by counties, cities, or other governmental organizations. See Chapter 11, especially Figure 11.13, for more.

sexual reproduction

Reproduction in which two separate individuals contribute genetic material to produce offspring. In simple terms, two parents make one new individual. Most animals employ this strategy. Contrast with asexual reproduction. See Chapter 6 for more.

sick-building syndrome

An indoor air pollution phenomenon. Many people living or working in the same building become stricken with illnesses (often respiratory conditions) that are difficult to trace to a source. Materials in their common space are often to blame, although establishing cause and effect can be difficult. See Chapter 14 for more.

side effect

Unintended, and often harmful, consequence of ingestion of a therapeutic drug. See Chapter 15 for more.

sinkhole

A depression or opening in the Earth's surface resulting from a lack of underlying support. A common cause of sinkholes is dissolution of limestone beneath the surface. See Chapter 14, especially Box 14.2, for more.

smog

A local air pollution problem generally seen in urban areas. Airborne products of fossil fuel combustion can get trapped near their sources and profoundly degrade air quality. See Chapter 14 for details.

soil

Natural materials that are produced from chemical, biological, and physical transformations of rocks and organisms. True soil consists of layers or soil horizons. Soils consist of inorganic and organic materials + air and water. See Chapter 9 for details.

soil horizons

The horizontal layers that extend from Earth's surface downward for a few meters (it varies). Each layer is physically and chemically distinct from the other layers. See Chapter 9 for details.

soil organic matter, SOM

An important component of soil. It is a complex material made up of the partially decomposed remains of organisms. It supplies nutrients to plants and helps retain moisture and soil structure. See Chapter 9 for details.

soil profile

The layers in a true soil when viewed in cross section. See Chapter 9 for more.

soil texture

Refers to the relative proportion of sand, silt, and clay in a given soil sample. See Chapter 9 for more.

solar energy

High-grade energy emitted from the sun. It provides nearly all of Earth's energy. See Chapters 4 and 10 for more.

solar energy, direct

Refers to ways humans use the sun to generate electricity or to heat water or spaces. Distinct from indirect usage of solar energy such as photosynthesis. See Chapter 10 for more.

solar farms

Areas at which many solar panels are concentrated; electricity generated at such a site could be distributed to multiple users. See Chapter 10 for more.

specialist

In ecology, an organism that occupies a narrow niche. It has strict requirements that can only be met in a small number of ways. Contrast with generalist. See Chapter 6 for more.

speciation

Broadly defined, the production of a new species. See Chapter 6 for more.

species

Refers to a group of organisms that are closely related enough that they can breed and produce fertile offspring. See Chapter 5 for more details.

species evenness

A way to quantify how many different species dominate in an area. Unlike species richness, this property reflects the relative importance of different species, not the total number of different species. For example, in a hypothetical ecosystem there might be 10,000 different species (high richness) but low evenness because only 1 of those species dominates the area--the other 9,999 are essentially invisible and inactive. In a second case, an ecosystem might have 10 organisms--low richness--but high evenness if the relative number of each species is about the same. See Chapter 6 for more.

species richness

The total number of different species in an area. See chapter 6 for more.

sphere

Environmental scientists divide the outer portion of Earth into four zones or units for the ease of study; each is known as a sphere. They include the atmosphere, hydrosphere, biosphere, and lithosphere. See Chapter 4 for more.

spontaneous generation

The name of a theory that was once used to explain how organisms could end up in food, water, and other places. In short, these organisms were thought to spring up within non-living matter without any external source or aid. It was used to account for disease and food spoilage. The theory of spontaneous generation persisted for centuries, despite many attempts to discredit it, including some that cited compelling evidence. It was finally debunked completely in the late 19th Century by Louis Pasteur and some other scientists. See Chapter 3, especially Box 3.5, for more.

standard deviation

A common strategy, but not the only way, to quantify uncertainty in a measurement. For example, 6.2 +/- 0.3 indicates the mean value from multiple readings (6.2) and the standard deviation (0.3). It can be reasonably understood to represent the range of possible values for a measurement; in the above example: 5.9 to 6.5. See also mean. See Chapter 2 for more.

start-up costs

A broad term referring to the money that must be spent to build a new facility. It is often used to describe challenges associated with power generation (Chapter 10) or waste management (Chapter 13) facilities, but it is appropriate for any new construction project.

steady state

In systems analysis, refers to a situation when the rate of input of a material to a system is equal to the rate of removal of that material from a system. It results in no net change in the amount of that material within a system. See Chapter 2 for more.

strain

A object's response to stress applied to it; strain can be understood to be the way an object changes or breaks. See Chapter 7 for more.

stratosphere

The layer of the atmosphere that lies above the troposphere. Here is where the bulk of atmospheric ozone is located. See Chapters 4 and 14 for more.

strength

The ability an object has to withstand stress applied to it without exhibiting strain. See Chapter 7 for more.

stress

The force applied to an object, as in tectonic stress applied to rocks. See Chapter 7 for more.

strip mining

One strategy to extract coal from the Earth. In short, rocks and soil on top of a coal seam are removed so the coal can be extracted at the surface. See Chapter 10 for more.

subatomic particles

A term that refers to objects that are smaller than and associated with atoms. These include neutrons, protons, and electrons. See Chapter 4 for more.

subduction

In geology, an interaction between lithospheric plates: the more dense one dives below the the less dense one. See Chapter 3 for more.

subduction zones

In geology, a convergent plate boundary at which more dense crust slides below less dense crust. See Chapter 3 for more.

subjective

An approach that is characterized by values and biases; contrast with objective. See Chapter 1 for more.

sulfurous smog

Sometimes called London smog. A local air pollution phenomenon caused by the specific conditions in London, England (and other places): cool damp conditions, persistent cloud cover, and high density of manufacturing, cars, and trucks (which release sulfur oxides and other primary pollutants). See Chapter 14 for more.

sunspot

A visible dark patch on the surface of the sun. The number of sunspots goes up and down in a regular 11-year cycle; there is a positive correlation between sunspot number and the amount of energy released by the sun. See Chapter 14 for more.

sustainability

Related to the way current usage of a resource affects future availability of that resource. Sustainable practices ensure that materials are not used up faster than they are replaced. See Chapter 1 for more.

symbiosis

In ecology, a general term referring to two or more organisms living together. Symbiotic and symbiotically are related words. See Chapter 5 for more.

system

Formally defined as a collection of subunits that work together. A system responds to inputs and produces outputs. See Chapter 2 for more.

systematic errors

In science, cause all measurements of a particular phenomenon to be inaccurate in the same direction (i.e., either too high or too low). For example, if a balance is poorly calibrated, it could consistently read 0.2 g too high for all objects assessed. Contributes to scientific uncertainty. Compare to random errors. See Chapter 2 for details.

target system

In toxicology, the part of the body affected by a xenobiotic. See Chapter 15 for details.

tectonic plates

A geology term referring to distinct units of the lithosphere. Several of these large plates make up the lithosphere, and they move around, on top of the asthenosphere, with respect to each other. Their movements are responsible for many important phenomena, including continental drift, earthquakes, and volcanoes. See Chapter 3 for more.

teratogen

In toxicology, a xenobiotic that induces birth defects. See Chapter 15 for details.

terrestrial

An adjective that refers to organisms that live primarily on land (i.e., not in water). Compare to aquatic organisms. See Chapter 4 for more.

thalidomide

A very important historical teratogen. It was used to treat morning sickness in the late 1950s and early 1960s but caused severe birth defects. See Chapter 15 for details.

The Montreal Protocol

An international agreement designed to eliminate the release of CFCs into the atmosphere because of the way they contribute to ozone destruction. It was initiated in 1987 by a small number of countries (including the U.S.) and signed by many more nations in subsequent years. See Chapter 14 for more.

theory

A scientific notion that has developed after many experiments; often, several generations of scientists have tested and refined such a notion before it is elevated to the status of theory. It is reasonable to view a theory as a conclusion that has withstood many attempts to discredit it. Note that "theory" is used differently in science than it is in everyday life. In the former, it is taken very seriously, whereas in the latter it indicates an idea that has little or no evidence to support it. Even well-supported theories could be modified with continued study, but a lot of new data would need to come to light in order to discredit one of them. Compare to conclusion and law. See Chapter 2 for more.

therapeutic

An adjectival word that refers to a substance that provides a benefit to the person who ingests it. Generally, therapeutic substances are assumed to be medically necessary. Compare to recreational. See Chapter 15 for more.

thermal expansion

Refers to the way water takes up more space as its temperature goes up (at least above about 4 degrees C). See Chapter 14 for more.

thermal pollution

When non-consumptive off-stream water use introduces heat into a stream. See Chapter 11 for more.

thermodynamics, laws of

Describe how energy is used and converted among forms (e.g., potential, kinetic, and heat). See Chapter 4 for details.

Three Mile Island

A nuclear power plant near Harrisburg, PA (USA). An serious accident there in 1979 changed the way Americans viewed nuclear power. See Chapter 1 for more.

threshold

In toxicology, the dose of a xenobiotic below which no response is observable. See Chapter 15 for details.

tight gas

Reservoirs of oil and natural gas that are locked in the pores of shales and other hard-to-access geologic structures are said to be "tight". They can only be recovered using hydraulic fracturing or other unconventional method. See Chapter 10 for more.

tight petroleum

A term that refers to oil and natural gas that is difficult to extract from their reservoirs because it is tied up in solid rocks or sands. See Chapter 10 for more.

tipping point

An informal term that refers to rapid, non-linear acceleration of change once a certain level of some property is reached. In the case of global warming and climate change, a tipping point could be reached at a particular concentration of carbon dioxide in the atmosphere or average global temperature. See Chapter 14 for more.

topsoil

The layer, or soil horizon, that contains the most soil organic matter (SOM). It is is at or close to the surface. This is the most fertile part of the soil profile and is where plants grow; when it is lost, the remaining horizons cannot support plants, including crops. See Chapter 9 for details.

total fertility rate, TFR

In demography, the number of live births per woman during childbearing years; it is generally expressed as a number such as 2.3 or 5.6. See Chapter 8 for more.

toxicology

The science of poisons. See Chapter 15 for more.

toxin

A poisonous substance produced by an organism. See Chapter 15 for details.

transform plate boundary

In geology, where two lithospheric plates slide side by side. See Chapter 3 for more.

translocation

Movement of material from one place to another; often used to refer to movement among reservoirs. See Chapters 1 and 4 for more.

transpiration

A pathway of the hydrologic cycle. Water is drawn into plants via their roots. The water moves upwards and is released through openings in leaves as a vapor. See Chapter 4 for more.

treatment (of polluted soil, water)

A term that refers to clean up of an existing contaminated parcel of water or soil. Various strategies can be employed to clean up, or remediate, polluted materials. See Chapters 11 and 15 for more.

trophic interactions

In ecology, refers to the hierarchical feeding relationships in ecosystems. See Chapter 5 for details.

trophic level

In ecosystems, a group of organisms that all feed the same number of steps away from the original source of energy. For example, plants are all at the first trophic level because they get their energy directly from the sun, whereas deer and other herbivores are at the second trophic level because they get their energy from eating organisms at the first trophic level. Consult Chapter 5 for more details.

trophic transfer

Refers to the movement of an environmental pollutant from one organism to another via predator-prey interactions. See Chapter 15 for more.

troposphere

The inner-most layer of the atmosphere, this is where weather events occur and is the portion of the atmosphere that is in direct contact with Earth's surface and living things. See Chapters 4 and 14 for more.

tsunami

An ocean wave that results from the sudden displacement of a large volume of water. Earthquakes are commonly responsible, but volcanic eruptions and meteorite impacts can also cause tsunamis. See Chapter 7 for more.

ultra-violet radiation, uv

A kind of radiation emitted by the sun. It is higher energy than visible light and can damage living cells (particularly cells of the skin). See Chapter 14 for more.

uncertainty

A property of any scientific measurement, it represents the size of the error term for a measurement. Can be thought of as the range of possible answers to a question (often, from lowest to highest). May be expressed as a number, e.g., 5.0 +/- 0.5 which indicates an average value (or otherwise representative number) of 5 with a possible range of 4.5 to 5.5. It arises from random and systematic errors. See Chapter 2 for more.

unconventional petroleum recovery

Refers to methods that are more costly and complicated than conventional methods. These can recover tight petroleum, fluids that are difficult to separate from the solid materials in which they are bound. Hydraulic fracturing is one example. See Chapter 10 for more.

unicellular

An organism that consists entirely on a single cell. Bacteria are unicellular. Contrast with multicellular. See Chapter 3 for more.

uniformitarianism

An approach to studying Earth's history characterized by the phrase 'the present is the key to the past.' In other words it holds that slow, steady processes (like those active today) were responsible for the features we see on present-day Earth. It requires that Earth be ancient to be plausible. Contrast with catastrophism. See Chapter 3 for more.

unsaturated zone

A layer that can be found beneath the Earth's surface; it is above the water table. It consists of solids and pores, and the pores are filled with some combination of water and air. See Chapter 4 for details.

uplift

In geology, refers to vertical movement upward of Earth's crust. It can be caused by one of several tectonic phenomena. See Chapter 3 for more.

uptake

A general term that refers to the entry of a xenobiotic into an organism. See Chapter 15 for details.

vaccination

Refers to the medical administration of a substance--generally, some derivative or relative of an infectious agent--to induce a protective immune response. Vaccinations against polio, measles, COVID-19 and many others have prevented millions of deaths worldwide during the past century.

variable

In simple terms, something that changes. In science, it refers to a stressor that is altered during an experiment. Generally, the way changes to a variable--for example temperature or light--affect a test subject are studied. See Chapter 2 for more.

vertebrate

This is animal with a backbone (e.g., humans, snakes, cats, mice, barracuda). Compare to invertebrate.

virus

An acellular biological entity. Viruses are not alive in the same sense that cellular plants, animals, fungi, bacteria, and protozoa are alive, but they carry genetic information and can reproduce using the machinery of cells in organism they invade. Viruses are extremely small, 100-1000 times smaller than most bacteria and 1000-10,000 times smaller than most eukaryotic cells. They are important agents of disease. See Chapter 3 for more.

volatilization

A process that converts a substance in the liquid phase into a gas. See Chapter 15 for more.

waste

A rather broad term that refers to materials unwanted by the organism or system that generates them. Individuals produce biological waste products that are partially decomposed remains of food. Such leftovers are toxic to the organism expelling them but can serve as nutrients for other organisms. The term also is applied to human systems and includes garbage from homes plus a range of other materials from industry, medical facilities, military installations, and farms. See Chapter 13 for more.

waste-to-energy

A broad term that refers to the generation of electricity through the utilization of products of human waste management. See Chapter 13 for more.

wastewater

Water contaminated by human activities (e.g., agriculture, industry, domestic use). Literally, water that contains waste products. It can be treated to remove the pollutants that render it unfit for use. See Chapter 11 for more.

water quality

Essentially, a quantifiable property of water that refers to how clean it is. Generally, how clean is related to how much material (biological, physical, chemical) is present and the intended use of the water. See Chapter 11 for more.

water table

The upper limit of the saturated zone. See Chapter 4 for details.

water-borne disease

A transmissible illness caused by microorganisms that live in contaminated water. See Chapter 7 for more.

water-holding capacity

A soil science term that refers to the amount of water retained in soil after inputs (i.e., precipitation) ceases. Water held in soil is vital to the survival of soil organisms, including crops, because it can meet demand until the next rainfall event. See Chapter 9 for more.

weather

Short-term temperature, precipitation, wind, and humidity conditions that affect an area. Weather changes daily and can fluctuate wildly. Compare to climate. See Chapter 14 for more.

weathering

In geology, a term that refers to chemical and physical processes that break down and transform rocks and the remains of living materials. Weathering products can be transported and incorporated into new rocks and soils. Compare to erosion. See Chapter 3 for more.

wind farm

An area in which many wind turbines are concentrated. See Chapter 10 for more.

wind turbine

A structure with spinning blades that captures the wind's energy and converts it into electricity. See Chapter 10, especially Figure 10.33, for more.

working face

The relatively small portion of a landfill that actively receives trash. See Chapter 13 for more.

xenobiotic

A substance that originates outside the body, that is, it is not produced by the body but from some source external to the body. See Chapter 15 for details.

zooplankton

Plankton that are consumers or chemoheterotrophs; they eat phytoplankton so live in surface waters. They tend to be several millimeters to a centimeter in size. See Chapter 5 for more.