Introduction to Oceanography

Introduction to Oceanography

PAUL WEBB

Cover image: shallow lagoon waters in Bora Bora, French Polynesia, photo by Paul Webb CC-BY.

Contents

Preface ix

Acknowledgements x

Chapter 1: Introduction to the Oceans

Chapter 1: Introduction to the Oceans 3

1.1 Overview of the Oceans 4

1.2 Continental Margins 7

*Portions modified from "Physical Geology" by Steven Earle**

1.3 Marine Provinces 10

1.4 Mapping the Seafloor 12

Chapter 2: Getting our Bearings

Chapter 2: Getting our Bearings 19

2.1 Latitude and Longitude 21

2.2 Measuring Speed 24

2.3 Map Projections 25

Chapter 3: The Origin and Structure of Earth

Chapter 3: The Origin and Structure of Earth 31

3.1 Origin of Earth and the Solar System 32

*Modified from Karla Panchuk in "Physical Geology" by Steven Earle**

3.2 Structure of Earth 37

*Portions modified from "Physical Geology" by Steven Earle**

3.3 Determining the Structure of Earth 42

*Modified from "Physical Geology" by Steven Earle**

Chapter 4: Plate Tectonics and Marine Geology

Chapter 4: Plate Tectonics and Marine Geology 51

*Modified from "Physical Geology" by Steven Earle**

4.1 Alfred Wegener and the Theory of Plate Tectonics 52

*Modified from "Physical Geology" by Steven Earle**

4.2 Paleomagnetic Evidence for Plate Tectonics 57

*Modified from "Physical Geology" by Steven Earle**

[4.3 Mechanisms for Plate Motion](#) 60
*Modified from "Physical Geology" by Steven Earle**
4.4 Plates and Plate Motions 63
*Modified from "Physical Geology" by Steven Earle**
4.5 Divergent Plate Boundaries 67
*Modified from "Physical Geology" by Steven Earle**
4.6 Convergent Plate Boundaries 74
*Modified from "Physical Geology" by Steven Earle**
4.7 Transform Plate Boundaries 77
*Modified from "Physical Geology" by Steven Earle**
4.8 Earthquakes and Plate Tectonics 79
*Modified from "Physical Geology" by Steven Earle**
4.9 Seamounts and Hot Spots 83
*Modified from "Physical Geology" by Steven Earle**
4.10 Coral Reefs 85
*Modified from "Physical Geology" by Steven Earle**
4.11 Hydrothermal Vents 87

Chapter 5: Chemical Oceanography

Chapter 5: Chemical Oceanography 93
5.1 Properties of Water 94
5.2 Origin of the Oceans 100
*Portions modified from Karla Panchuk in "Physical Geology" by Steven Earle**
5.3 Salinity Patterns 104
5.4 Dissolved Gases: Oxygen 112
5.5 Dissolved Gases: Carbon Dioxide, pH, and Ocean Acidification 115
5.6 Nitrogen and Nutrients 121
5.7 Classifying Elements in Seawater 124

Chapter 6: Physical Oceanography

Chapter 6: Physical Oceanography 127
6.1 Pressure 128
6.2 Temperature 130
6.3 Density 135
6.4 Sound 138
6.5 Light 142

Chapter 7: Primary Production

Chapter 7: Primary Production 147

7.1 Primary Production 148
7.2 The Producers 149
7.3 Factors Influencing Production 154
7.4 Patterns of Primary Production 157

Chapter 8: Oceans and Climate

Chapter 8: Oceans and Climate 163
8.1 Earth’s Heat Budget 165
8.2 Winds and the Coriolis Effect 169
8.3 Winds and Climate 175
8.4 Hurricanes 179
8.5 Climate Change 184
*Modified from "Physical Geology" by Steven Earle**

Chapter 9: Ocean Circulation

Chapter 9: Ocean Circulation 195
9.1 Surface Gyres 197
9.2 The Gulf Stream 200
9.3 The Ekman Spiral and Geostrophic Flow 204
9.4 Western Intensification 207
9.5 Currents, Upwelling and Downwelling 210
9.6 El Niño and La Niña 213
9.7 Langmuir Circulation 218
9.8 Thermohaline Circulation 220

Chapter 10: Waves

Chapter 10: Waves 231
10.1 Wave Basics 233
10.2 Waves at Sea 237
10.3 Waves on the Shore 246
10.4 Tsunamis 250

Chapter 11: Tides

Chapter 11: Tides 255
11.1 Tidal Forces 256
11.2 Dynamic Theory of Tides 262
11.3 Tide Classification 266

Chapter 12: Ocean Sediments

Chapter 12: Ocean Sediments 273
12.1 Classifying Sediments 274
12.2 Lithogenous Sediments 279
12.3 Biogenous Sediments 283
12.4 Hydrogenous Sediments 288
*Methane hydrate section modified from "Physical Geology" by Steven Earle**
12.5 Cosmogenous Sediments 293
12.6 Sediment Distribution 294

Chapter 13: Coastal Oceanography

Chapter 13: Coastal Oceanography 301
13.1 Beaches 302
13.2 Longshore Transport 305
*Modified from "Physical Geology" by Steven Earle**
13.3 Landforms of Coastal Erosion 309
*Modified from "Physical Geology" by Steven Earle**
13.4 Landforms of Coastal Deposition 314
*Modified from "Physical Geology" by Steven Earle**
13.5 Human Interference with Shorelines 319
13.6 Estuaries 323
13.7 Sea Level Change 330
*Modified from "Physical Geology" by Steven Earle**

Chapter 14: Ice

Chapter 14: Ice 335
14.1 Types of Ice 337
14.2 Icebergs 343
14.3 Ice and Climate Change 347

Glossary 351
For each entry the chapter in which the word first appears is shown in parentheses
About the Author 383

Preface

This book was originally written for my *Principles of Oceanography* course at Roger Williams University, a lower-level introductory course required for marine biology and environmental science majors. By design, this course does not go into great detail about marine biological topics, as our students will cover those topics in their other courses. For that reason, this book does not currently include sections on marine ecology, marine communities, or the diversity of marine life that are often found in other introductory texts. However, this book remains a work in progress, and it is hoped that over time, those sections may be added. I invite instructors who utilize this text to send suggestions, edits, updates, or sections you would like to see added to pwebb@rwu.edu. I especially encourage instructors who are willing to author additional sections in their areas of expertise to submit them for inclusion in future updates to this text.

Acknowledgements

This book arose through a fellowship from the Open Educational Resources (OER) Faculty Fellows Program at Roger Williams University. I am grateful to the faculty and staff of the Center for Teaching & Learning, the University Library, and the Instructional Design Office for their support and assistance throughout the fellowship, in particular Bob Shea, Linda Beith, and Lindsey Gumb.

I would like to extend my thanks to Steven Earle, whose book *Physical Geology* I have drawn from heavily in the geological sections of this text. Geology is not my area of expertise, so I am appreciative of the work of Dr. Earle in both creating his book and in making it available for others to use and adapt through the OER and Creative Commons ideals. I also thank the numerous other people cited throughout the text who have made their resources available through Creative Commons licensing. I hope that others will be able to similarly make use of the materials I have created here.

Thank you to Julio Romero for creating many of the figures throughout the book. Artwork credited to Julio is designated with (JR).

And finally, I thank my family, Alexis, Rhys and Kyle, for their support and encouragement throughout the ongoing process that is this book.

CHAPTER 1: INTRODUCTION TO THE OCEANS

Chapter 1: Introduction to the Oceans

Learning Objectives

After reading this chapter, you should:

- know the names and locations of the major oceans
- know the average depth of the ocean
- be able to identify the deepest region of the world ocean
- know the differences between passive and active continental margins
- know the geological features of a continental margin, and what is responsible for them
- know the various ocean provinces
- understand how oceanographers can map the seafloor

There are many reasons why people study oceanography. An understanding of ocean processes is obviously vital to oceanographers, marine biologists, or environmental scientists. However, there are numerous other scientific fields where the oceans play an important role. The oceans are a major contributor to global climate patterns, and can give us clues to past climate conditions. Knowledge of oceanography is important for resource or energy extraction, such as commercial fishing or aquaculture, oil and gas exploration, and clean energy resources such as wind, wave, or tidal energy. The oceans are the major route for international trade through commercial shipping, and are still a significant factor in the transportation of people across the seas. But even the recreational user benefits from an understanding of the ocean, from winds and currents for the casual sailor, tides and habitat conditions for a fisherman, to wave patterns for surfers. And finally, for anyone who has ever stood on the shore and gazed out at the ocean with a sense of wonder at what lies beneath the surface, the study of oceanography can begin to reveal some of the ocean's mysteries.

This chapter begins with a basic overview of the world oceans, before discussing some of the ways that the features of the ocean can be classified. Finally, we will examine the techniques that are used, and have been used, to map the ocean floor.

1.1 Overview of the Oceans

Let's begin by looking at a few basic facts about the oceans. We often think of Earth in terms of its land area, but in reality 71% of the Earth's surface is covered by oceans, while only 29% is land. Oceans cover an area of 139 million miles^2 or 361 million km^2, and contain a volume of about 1.37 billion km^3 of water. All of this water is not distributed equally over the Earth; 61% of the Northern Hemisphere is covered by oceans, while in the Southern Hemisphere the oceans cover 81% of the surface area (Figure 1.1.1).

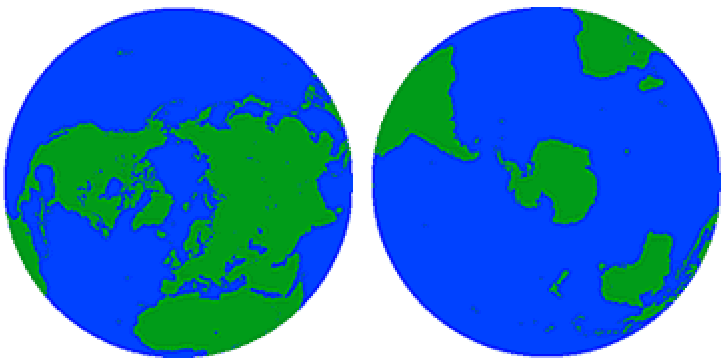


Figure 1.1.1 *Ocean cover in the Northern (left) and Southern (right) Hemispheres.*

Various sources differ in the number of recognized ocean basins. Historically the major oceans were recognized as the Pacific, Atlantic, Indian, and Arctic Oceans. More recently, the Southern Ocean has been recognized as fifth named ocean, comprising all of the water from the coast of Antarctica to 60^o S (Figure 1.1.2). In 2000 these boundaries were submitted to the International Hydrographic Organization for official recognition, but several countries do not recognize it as a separate ocean, but rather as the southern extension of the other major oceans. The Southern Ocean has its own unique characteristics, so for the purposes of this book we will include it as a separate ocean.

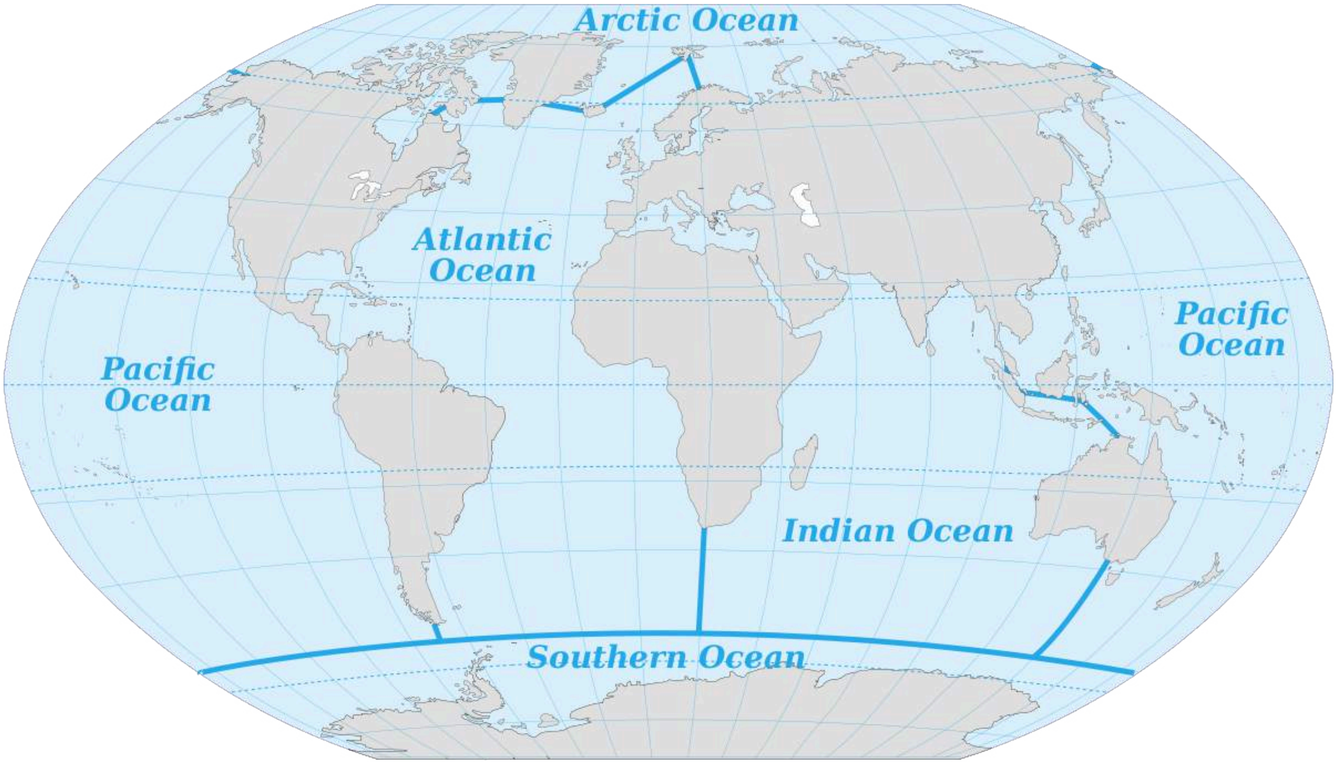


Figure 1.1.2 *Map of the world oceans (By Pinpin [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 3.0], via Wikimedia Commons).*

The oceans account for vast amounts of water, containing 97% of the water on Earth's surface, with over half of the water in the Pacific alone (Table 1.1.1).

Table 1.1.1 Percentage of Earth's water in various locations

Pacific	52%
Atlantic	25%
Indian	20%
Ice	2%
Ground water	0.6%
Atmosphere, lakes & rivers	0.01%

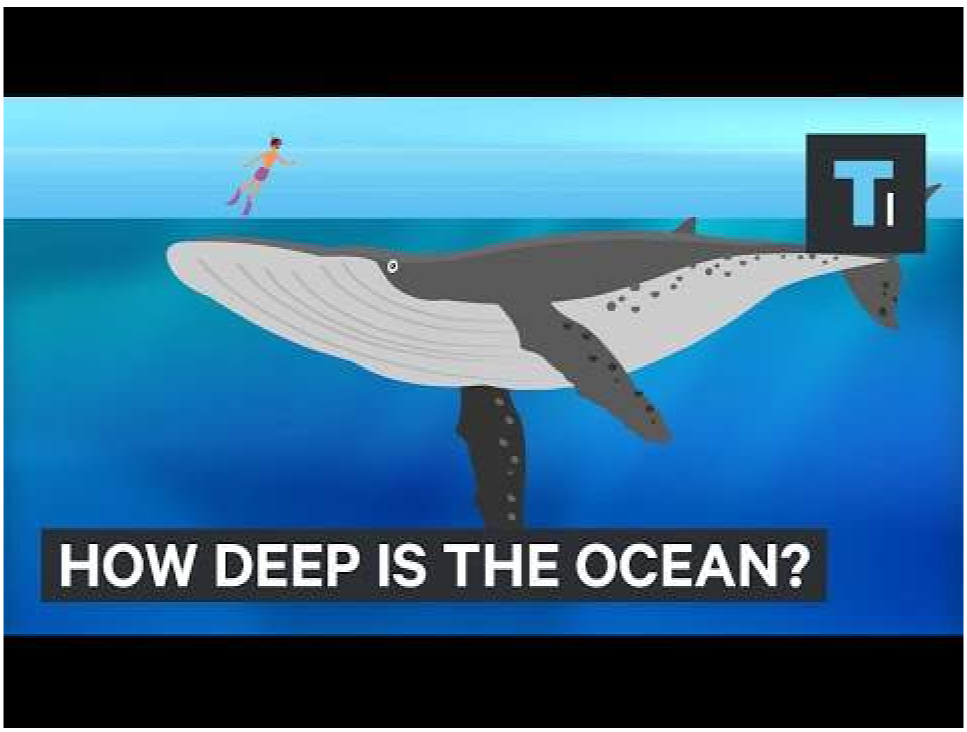
The average depth of the world ocean is about 3800m (12,500 ft), which is about four times deeper than the average land elevation is high (840m or 2800 ft). In fact Mt. Everest, the highest point on land, is 8848m (29,028 ft) high, while the deepest part of the ocean, the Challenger Deep of the Marianas Trench is 11,022m (36,200 ft) deep. So you could submerge Mt. Everest in the Marianas Trench and it would still be covered by over 2 km of water! Because there is so much more water on Earth than there is land, if you could smooth out the land elevation the entire Earth would still be covered by water about 2700 m deep.

Of the major ocean basins, the Pacific is the largest (almost as large as all of the others combined), and is the deepest (Table 1.1.2).

Table 1.1.2 Area and depth of the major oceans

	Area (million km^2)	Average depth (m)
Pacific	166	4282
Atlantic	87	3926
Indian	73	3963
Arctic	14	1205
Southern	20	4000

Watch the video below for some perspective on the size and depth of the oceans.



A YouTube element has been excluded from this version of the text. You can view it online here: https://rwu.pressbooks.pub/webboceanography/?p=24

1.2 Continental Margins

*Portions modified from "Physical Geology" by Steven Earle**

Continental margins refer to the region of transition from the land to the deep seafloor, i.e. between continental and oceanic crust. In an **active continental margin**, the boundary between the continent and the ocean is also a tectonic plate boundary, so there is a lot of geological activity around the margin. The west coast of the United States is an example of an active margin, where the coastline corresponds with the boundary between the Pacific and North America Plates. A **passive continental margin** occurs where the transition from land to sea is not associated with a plate boundary. The east coast of the United States is a good example; the plate boundary is located along the mid Atlantic ridge, far from the coast. Passive margins are less geologically active. Figure 1.2.1 shows an idealized passive margin. When examining this figure, and others like it, note that there is significant vertical exaggeration; the depth scale covers approximately 5000 m, while the horizontal scale extends around 300 km. This makes the features look much steeper than they actually are. The bar at the bottom of Figure 1.2.1 shows what a passive margin would look like without this exaggeration; there is a much more gradual transition to depth.

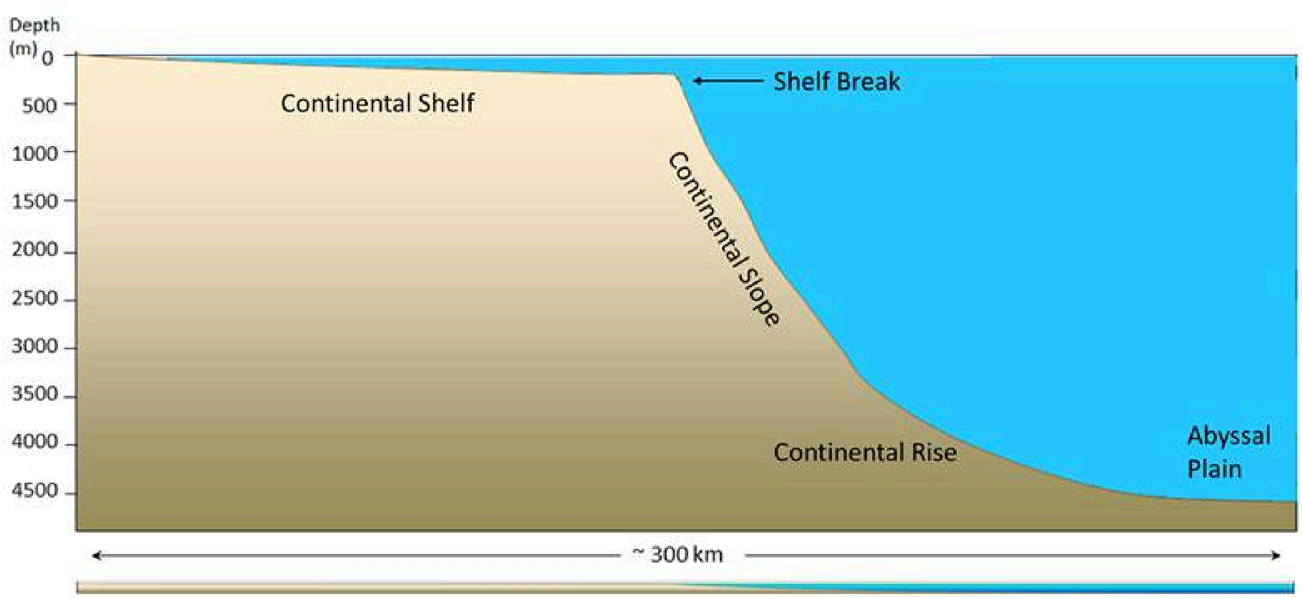


Figure 1.2.1 *Features of a passive continental margin (modified by PW from Steven Earle, "Physical Geology").*

The **continental shelf** is the shallow, flooded edge of the continent. Geologically the shelf is still part of the continental crust, but it is often overlaid with marine sediments. On average, the shelf extends about 80 km from the coast; some margins have very little shelf, while the Siberian Shelf in the Arctic extends roughly 1500 km. The depth of the shelf generally remains below about 150 m, and the floor of the shelf is fairly flat. The flat topography is the result of changes in sea level; throughout history the shelves have been both submerged and exposed, and as sea level rose and fell, wave action, ice sheets, and other erosional processes smoothed out the shelf surface. Wave action and the movement of sediments over the shelf have continued this smoothing process. Continental shelves only make up about 6% of the ocean's surface area, but they are biologically one of the richest parts of the ocean; their shallow depth prevents nutrients from sinking out, and their proximity to the coast provides significant nutrient input. The continental shelf ends at the **shelf break**, which is the point where the angle of the seafloor begins to get steeper. The shelf break averages about 135 m deep.

After the shelf break, the seafloor takes on a steeper angle (about 4^{o}) as it descends to the deep ocean. This steeper portion of the margin is the **continental slope**, and it extends from the shelf break down to 3000-5000m. In some parts of the ocean, large submarine canyons have been carved into the continental slope; for example, Monterey Canyon in Monterey Bay, California, is a submarine canyon similar in size to the Grand Canyon! These canyons may be carved out by **turbidity currents**, which are essentially landslides of sediment, rocks, and other debris down the face of the slope.

At the bottom of the slope is the **continental rise**. This area represents where the continental crust meets the oceanic crust, as the slope begins to level off to become the deep ocean floor. The rise consists of a thick layer of accumulated sediment coming from the continent, so it is difficult to tell where the slope ends and the rise begins.

After the rise comes the **abyssal plain**, or the deep ocean floor, lying between 4500 – 6000 m. The abyssal plain includes most of the ocean floor, and is the flattest region on Earth. It is flat due to millions of years of sediment accumulation on the bottom, which buries many bottom features (Figure 1.2.2).

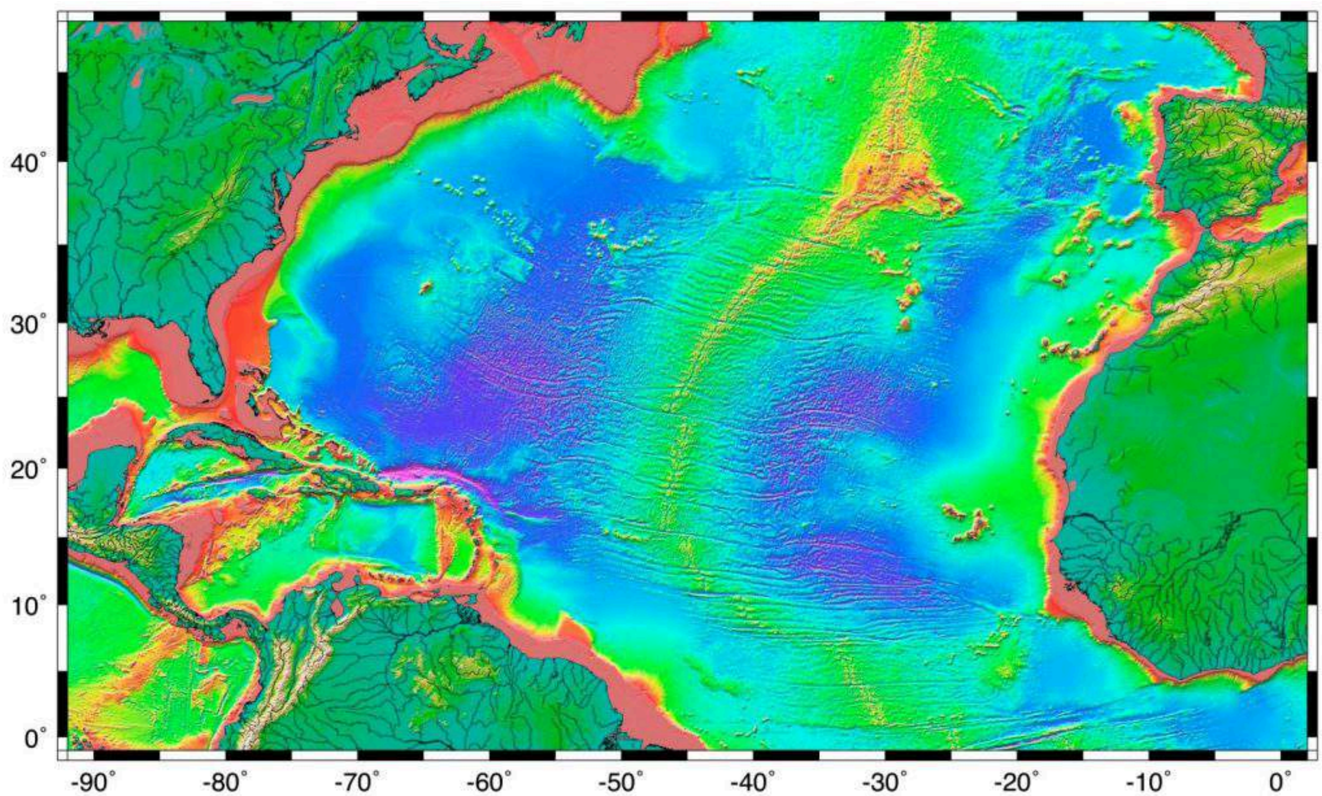


Figure 1.2.2 *Topography of the North Atlantic. Shallow continental shelf regions are shown in red, and the abyssal plain is shown in blue. Along the east coast of the United States the continental slope can be seen in green (Steven Earle, "Physical Geology").*

Passive margins, as described above, have wide shelves, gentle slopes, and a well-developed rise. Since passive margins are not plate boundaries, they experience long periods of relative stability which can lead to the development of these features. Active margins have similar features to passive margins, but the plate boundary affects the properties of the features. Active margins, like the Pacific coast of North America, have narrower shelves, steeper slopes, and little to no rise, particularly in convergent boundaries. Trenches associated with subduction zones act as sediment traps, preventing the accumulation of a continental rise, and keeping sediments off of the abyssal plains.

1.3 Marine Provinces

In section 1.2 we learned about the regions that make up the continental margins. So before we leave this topic, we will look at some of the other ways we can categorize the ocean environments.

The first major distinction is between the pelagic and benthic zones. The **pelagic** zone refers to the water column, where swimming and floating organisms live. The **benthic** zone refers to the bottom, and organisms living on and in the bottom are known as the **benthos**.

The pelagic zone is divided into two provinces: the **neritic** province corresponds to all of the water from the low tide line to the shelf break, while the **oceanic** province represents all of the other water in the open ocean regions.

The oceanic province is divided into depth zones (Figure 1.3.1):

- 0-200 m is the **epipelagic** zone ("epi" = "upon", as in on top of the pelagic zone). This is the region where enough light penetrates the water to support photosynthesis (see section 7.3), so it is also called the **euphotic** or **photic** zone.
- 200-1000 m is the **mesopelagic** zone ("meso" = "middle"). There is some light here, but not enough for photosynthesis, so it is called the **dysphotic zone**, or the **twilight zone**.
- 1000-4000 m is the **bathypelagic** zone ("bathy" = "deep"). There is no light at these depths, so it is referred to as the **aphotic** zone. About 75% of the living space in the ocean lies at these depths or deeper.
- 4000-6000 m is the **abyssopelagic** or **abyssalpelagic** zone, which extends to the seafloor in most areas.
- 6000 m and below is the **hadopelagic** or **hadalpelagic** zone (named for Hades or "hell"). This refers to the water in deep ocean trenches.

Inhabitants of these regions are referred to according to their habitat, for example mesopelagic fish, or epipelagic squid.

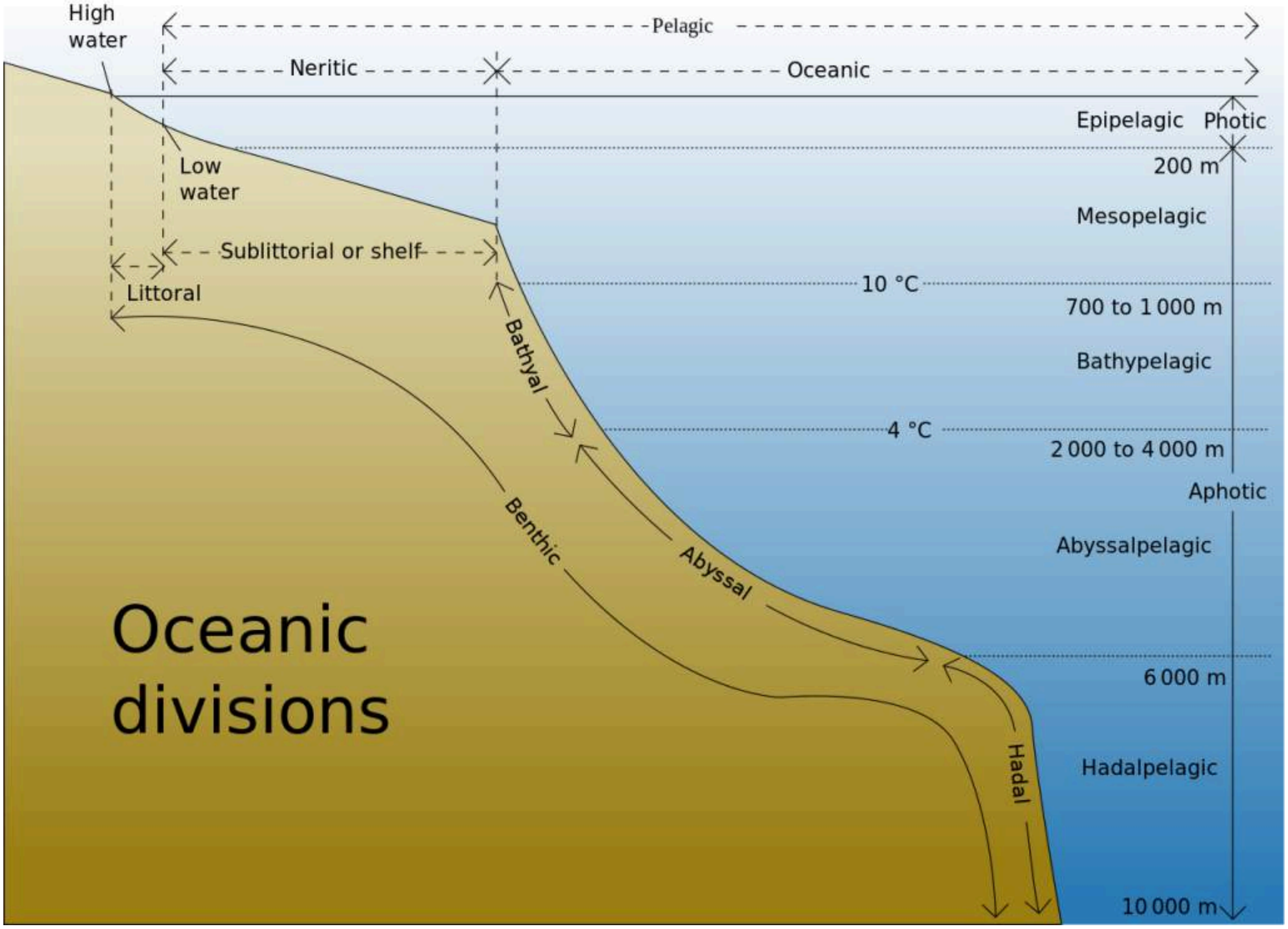


Figure 1.3.1 *The major benthic and pelagic oceanic divisions (K. Aainsqatsi at en.wikipedia [Public domain, GFDL (http://www.gnu.org/copyleft/fdl.html) or CC-BY-SA-3.0], via Wikimedia Commons).*

The benthic environment is also divided into zones, most of which correspond to the pelagic divisions:

- The **supralittoral** zone lies above the high tide line. Also called the spray zone, it is only submerged during storms or unusually high waves.
- The **littoral** zone is the region between the high and low tides. Thus it is also referred to as the **intertidal** zone.
- Below the littoral zone is the **sublittoral** (shelf) zone, extending from the low tide mark to the shelf break, essentially covering the continental shelf.
- The **bathyal** zone extends along the bottom from the shelf break to 4000m, so it generally includes the continental slope and rise.
- The **abyssal** zone is found between 4000-6000 m, including most of the abyssal plains. The abyssal zone represents about 80% of the benthic environment.
- The **hadal** zone includes all benthic regions deeper than 6000 m, such as in the bottom of trenches.

1.4 Mapping the Seafloor

The previous sections included some information on the depths of the oceans in various places. So how are we able to map the ocean floor to ascertain these vast depths?

To map the ocean floor we need to know the depth at a number of places. The process of measuring the depths is known as **bathymetry**. These measurements were first made through **soundings**, where a weighted line (lead line) was let out by hand until it touched the bottom, and the depth could be recorded from the length of the line (Figure 1.4.1). This technique led to the fathom as a unit of depth; as sailors hauled in the sounding line they would stretch it out to cover their arm span. The average arm span of a sailor was about six feet, so one fathom equals six feet, and the sailors could simply count the number of “arm spans” as they pulled in the line.

Figure 1.4.1 *Lead line survey from a catamaran hull in Alaska, 1942 (http://celebrating200years.noaa.gov/transformations/hydrography/image7.html).*

This technique had a number of drawbacks, and was usually limited to shallower water. It was very time consuming, and only gave depth data for a single point, so many individual soundings were needed to map an area. It could also be error-prone; in deep water it could be difficult to determine when the weight hit the bottom as the weight of the line itself could cause the line to keep sinking, and currents could deflect the line away from vertical, thus overestimating the depth. In later years, winches and heavy steel cables were used for deeper water, but this did not solve all of the problems inherent in the sounding method, and also added the constraint of excessive weight of the equipment.

In the 19th century, a number of modifications were made to this simple design. In 1802 the British clockmaker Edward Massey invented a mechanical device that was attached to the sounding line; as the device sank, a rotor turned a dial which locked in place when the line hit bottom (Figure 1.4.2). The line could then be reeled in and the depth read from the dial. In 1853 American sailor John Mercer Brooke developed a cannonball weight attached to a twine reel. The cannonball was dropped over the side and allowed to free-fall to the bottom; by timing the fall rate (the rate at which the twine unspooled) and noting when the rate changed as the cannonball hit the bottom, the water depth could be calculated. When it hit bottom, the cannonball was released and the line could be hauled back in, bringing with it a sample of mud in the iron bar that held the cannonball, thus confirming that the bottom had been reached.

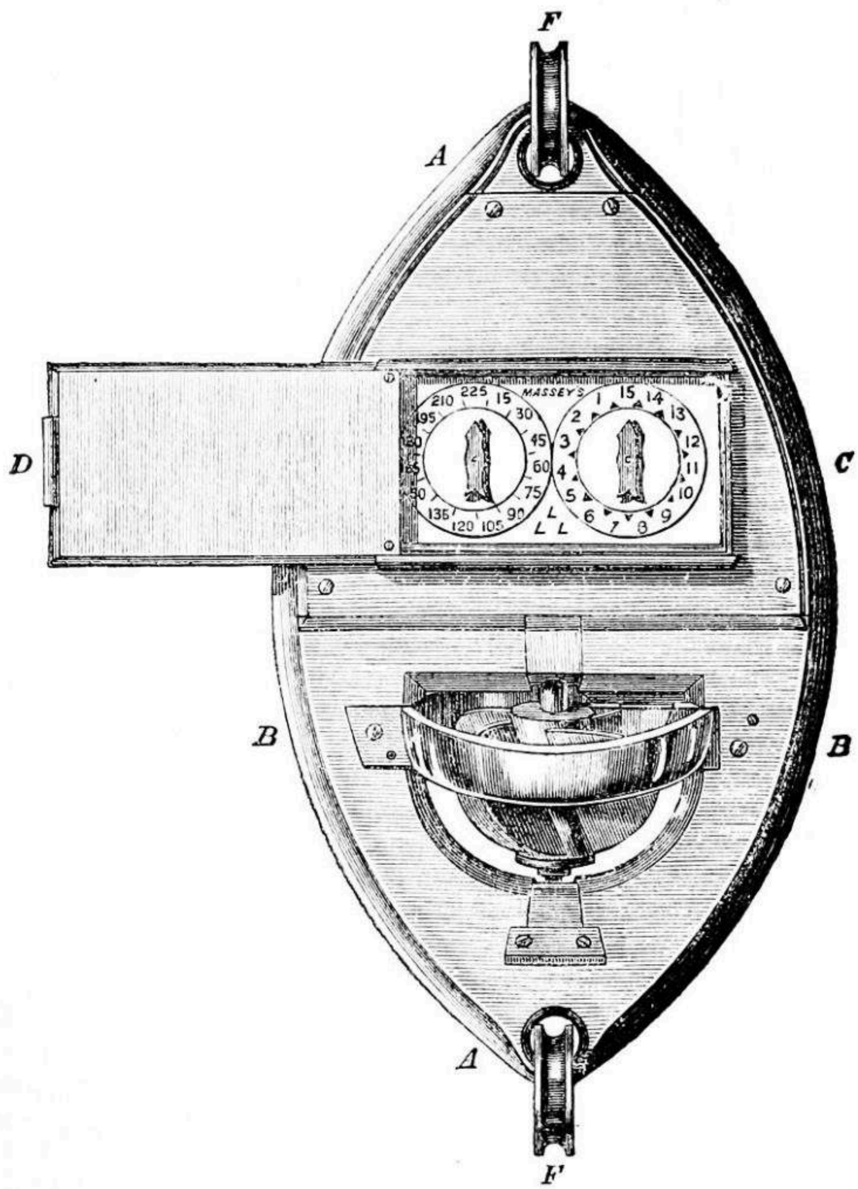


Figure 1.4.2 *Massey's sounding machine (Public domain, via Wikimedia Commons).*

After the *Titanic* disaster in 1912, there was an effort to develop better methods of detecting icebergs from a ship. This led to the development of **sonar** (SOund Navigation And Ranging) technology, which was soon applied to mapping bathymetry. A sonar device called an echosounder sends out a pulse of sound, then listens for the returning echo. The timing of the returning echo is used to calculate depth. We know that the speed of sound in water is approximately 1500 m/s (see section 6.4). Since the returning echo traveled to the bottom and back, the water depth corresponds to half the time it takes for an echo to return, multiplied by the speed of sound in water (Figure 1.4.3):

depth=1/2*(two-way travel time)*(speed of sound in water)

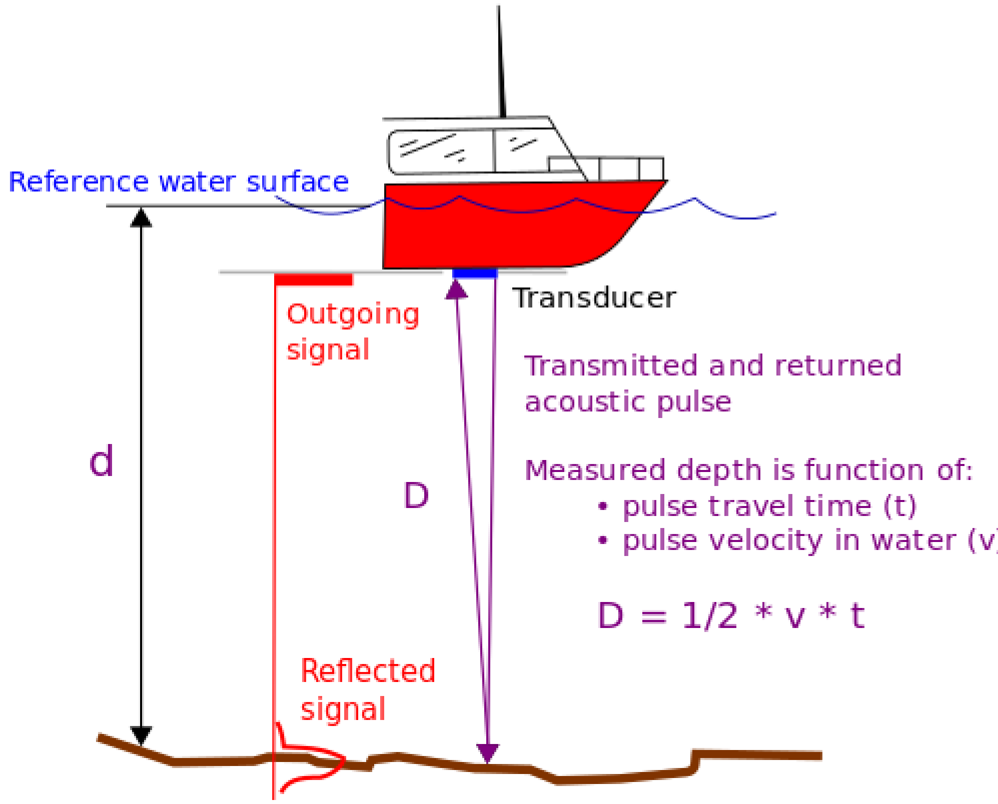


Figure 1.4.3 *Measuring depth using an echosounder (Public domain via Wikimedia Commons).*

Echosounders allowed a fast, continuous record of bathymetry under a moving ship. However, they only give the depth directly under a ship's path. Today, high resolution seafloor maps are made through multibeam or side scan sonar, either from a ship or from a towed transmitter (Fig. 1.4.4). Multibeam sonar produces a fan-shaped acoustic field allowing a much a wider area (>10 km wide) to be mapped simultaneously.

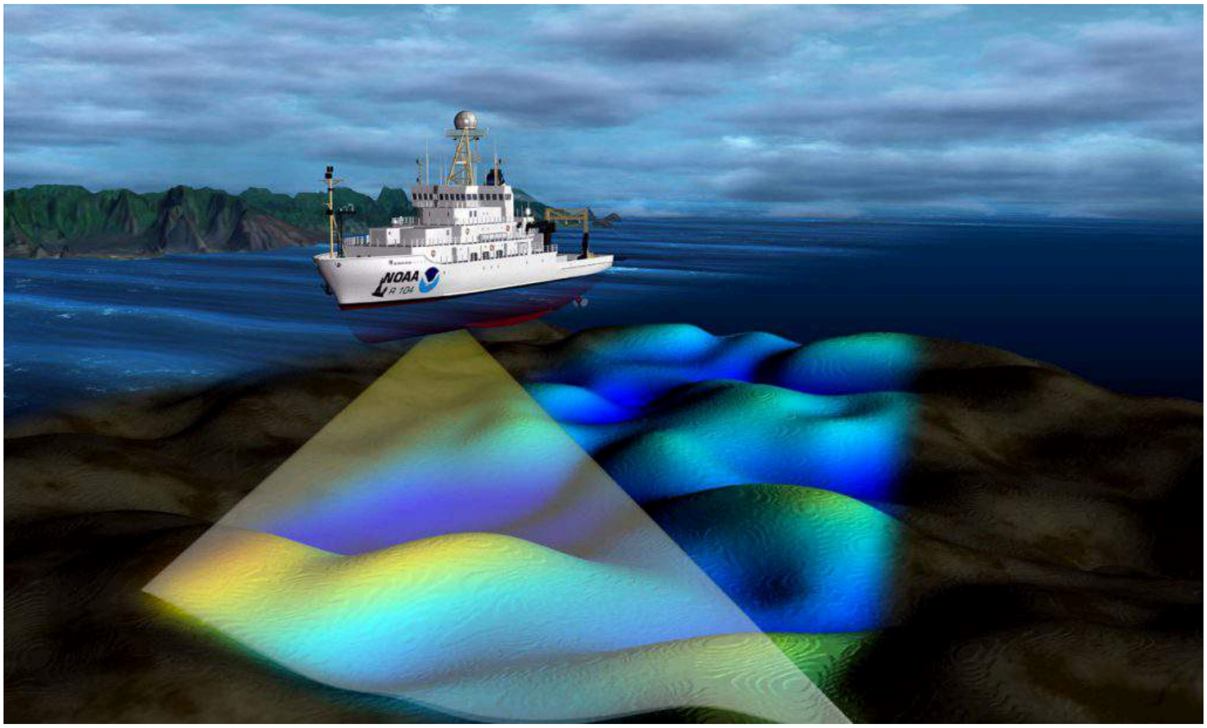
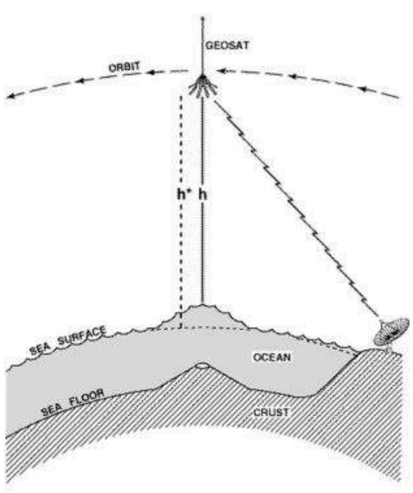


__Figure 1.4.4__ Multibeam sonar (NOAA).

Large-scale mapping of the ocean floor is also carried out by satellites (originally SEASAT, then GEOSAT, now the Jason satellites) which use radio waves to measure the height of the sea surface (radar altimetry). The sea surface is not flat; gravity causes it to be slightly higher over elevated features on the ocean floor, and slightly lower over trenches and other depressions. Satellites send out radio waves, and similar to an echosounder, can use the returning waves to detect differences in sea surface height down to 3-6 cm (Figure 1.4.5). These differences in sea surface heights allow us to determine the topography under the surface. Unlike the old lead line technology, where hundreds of soundings were necessary to map a small area, each taking an hour or more to complete, the current satellites can map over 90% of Earth's ice-free sea surface every 10 days!



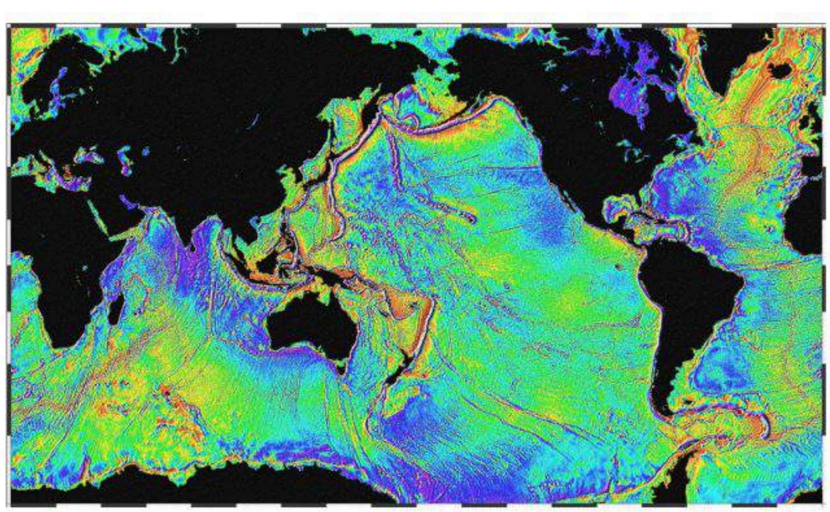


__Figure 1.4.5__ Radar altimetry (left) and a map of the seafloor produced by radar altimetry satellites (right) (NOAA).

Additional links for more information:

- NOAA page on side scan sonar: http://www.nauticalcharts.noaa.gov/hsd/SSS.html
- NOAA page on exploring the ocean using satellite altimetry data: https://www.ngdc.noaa.gov/mgg/bathymetry/predicted/explore.HTML

CHAPTER 2: GETTING OUR BEARINGS

Chapter 2: Getting our Bearings

Learning Objectives

After reading this chapter, you should be able to:

- explain the concept of latitude and longitude, and be able to make calculations based on these coordinates
- explain the advantages and limitations of the various map projections used in oceanography

People have been navigating the oceans for thousands of years, for exploration, travel, acquiring food, and transporting goods. To do so requires some form of map and the ability to tell direction. Therefore, various systems of navigation have been around for centuries.

Early Pacific Islanders used stick charts, where shells indicated islands, and bent sticks represented wave and current patterns around the islands (Figure 2.1).

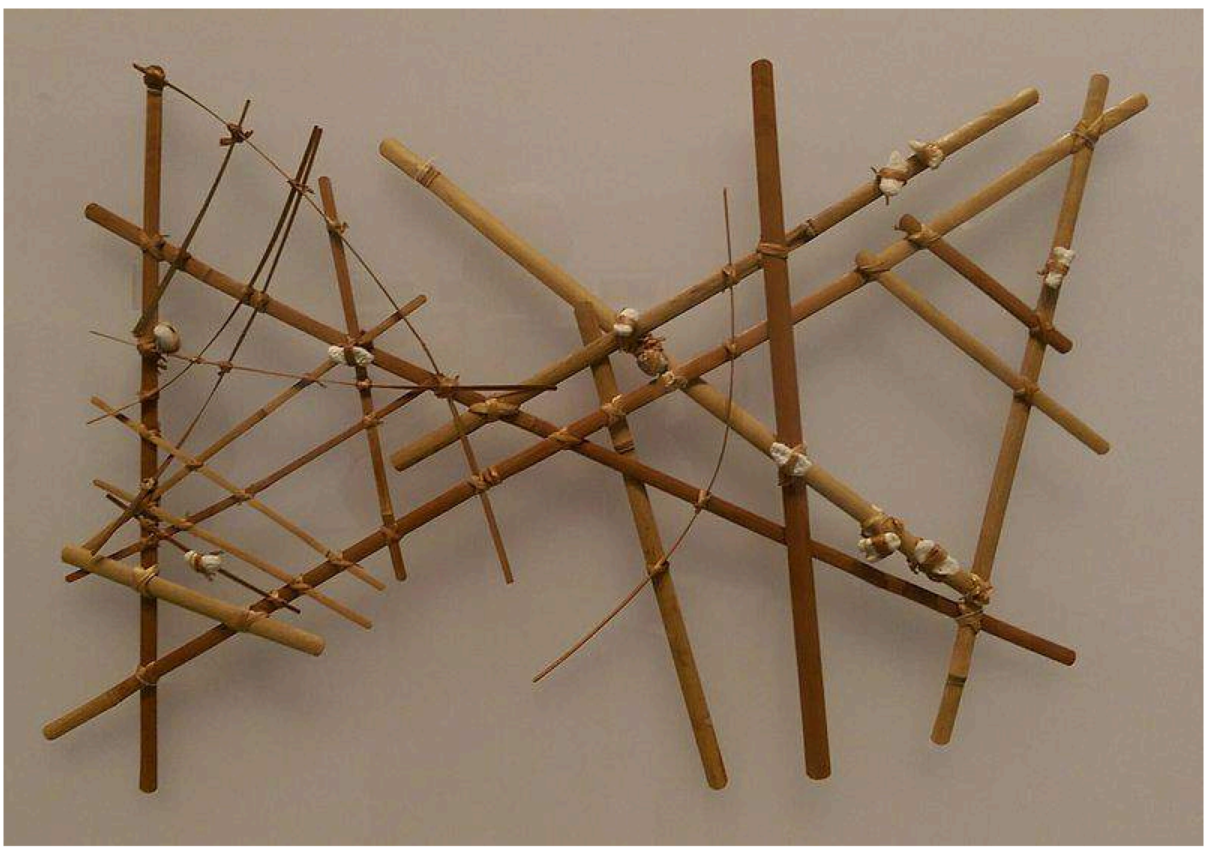


Figure 2.1 *A Micronesian navigational chart from the Marshall Islands, made of wood, sennit fiber and cowrie shells (By Cullen328 – Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=46844500).*

The first Western civilization known to have developed navigation at sea were the Phoenicians, about 4,000 years ago (c. 2000 B.C.E.). Phoenician sailors accomplished navigation by using primitive charts and observations of the sun and stars to determine directions. They explored the Mediterranean and Red Seas, and even circumnavigated Africa in 590 B.C.E.

Ptolemy was a Greek-Egyptian writer, mathematician and scientist living in Alexandria, Egypt. He produced maps of the known world in around 150 AD, including the locations of major cities, and the first known use of lines of latitude and longitude (Figure 2.2). This early coordinate system forms the basis of geolocation techniques still in use today.

Figure 2.2 *A 15th century world map based on Ptolemy's Geography (Public domain via Wikimedia Commons).*

2.1 Latitude and Longitude

Any point on Earth can be defined by the intersection of its lines of latitude and longitude. **Latitude** is measured as the angle from the equator, to the Earth's center, to your position on the Earth's surface (Figure2.1.1). It is expressed as degrees north or south of the equator (0^o), with the poles at a latitude of 90^o. Thus the poles are referred to as high latitude, while the equatorial region is considered low latitude. Lines of equal latitude are always the same distance apart, and so they are called **parallels** of latitude; they never converge. However, the parallels of latitude do get shorter as they approach the poles.

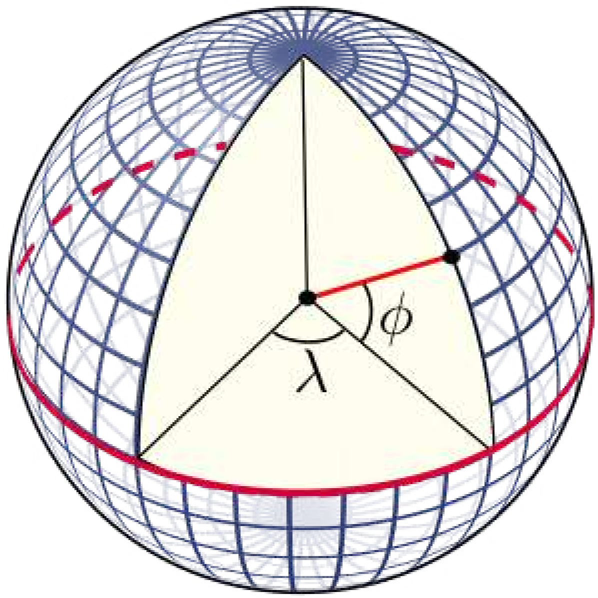


Figure 2.1.1 *The latitude of a point on the Earth's surface is determined by the angle (ø) between the point and the equator, passing through Earth's center (Peter Mercator [Public domain], via Wikimedia Commons).*

One degree of latitude is divided into 60 minutes ('). One minute of latitude equals one **nautical mile**, which is equal to 1.15 land miles (1.85 km). Each minute of latitude is further divided into 60 seconds ("). So traditionally, positions have been expressed as degrees/minutes/seconds, e.g. 36^o 15' 32" N. However, with modern digital technology, positions are increasingly expressed as decimals, such as 36^o 15.25 N', or 36.2597^o N.

In the Northern Hemisphere, latitude can be determined by the angle of the North Star (Polaris) from the horizon. The North Star always sits over the North Pole. Here, if a person looks straight ahead towards the horizon, the star would be directly overhead, creating a 90^o angle; thus the latitude at the North Pole is 90^o N. At the equator looking north, the star is in the same direction as the horizon, so the angle between them is 0^o, and thus the equatorial latitude is also 0^o. At any other point in the Northern Hemisphere, the angle between the horizon and the star will give the latitude.

Early mariners used an instrument called an astrolabe to calculate this angle. Later the sextant was developed, which allowed more accurate measurements (Fig. 2.1.2).

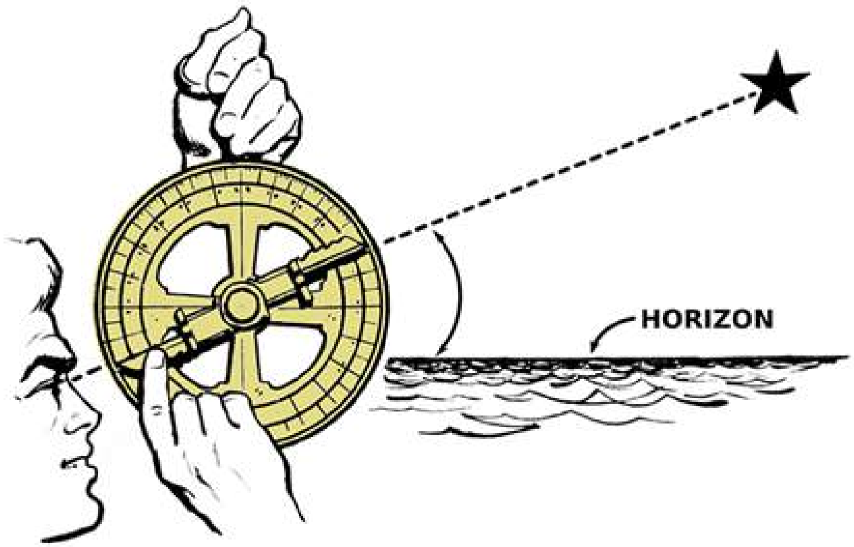


__Figure 2.1.2__ An astrolabe (left) and sextant (right) (Public domain via Wikimedia Commons).

There is no direct analogue to the North Star in the Southern Hemisphere that is useful for determining latitude. However, the Southern Cross and Centaurus constellations can be used to find the south celestial pole. If a line is drawn through the long axis of the Southern Cross, and another line is drawn between the two brightest stars in Centaurus, the two lines will intersect at the south celestial pole.

Longitude measures the distance east or west of an imaginary reference point, the **prime meridian** (0^{o}), which is now defined as the line passing through Greenwich, England (although throughout history the prime meridian has also been located in Rome, Copenhagen, Paris, Philadelphia, the Canary Islands, and Jerusalem; unlike the equator, the prime meridian's location is fairly arbitrary). Your longitude represents the angle east or west between your location, the center of the Earth, and the prime meridian (Fig. 2.1.3).

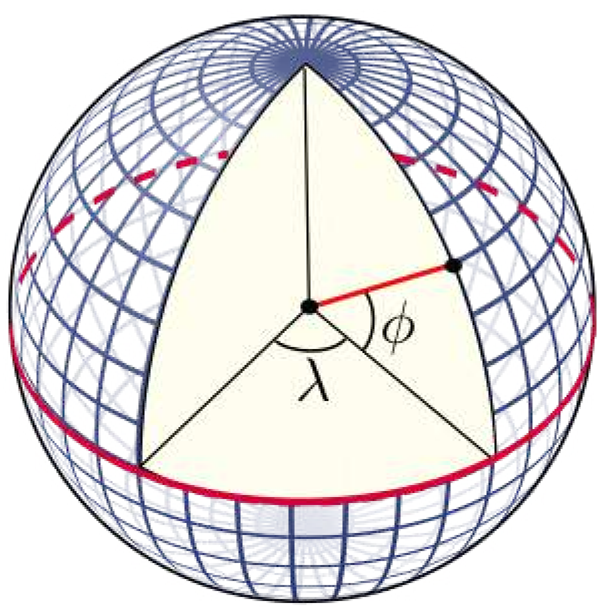


__Figure 2.1.3__ Longitude is determined as the angle (λ) between the prime meridian and your position (Peter Mercator [Public domain], via Wikimedia Commons).

As you move east and west from the prime meridian, eventually you reach 180° E and W on the opposite side of the globe from Greenwich. This point is the International Date Line. Lines of longitude are called **meridians** of longitude, or great circles. All circles of longitude are the same length, and are not parallel like lines of latitude; they converge as they near the poles. Therefore, while one minute of latitude always equals one nautical mile, the length of one minute of longitude will decline from the equator to to poles, where it will ultimately decline to zero.

Measuring longitude requires accurate time at your current location, and also the time at some distant point like a home port at the same instant. The time difference can be used to calculate longitude. This is because the Earth takes 24 hours for a complete 360° rotation. So in one hour, the Earth rotates through 1/24 of 360°, or 15°. Therefore, for each hour of time difference between two locations, there is a 15° difference in longitude.

Accurate measurements of latitude using the North Star have been made since at least the third century B.C.E. Because longitude measurements required accurate timekeeping, it wasn't until the mid-18th century that longitude was easily and precisely measured at sea. Before then, sailors would often sail north or south to get to the desired latitude, then just head east or west until they reached the target longitude. Solving the longitude problem was so important that the British government passed the Longitude Act in 1714, offering a £20,000 prize to anyone who could devise a method of measuring longitude at sea to within half a degree. Many unsuccessful solutions were proposed, including astronomical observations, but it was a clock maker, John Harrison, who developed a series of clocks that eventually satisfied the criteria. The first version (the H1) weighed over 80 lbs, but his final timepieces, the H4 and H5, could be held in the palm of one hand. Ironically, even though his clocks satisfied the criteria, Harrison was never named as the winner of the longitude prize, and in fact no winner was ever officially determined. With accurate timepieces now available, a ship could have one clock set for Greenwich time (or some other home location), and another clock set to local time, which could be reset each day by observing the sun. The time difference between the two clocks could be used to calculate longitude.

Today we use **GPS** (Global Positioning System) technology to determine latitude and longitude, and even the smallest smart phones and smart watches can use GPS to calculate position. GPS works through a system of orbiting satellites that constantly emit signals containing the time and their position. A GPS receiver receives these signals from multiple satellites, and triangulates the signals to calculate position. The system needs 24 satellites to be functional at one time; as of 2015, the system consisted of about 32 operational satellites, able to give a position with an accuracy of 9 meters (30 feet) or less.

Additional links for more information:

- For more about John Harrison and the longitude problem, see http://www.rmg.co.uk/explore/astronomy-and-time/time-facts/harrison
- For more about the workings of GPS technology, visit https://www.gps.gov/

2.2 Measuring Speed

In addition to their position and the depth of the water, it was also important for mariners to know their speed at sea. Early sailors used **chip logs**, which were planks of wood attached to long spooled lines containing knots at regular intervals. The plank was thrown overboard, while the spool remained on the ship. Once in the water, the plank encountered drag, which held the plank roughly stationary in the water while the attached line unspooled. The rate at which the line unspooled indicated how fast the ship moved away from the plank, and thus the ship's speed. The rate at which the line unspooled was measured by counting the number of knots that passed through the sailor's hands in a certain amount of time. Since it was originally measured as the number of knots per unit time, speed over water traditionally uses the unit of "knots" to indicate speed. One knot (kt) = 1 nautical mile per hour = 1.15 mph = 1.85 kph.

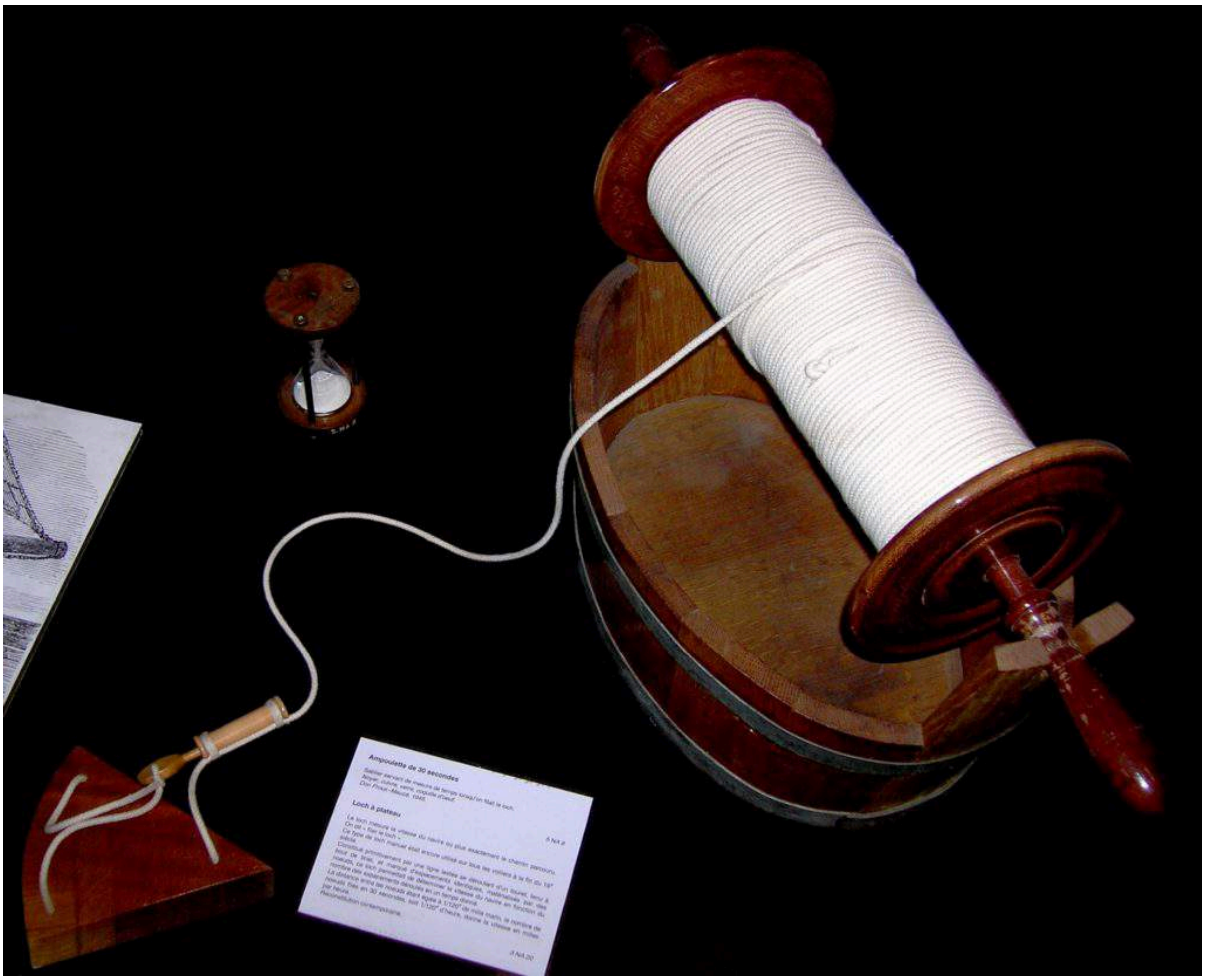


Figure 2.1.1 *Chip log (Rémi Kaupp – Personal photograph taken in the Musée de la Marine, Paris, CC BY-SA 3.0, via Wikimedia Commons).*

2.3 Map Projections

It's impossible to study oceanography without looking at maps, so it is important to recognize the strengths and weaknesses of the types of maps you might encounter. It is difficult to accurately represent a three-dimensional spherical object like the Earth on a flat, two-dimensional map or chart. Therefore, two-dimensional maps are distorted in representing the Earth's true area, direction, distance, and shape. Only a globe is accurate in all of these variables, but globes are impractical to use in the field, and impossible to reproduce in a book. Because of these limitations, we use different map **projections** to represent the Earth, depending on the needs of the presenter. Below are some of the more common map projections used in oceanography.

In a **Mercator projection**, latitude and longitude are both represented as straight, parallel lines intersecting at right angles (Figure 2.3.1). This projection is good for navigation as directions are preserved; for example, on any point on the map, north points to the top of the chart. This makes Mercator projections the standard for navigational charts. The drawback to this projection is that size and distance are distorted at high latitudes. This is because the distance between lines of longitude declines as you approach the poles, but they remain constant on a Mercator projection. The poles, which are represented by a point on a globe, are expanded to have the same circumference as the equator. This exaggerates distances, and thus area, at high latitudes. For example, South America is really nine times larger than Greenland, but on a Mercator map they appear to be the same size (Figure 2.3.1).

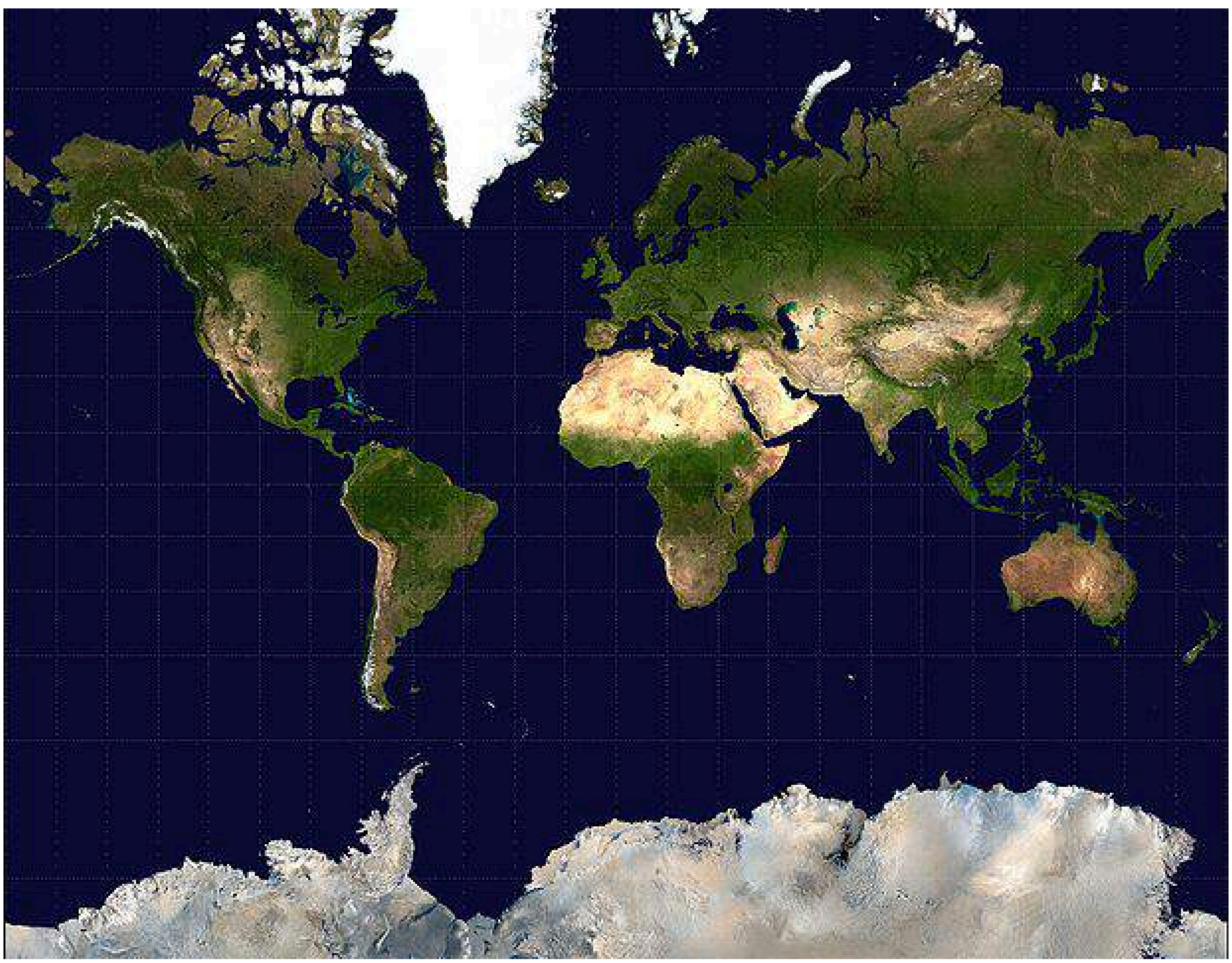


Figure 2.3.1 *Mercator projection (By Mdf (Own work) [Public domain], via Wikimedia Commons).*

The **Goode homolosine** projection is often used to represent the entire globe (Figure 2.3.2). An advantage of this projection is that it does not exaggerate distance and area as much as the Mercator projection. But there are significant disadvantages too; obviously there is the problem of the oceans (and Greenland) being split apart in the figure below. Other versions of this projection may keep the oceans somewhat intact, but then the continents are disrupted. There is no way to keep both the oceans and the continents intact with this projection. The homolosine projection is also useless for navigation, as the lines of longitude point in different directions over various parts of the map.

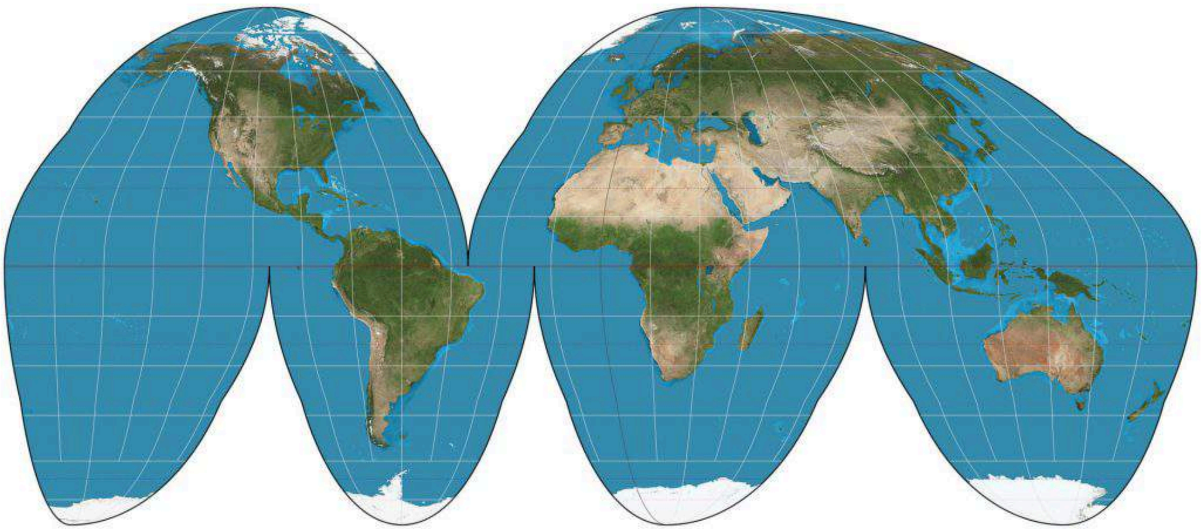


Figure 2.3.2 *Goode homolosine projection (By Strebe (Own work) [CC BY-SA 3.0],via Wikimedia Commons).*

The **Robinson planisphere projection** (Figure 2.3.3) keeps latitude horizontal, but shows some convergence of longitude. There is still some distortion, but not as much as in a Mercator. This projection is used mostly for data presentation.

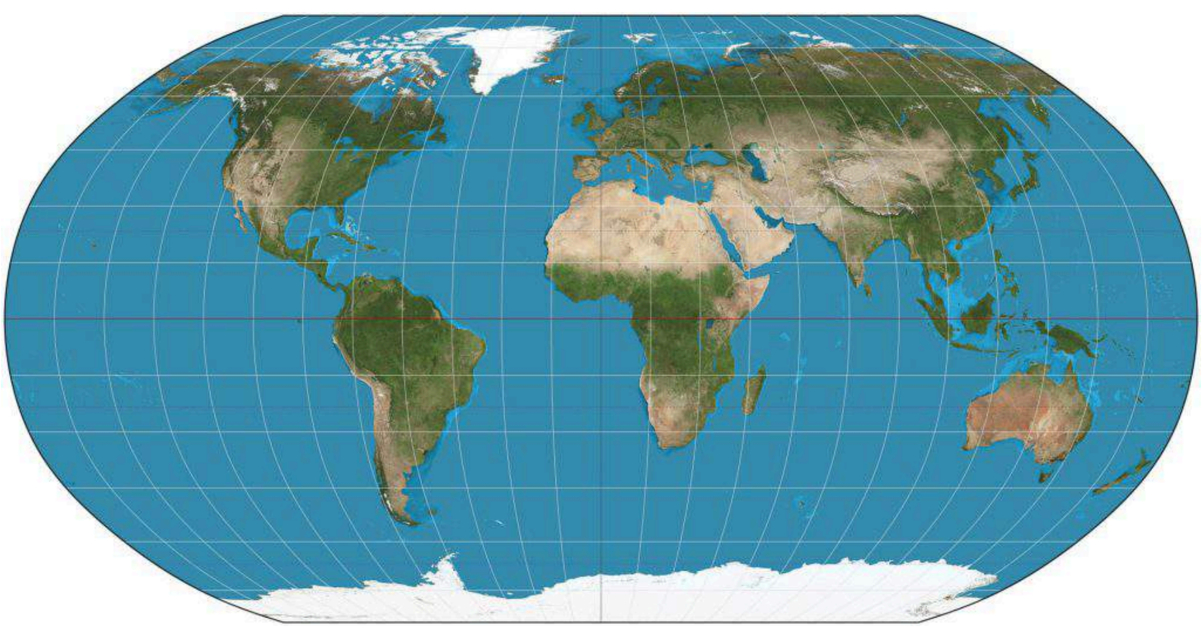


Figure 2.3.3 *Robinson planisphere projection. (By Strebe (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

In oceanography, our use of maps is not limited to viewing the Earth's surface; we also need to see what's at the bottom of the ocean. Other map types include **bathymetric** maps (Figure 2.3.4). These are similar to topographic maps for terrestrial locations, with lines connecting areas of equal depth. The closer together the lines, the steeper the feature. In the example below, the steep continental slope is represented by the high density of depth contours as the colors transition from light blue to dark blue. The well-spaced dark blue lines in the bottom right of the figure represent the relatively flat deep seafloor.

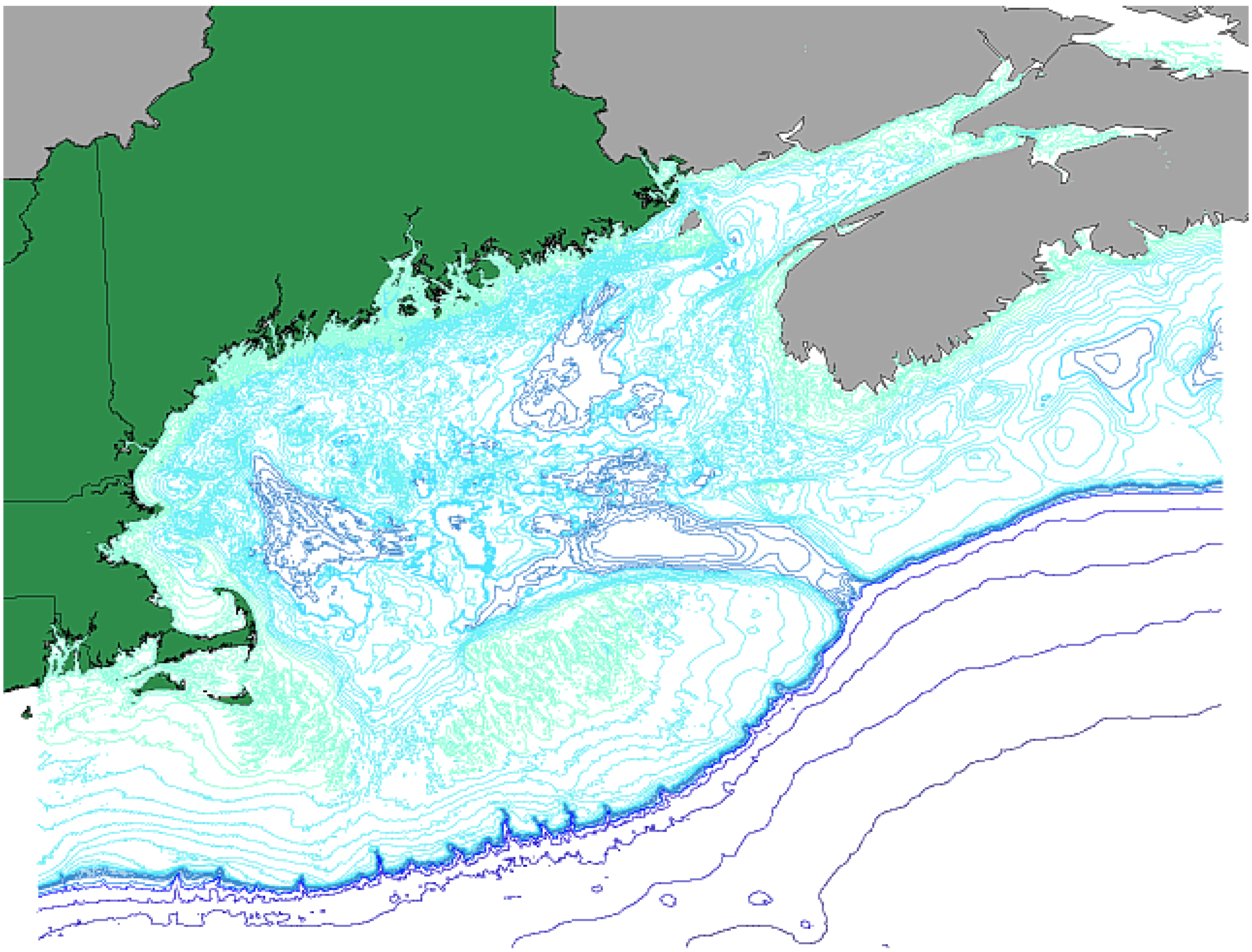


Figure 2.3.4 *Bathymetric map of the Gulf of Maine (USGS).*

Physiographic maps present bathymetry data as a 3D relief map to show ocean features (Figure 2.3.5). It is important to note that they tend to show significant vertical exaggeration. In the example below, there are several hundred kilometers of coastline, while the change in depth from the continental shelf to the seafloor is only a few km.

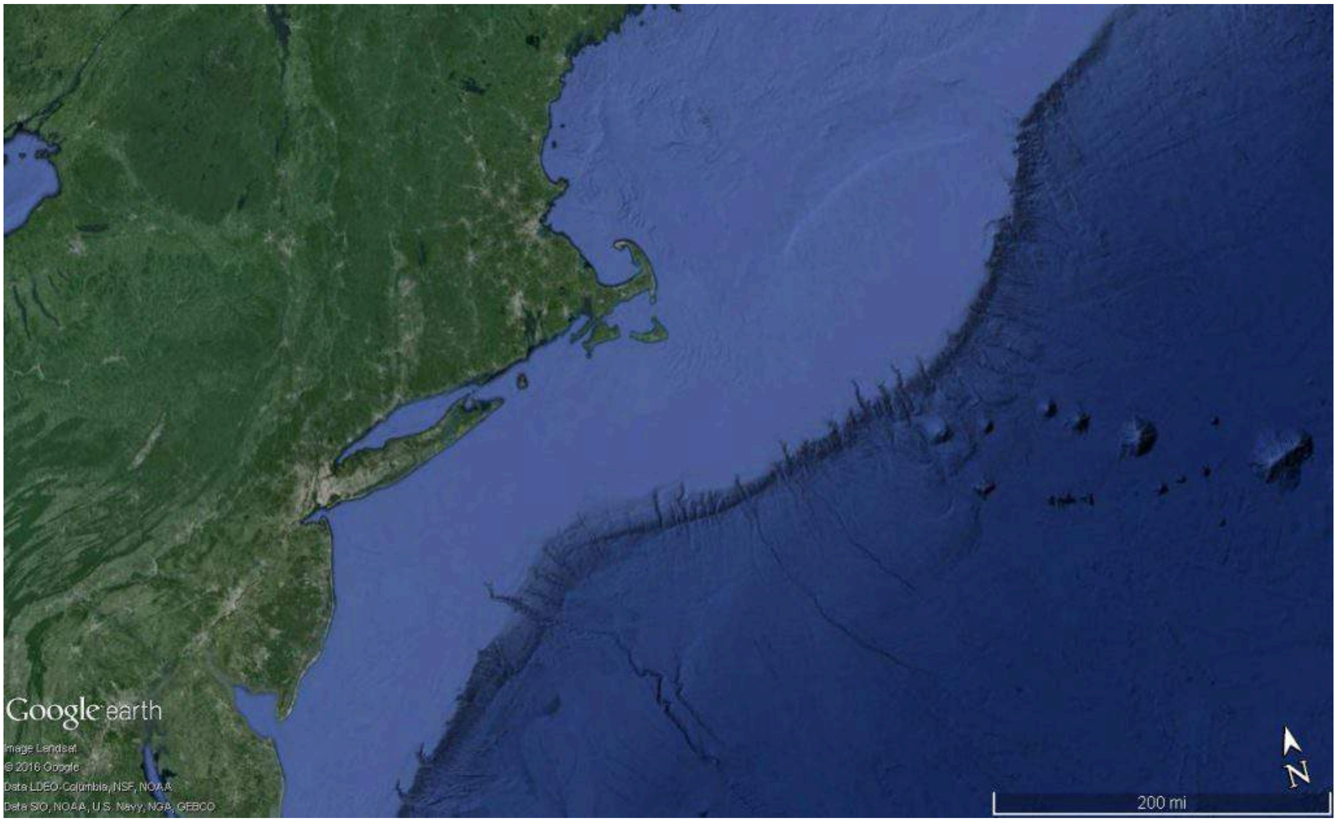


Figure 2.3.5 *Physiographic map of the southern New England coast (Google Earth, Map Data: SIO, NOAA, US Navy, NGA, GEBCO).*

Additional web links for more information:

- For more on map projections: https://www.axismaps.com/guide/general/map-projections/
- Examples of many types of map projections: https://en.wikipedia.org/wiki/List_of_map_projections

CHAPTER 3: THE ORIGIN AND STRUCTURE OF EARTH

Chapter 3: The Origin and Structure of Earth

Learning Objectives

After reading this chapter you should be able to:

- know the age of the universe, the solar system, and Earth
- explain the processes responsible for the early formation of Earth
- know the various layers of the Earth, including their composition
- know the differences between oceanic and continental crust
- explain the difference between lithosphere and asthenosphere
- explain the concept of isostasy
- explain how indirect methods can be used to investigate the interior of Earth

To understand the geological processes occurring in the ocean, it is important to recognize some of the phenomena that led to the formation and structure of the Earth. In this chapter we will start at the very beginning, with a discussion of the Big Bang and the origin of the universe and our solar system. From there, we will investigate the formation of the Earth, and the reasons behind its interior and exterior structure. Finally, we will end the chapter by attempting to answer the question of how we can know what is happening deep within the Earth's interior.

3.1 Origin of Earth and the Solar System

*Modified from Karla Panchuk in "Physical Geology" by Steven Earle**

According to the **Big Bang theory**, the universe blinked violently into existence 13.77 billion years ago (Figure 3.1.1). The Big Bang is often described as an explosion, but imagining it as an enormous fireball isn't accurate. The Big Bang involved a sudden expansion of matter, energy, and space from a single point. The kind of Hollywood explosion that might come to mind involves expansion of matter and energy *within* space, but during the big bang, space *itself* was created.

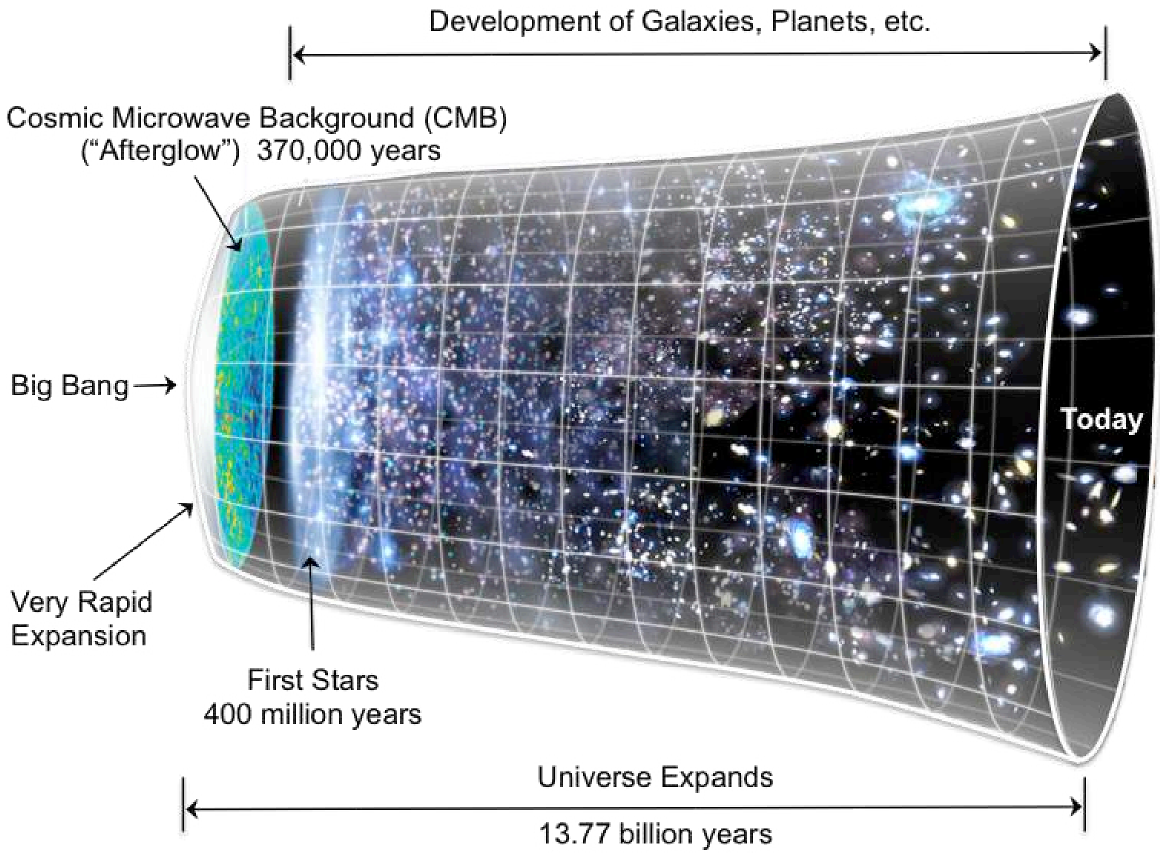


***Figure 3.1.1** The Big Bang and development of the universe (Steven Earle, "Physical Geology").*

At the start of the Big Bang, the universe was too hot and dense to be anything but a sizzle of particles smaller than atoms, but as it expanded, it also cooled. Eventually some of the particles collided and stuck together. Those collisions produced hydrogen and helium, the most common elements in the universe, along with a small amount of lithium. Gravity caused clouds of these early elements to coalesce into stars, and it was inside these stars that heavier elements were formed

Our solar system began to form around 5 billion years ago, roughly 8.7 billion years after the Big Bang. A **solar system** consists of a collection of objects orbiting one or more central stars. All solar systems start out the same way. They begin in a cloud of gas and dust called a **nebula**. Nebulae are some of the most beautiful objects that have been photographed in space, with vibrant colors from the gases and dust they contain, and brilliant twinkling from the many stars that have formed within them (Figure 3.1.2). The gas consists largely of hydrogen and helium, and the dust consists of tiny mineral grains, ice crystals, and organic particles.

Figure 3.1.2 *Photograph of a nebula. The Pillars of Creation within the Eagle Nebula viewed in visible light (left) and near infrared light (right). Near infrared light captures heat from stars, and allows us to view stars that would otherwise be hidden by dust. This is why the picture on the right appears to have more stars than the picture on the left [NASA, ESA, and the Hubble Heritage Team (STScI/AURA) http://bit.ly/1Dm2X5a].*

A solar system begins to form when a small patch within a nebula (small by the standards of the universe, that is) begins to collapse upon itself. Exactly how this starts isn't clear, although it might be triggered by the violent behavior of nearby stars as they progress through their life cycles. Energy and matter released by these stars might compress the gas and dust in nearby neighborhoods within the nebula. Once it is triggered, the collapse of gas and dust within that patch continues for two reasons. One of those reasons is that gravitational force pulls gas molecules and dust particles together. But early in the process, those particles are very small, so the gravitational force between them isn't strong. So how do they come together? The answer is that dust first accumulates in loose clumps for the same reason dust bunnies form under your bed: static electricity. As the small patch within a nebula condenses, a star begins to form from material drawn into the center of the patch, and the remaining dust and gas settle into a disk that rotates around the star. The disk is where planets eventually form, so it's called a **protoplanetary disk**. In Figure 3.1.3 the image in the upper left shows an artist's impression of a protoplanetary disk, and the image in the upper right shows an actual protoplanetary disk surrounding the star HL Tauri. Notice the dark rings in the protoplanetary disk. These are gaps where planets are beginning to form. The rings are there because incipient planets are beginning to collect the dust and gas in their orbits. There is an analogy for this in our own solar system, because the dark rings are akin to the gaps in the rings of Saturn (Fig. 3.1.3, lower left), where moons can be found (Fig. 3.1.3, lower right).

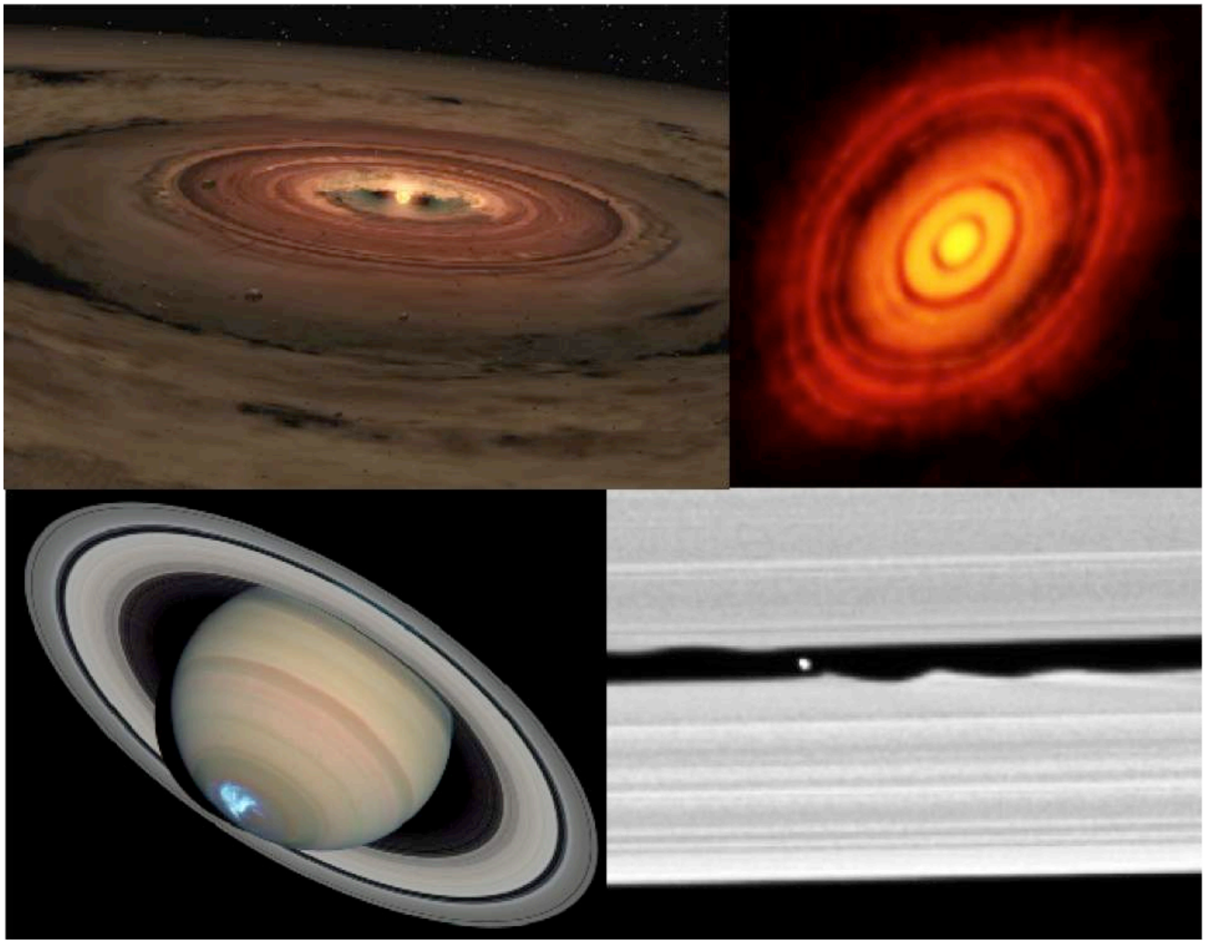


Figure 3.1.3 *Protoplanetary disks and Saturn's rings. Upper left: An artists impression of a protoplanetary disk containing gas and dust, surrounding a new star. [NASA/ JPL-Caltech, http://1.usa.gov/1E5tFJR] Upper right: A photograph of the protoplanetary disk surrounding HL Tauri. The dark rings within the disk are thought to be gaps where newly forming planets are sweeping up dust and gas. [ALMA (ESO/NAOJ/NRAO) http://bit.ly/1KNCq0e]. Lower left: A photograph of Saturn showing similar gaps within its rings. The bright spot at the bottom is an aurora, similar to the northern lights on Earth. [NASA, ESA, J. Clarke (Boston University), and Z. Levay (STScI) http://bit.ly/1IfSCX5] Lower right: a close-up view of a gap in Saturn's rings showing a small moon as a white dot. [NASA/JPL/Space Science Institute, http://1.usa.gov/1g2EeYw].*

In general, planets can be classified into three categories based on what they are made of (Fig. 3.1.4). **Terrestrial planets** are those planets like Earth, Mercury, Venus, and Mars that have a core of metal surrounded by rock. **Jovian planets** (also called **gas giants**) are those planets like Jupiter and Saturn that consist predominantly of hydrogen and helium. **Ice giants** are planets such as Uranus and Neptune that consist largely of water ice, methane (CH_4) ice, and ammonia (NH_3) ice, and have rocky cores. Often, the ice giant planets Uranus and Neptune are grouped with Jupiter and Saturn as gas giants; however, Uranus and Neptune are very different from Jupiter and Saturn.

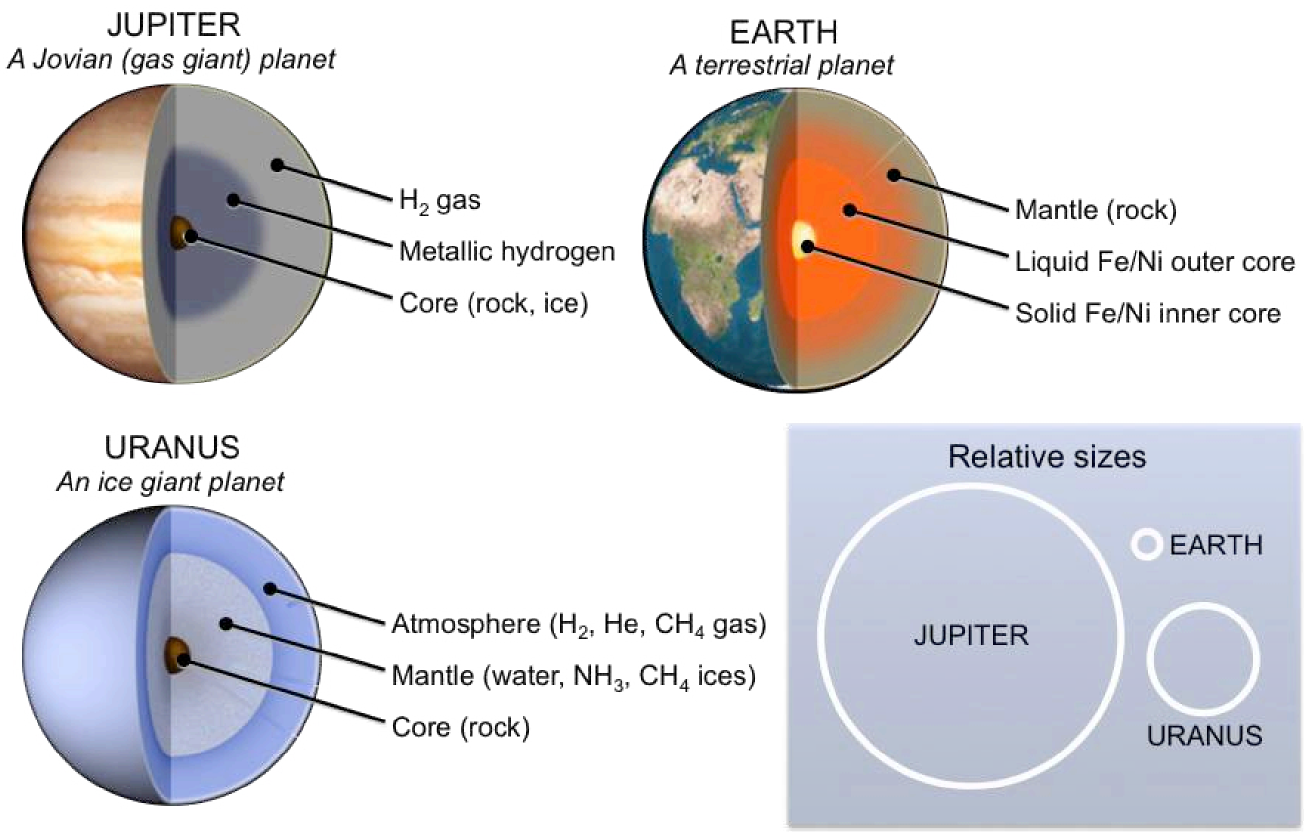


Figure 3.1.4 *Three types of planets. Jovian (or gas giant) planets such as Jupiter consist mostly of hydrogen and helium. They are the largest of the three types. Ice giant planets such as Uranus are the next largest. They contain water, ammonia, and methane ice. Terrestrial planets such as Earth are the smallest, and they have metal cores covered by rocky mantles. [KP, after public domain images by Francesco A, Wolfman SF (http://bit.ly/1eP75P4), and NASA (http://1.usa.gov/1gFVsf6, http://1.usa.gov/1M89jI3)].*

These three types of planets are not mixed together randomly within our solar system. Instead they occur in a systematic way, with terrestrial planets closest to the sun, followed by the Jovian planets and then the ice giants. Part of the reason for this arrangement is the **frost line** (also referred to as the **snow line**). The frost line separated the inner part of the protoplanetary disk closer to the sun, where it was too hot to permit anything but silicate minerals and metal to crystallize, from the outer part of the disk farther from the Sun, where it was cool enough to allow ice to form. As a result, the objects that formed in the inner part of the protoplanetary disk consist largely of rock and metal, while the objects that formed in the outer part consist largely of gas and ice. The young sun also blasted the solar system with raging **solar winds** (winds made up of energetic particles), which helped to drive lighter molecules toward the outer part of the protoplanetary disk.

The objects in our solar system formed by **accretion**. Early in this process, mineral and rock particles collected in fluffy clumps because of static electricity. As the mass of the clumps increased, gravity became more important, pulling material from farther away and growing these solid masses into larger and larger bodies. Eventually the mass of the objects became large enough that their gravity was strong enough to hang onto gas molecules, because gas molecules are very light.

Our Earth formed though this process of accretion about 4.6 billion years ago. The early Earth was very hot and had a molten, fluid composition, with lost of geological and volcanic activity on the surface. The Earth's heat came from a variety of processes:

- Heat came from the decay of radioactive elements within the Earth, specifically the decay of 235U, 238U, 40K, and 232Th, which are primarily present in the mantle. The total heat produced that way has been decreasing over time (because these isotopes are getting used up), and is now roughly 25% of what it was when Earth formed. This means that Earth's interior is slowly becoming cooler.

- Heat came from the thermal energy already contained within the objects that accreted to form the Earth.
- Heat came from collisions. When objects hit Earth, some of the energy from their motion went into deforming Earth, and some of it was transformed into heat. (The very worst collision that Earth experienced was with a planet named Theia, which was approximately the size of Mars. Not long after Earth formed, Theia struck Earth. When Theia slammed into Earth, Theia's metal core merged with Earth's core, and debris from the outer silicate layers was cast into space, forming a ring of rubble around Earth. The material within the ring coalesced into a new body in orbit around Earth, giving us our moon. Remarkably, the debris may have coalesced in 10 years or fewer! This scenario for the formation of the moon is called the **giant impact hypothesis**.)
- As Earth became larger, its gravitational force became stronger. This increased Earth's ability to draw objects to it, but it also caused the material making Earth to be compressed, rather like Earth giving itself a giant gravitational hug. Compression causes materials to heat up.

Heating had a very important consequence for Earth's structure. As Earth grew, it collected a mixture of silicate mineral grains as well as iron and nickel. These materials were scattered throughout Earth. That changed when Earth began to heat up: it got so hot that both the silicate minerals and the metals melted. The metal melt was much denser than the silicate mineral melt, so the metal melt sank to Earth's center to become its core, and the silicate melt rose upward to become Earth's crust and mantle. In other words, Earth unmixed itself. The separation of silicate minerals and metals into a rocky outer layer and a metallic core, respectively, is called **differentiation**. Gravity has since pulled Earth into an almost spherical shape with a radius of 6371 km, and a circumference of about 40,000 km. However, it is not a perfect sphere, as the Earth's rotation causes an equatorial bulge, so that the Earth's circumference is 21 km (0.3%) wider at the equator than it is pole to pole. Thus it is technically an "oblate spheroid."

If we were to take an inventory of the elements that make up Earth, we would find that 95% of Earth's mass comes from only four elements: oxygen, magnesium, silicon, and iron. Most of the remaining 5% comes from aluminum, calcium, nickel, hydrogen, and sulphur. We know that the Big Bang made hydrogen, helium, and lithium, but where did the rest of the elements come from? The answer is that the other elements were made by stars. The heat and pressure within stars cause smaller atoms to smash together and fuse into new, larger atoms. For example, when hydrogen atoms smash together and fuse, helium is formed. Large amounts of energy are released when some atoms fuse and that energy is what causes stars to shine.

It takes larger stars to make elements as heavy as iron and nickel. Our Sun is an average star; after it uses up its hydrogen fuel to make helium, and then some of that helium is fused to make small amounts of beryllium, carbon, nitrogen, oxygen, and fluorine, it will be at the end of its life. It will stop making atoms and will cool down and bloat until its middle reaches the orbit of Mars. In contrast, large stars end their lives in spectacular fashion, exploding as supernovae and casting off newly formed atoms —including the elements heavier than iron — into space. It took many generations of stars creating heavier elements and casting them into space before heavier elements were abundant enough to form planets like Earth.

3.2 Structure of Earth

*Portions modified from "Physical Geology" by Steven Earle**

In the previous section we learned that materials in the early Earth were sorted through the process of differentiation, with denser materials like iron and nickel sinking to the center, and lighter materials (oxygen, silicon, magnesium) remaining near the surface. As a result, the Earth is composed of layers of different composition and increasing density as you move from the surface to the center (Figure 3.2.1).

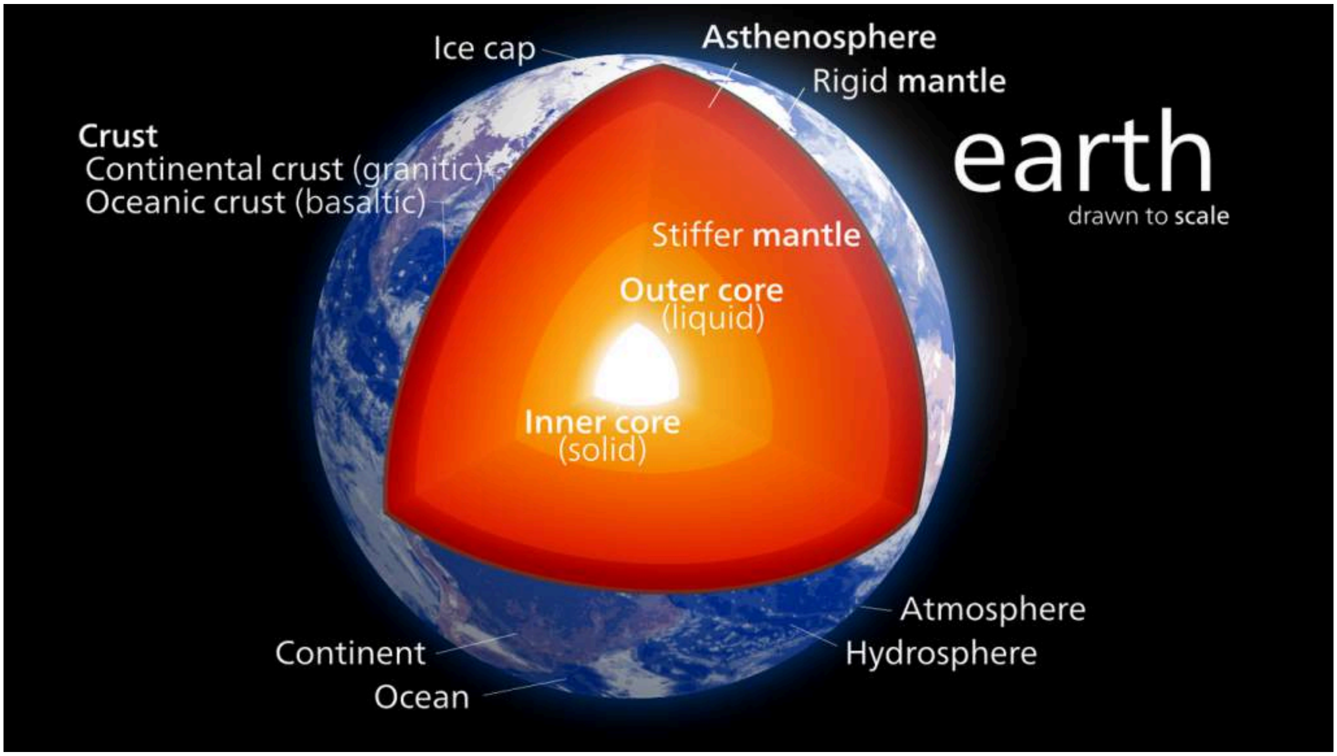


Figure 3.2.1 *Interior structure of Earth (By Kelvinsong (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

The traditional view based on chemical composition recognizes four distinct layers:

The **inner core** lies at the center of the Earth, and is about 1200 km thick. It is composed primarily of iron alloys and nickel, with about 10% comprised of oxygen, sulfur or hydrogen. The temperature in the inner core is about 6000 ^{o}C (10,800 ^{o}F), which is roughly the temperature of the surface of the sun (section 3.1 explains the sources of this intense heat). Despite the high temperature that should melt these metals, the extreme pressure (from literally the weight of the world) keeps the inner core in the solid phase. The solid metals also make the inner core very dense, at about 17 g/cm^3, giving the inner core about one-third of the Earth's total mass.

The **outer core** sits outside of the inner core. It has the same composition as the inner core, but it exists as a fluid, rather than a solid. The temperature is 4000-6000 ^{o}C, and the metals remain in the liquid state because the pressure is not as great as in the inner core. It is the movement of the fluid iron in the outer core that creates Earth's magnetic field (see section 4.2). The outer core is 2300 km thick, and has a density of 12 g/cm^3.

The **mantle** extends from the outer core to just under Earth's surface. It is 2900 km thick, and contains about

80% of the Earth's volume. The mantle consists of iron and magnesium silicates and magnesium oxides, so it is more similar to the rocks of Earth's surface than to the materials in the core. The mantle has a density of 4.5 g/cm^3, and temperatures in the range of 1000-1500 ^{o}C. The uppermost layer of the mantle is more rigid, while the deeper regions are fluid, and it is the motion of fluid materials in the mantle that is responsible for plate tectonics (see section 4.3). Magma that rises to the surface through volcanoes originates in the mantle.

The outermost layer is the **crust**, which forms the solid, rocky surface of the Earth. The crust averages 15-20 km thick, but in some places, such as under mountains, the crust can reach thicknesses of up to 100 km. There are two main types of crust; **continental crust** and **oceanic crust** that differ in a number of ways. Continental crust is thicker than oceanic crust, averaging 20-70 km thick, compared to 5-10 km for oceanic crust. Continental crust is less dense than oceanic crust (2.7 g/cm^3 vs. 3 g/cm^3), and it is much older. The oldest rocks in continental crust are about 4.4 billion years old, while the oldest oceanic crust only goes back about 180 million years. Finally, the two types of crust differ in their composition. Continental crust is made largely of granite. This is because underground or surface magmas can cool slowly, which allows time for crystal structures to form before the rocks solidify, which leads to granite. Oceanic crust is mostly composed of basalts. Basalts also form from cooling magmas, but they cool in the presence of water, which makes them cool much faster and does not allow time for crystals to form.

Based on physical characteristics, we can also divide the outermost layers of Earth into the **lithosphere** and **asthenosphere**. The lithosphere consists of the crust and the cool, rigid, outer 80-100 km of the mantle. The crust and outer mantle moves together as a unit, so they are combined together into the lithosphere. The asthenosphere lies below the lithosphere, from about 100-200 km to about 670 km deep. It includes the more "plastic" softer region of the mantle, where fluid movements can occur. The solid lithosphere is thus floating on the fluid asthenosphere.

Isostasy

To help explain how the lithosphere is floating on the asthenosphere, we need to examine the concept of **isostasy**. Isostasy refers to the way a solid will float on a fluid. The relationship between the crust and the mantle is illustrated in Figure 3.2.2. On the right is an example of a non-isostatic relationship between a raft and solid concrete. It's possible to load the raft up with lots of people, and it still won't sink into the concrete. On the left, the relationship is an isostatic one between two different rafts and a swimming pool full of peanut butter. With only one person on board, the raft floats high in the peanut butter, but with three people, it sinks dangerously low. We're using peanut butter here, rather than water, because its viscosity more closely represents the relationship between the crust and the mantle. Although it has about the same density as water, peanut butter is much more viscous (stiff), and so although the three-person raft will sink into the peanut butter, it will do so quite slowly.

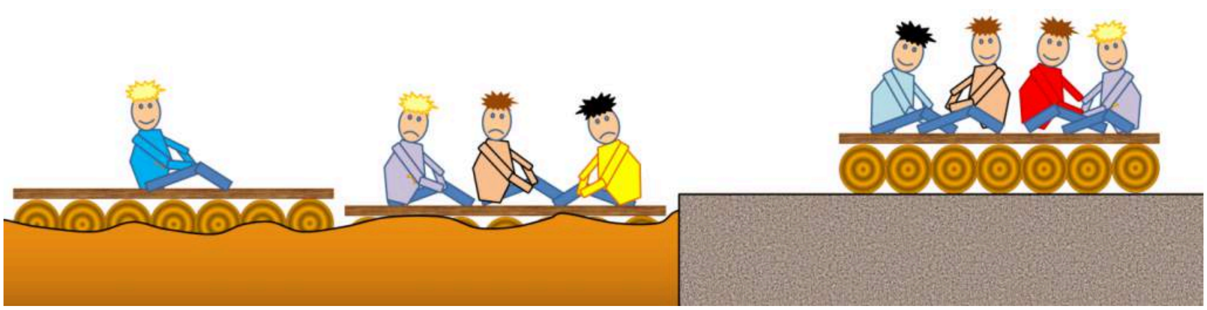


Figure 3.2.2 *Demonstrating isostasy (Steven Earle, "Physical Geology").*

The relationship of Earth's crust to the mantle is similar to the relationship of the rafts to the peanut butter. The raft with one person on it floats comfortably high. Even with three people on it the raft is less dense than the peanut butter, so it floats, but it floats uncomfortably low for those three people. The crust, with an average density of around 2.6 grams per cubic centimeter (g/cm^3), is less dense than the mantle (average density of approximately 3.4 g/cm^3 near the surface, but more than that at depth), and so it is floating on the "plastic" mantle. When more weight is added to the crust, through the process of mountain building, it slowly sinks deeper into the mantle and the mantle material that was there is pushed aside (Figure 3.2.3, left). When that weight is removed by erosion over tens of millions of years, the crust rebounds and the mantle rock flows back (Figure 3.2.3, right).

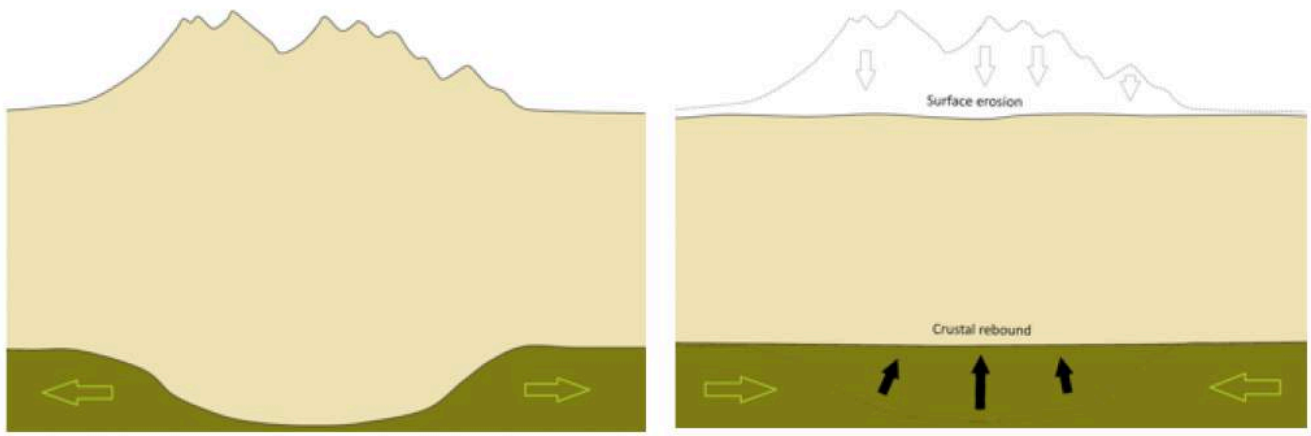


Figure 3.2.3 *Isostatic rebound when mass is removed from the crust (Steven Earle, "Physical Geology").*

The crust and mantle respond in a similar way to glaciation. Thick accumulations of glacial ice add weight to the crust, and as the mantle beneath is squeezed to the sides, the crust subsides. When the ice eventually melts, the crust and mantle will slowly rebound, but full rebound will likely take more than 10,000 years. Large parts of Canada are still rebounding as a result of the loss of glacial ice over the past 12,000 years, and as shown in Figure 3.2.4, other parts of the world are also experiencing isostatic rebound. The highest rate of uplift is in within a large area to the west of Hudson Bay, which is where the Laurentide Ice Sheet was the thickest (over 3,000 m). Ice finally left this region around 8,000 years ago, and the crust is currently rebounding at a rate of nearly 2 cm/ year.

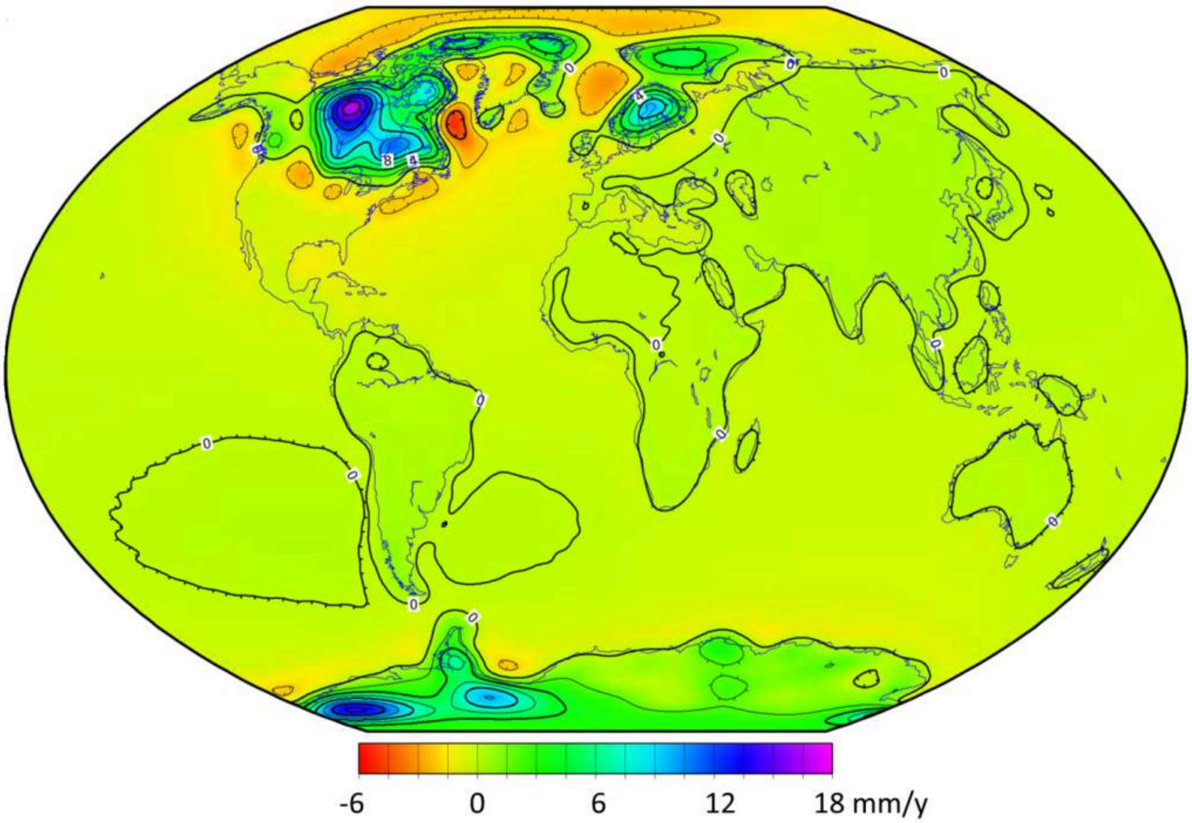


__Figure 3.2.4__ Global rates of isostatic adjustment (Steven Earle, "Physical Geology").

Since continental crust is thicker than oceanic crust, it will float higher and extend deeper into the mantle than oceanic crust. Crust is thickest where there are mountains, so the Moho will be deeper under mountains than under the oceanic crust. Since oceanic crust is also denser than continental crust, it floats lower on the mantle. Since the oceanic crust lies lower than the continental crust, and since water flows downhill to reach the lowest point, this explains why water has accumulated over the oceanic crust to form the oceans.

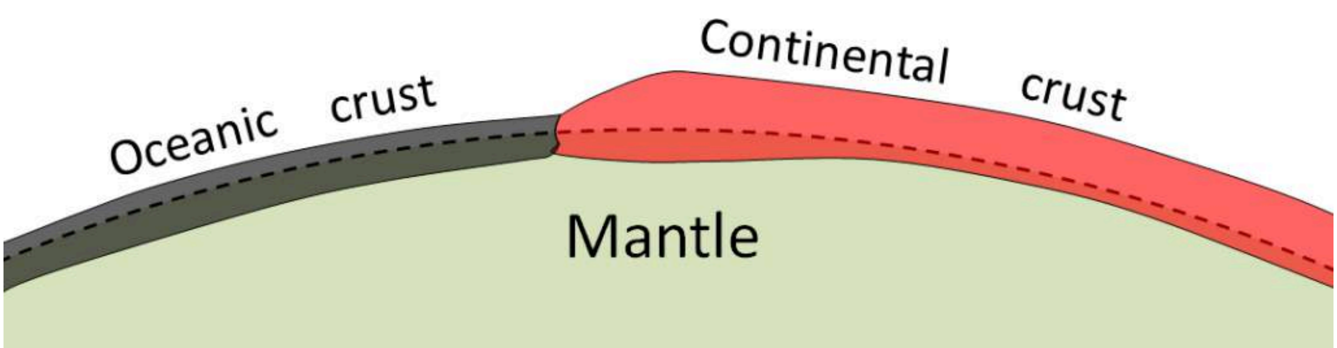


__Figure 3.2.5__ Thinner, denser oceanic crust floats lower on the mantle than thicker, less dense continental crust (Steven Earle, "Physical Geology").

3.3 Determining the Structure of Earth

*Modified from "Physical Geology" by Steven Earle**

The previous section described the properties and composition of Earth's interior, which begs the question: how can we know what conditions are like deep in the Earth? It's easy to sample the crust through drilling, and mantle material often comes to the surface as magma, but the farthest we have been able to drill into the crust so far is only about 12 km; this for a planet with a radius of 6370 km! So to understand the composition and structure of the Earth's deep interior, we need to use indirect methods such as seismology.

Seismology is the study of vibrations within the Earth. These vibrations are caused by various events, including earthquakes, extraterrestrial impacts, explosions, storm waves hitting the shore, and tidal effects. Of course, seismic techniques have been most widely applied to the detection and study of earthquakes, but there are many other applications, and arguably seismic waves provide the most important information that we have concerning Earth's interior. Before going any deeper into Earth, however, we need to take a look at the properties of seismic waves. The types of waves that are useful for understanding Earth's interior are called **body waves**, meaning that, unlike the surface waves on the ocean, they are transmitted through Earth materials.

Imagine hitting a large block of strong rock (e.g., granite) with a heavy sledgehammer. At the point where the hammer strikes it, a small part of the rock will be compressed by a fraction of a millimeter. That compression will transfer to the neighboring part of the rock, and so on through to the far side of the rock, from where it will bounce back to the top — all in a fraction of a second. This is known as a compression wave, and it can be illustrated by holding a loose spring (like a Slinky) that is attached to something (or someone) at the other end. If you give it a sharp push so the coils are compressed, the compression propagates (travels) along the length of the spring and back (Fig. 3.3.1). You can think of a compression wave as a "push" wave — it's called a **P-wave** (although the "P" stands for "primary" because P-waves are the first to arrive at seismic stations). In a P-wave the motion of the particles is parallel to the direction of wave propagation.

When we hit a rock with a hammer, we also create a different type of body wave, one that is characterized by back-and-forth vibrations (as opposed to compressions). This is known as a shear wave (**S-wave**, where the "S" stands for "secondary"), and an analogy would be what happens when you flick a length of rope with an up-and-down motion. As shown in Figure 3.3.1, a wave will form in the rope, which will travel to the end of the rope and back. In this case, the motion of the particles is perpendicular to the direction the wave travels.

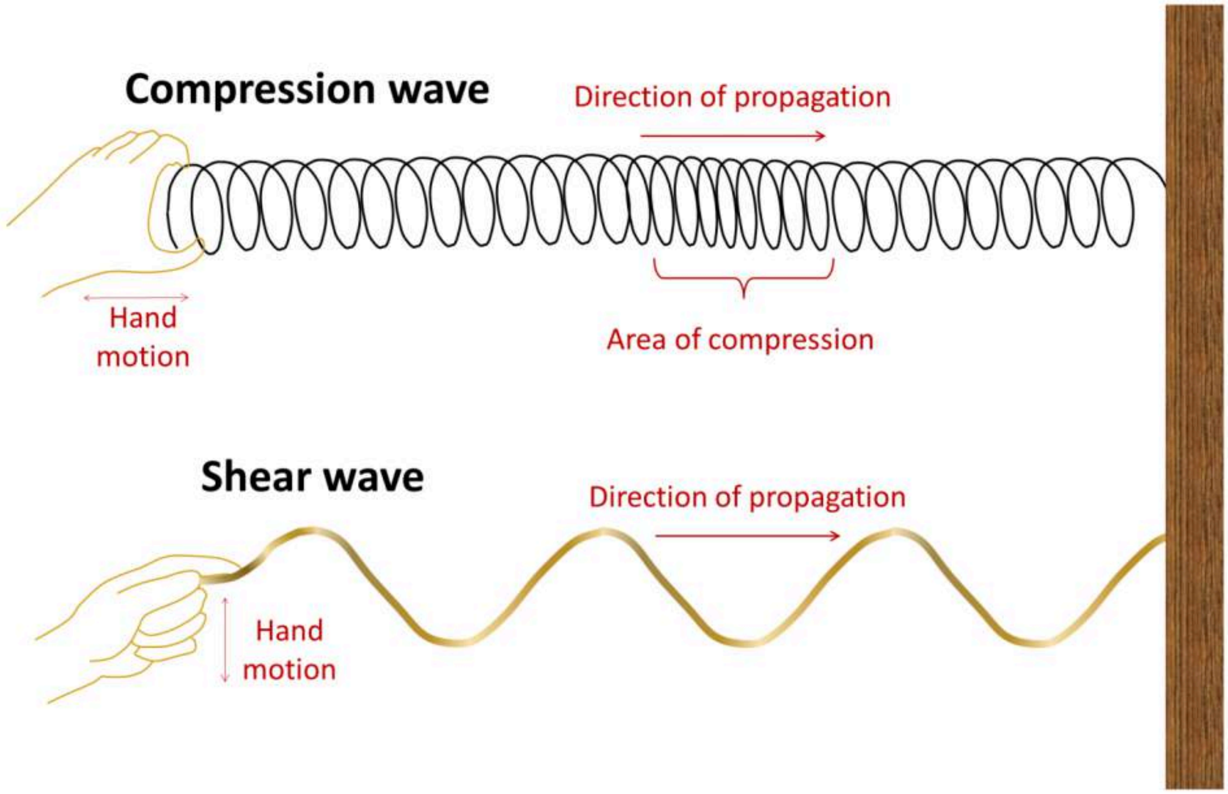


Figure 3.3.1 *Representations of a compression wave (P-wave, top) and a shear wave (S-wave, bottom) (Steven Earle, "Physical Geology").*

Compression waves and shear waves travel very quickly through geological materials. As shown in Figure 3.2.2, typical P-wave velocities are between 0.5 km/s and 2.5 km/s in unconsolidated sediments, and between 3.0 km/s and 6.5 km/s in solid crustal rocks. Of the common rocks of the crust, velocities are greatest in basalt and granite. S-waves are slower than P-waves, with velocities between 0.1 km/s and 0.8 km/s in soft sediments, and between 1.5 km/s and 3.8 km/s in solid rocks.

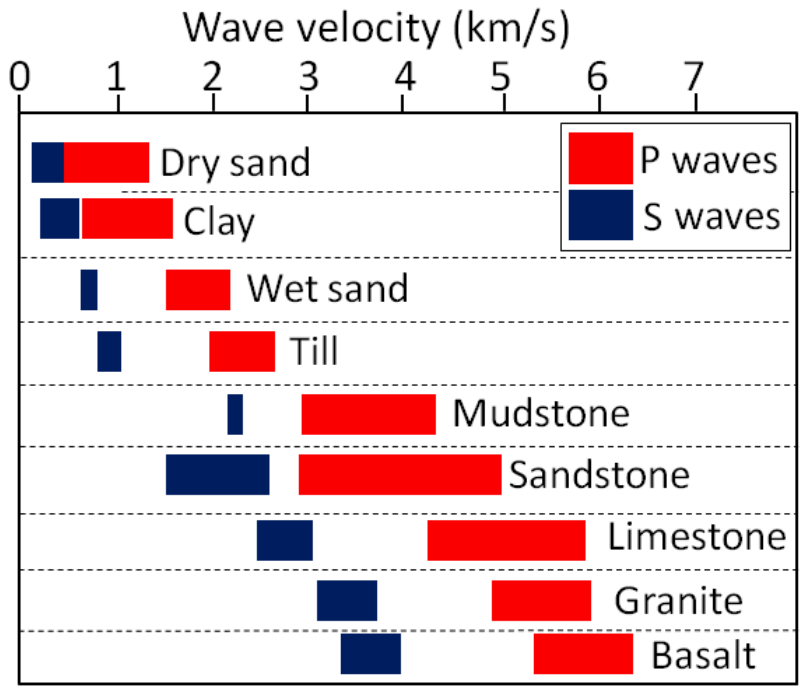


Figure 3.2.2 *Typical velocities of P waves (red) and S waves (blue) in sediments and in solid crustal rocks (Steven Earle, "Physical Geology").*

Mantle rock is generally denser and stronger than crustal rock and both P- and S-waves travel faster through the mantle than they do through the crust. Moreover, seismic-wave velocities are related to how tightly compressed a rock is, and the level of compression increases dramatically with depth. Finally, seismic waves are affected by the phase state of rock. They are slowed if there is any degree of melting in the rock. If the material is completely liquid, P-waves are slowed dramatically and S-waves are stopped altogether.

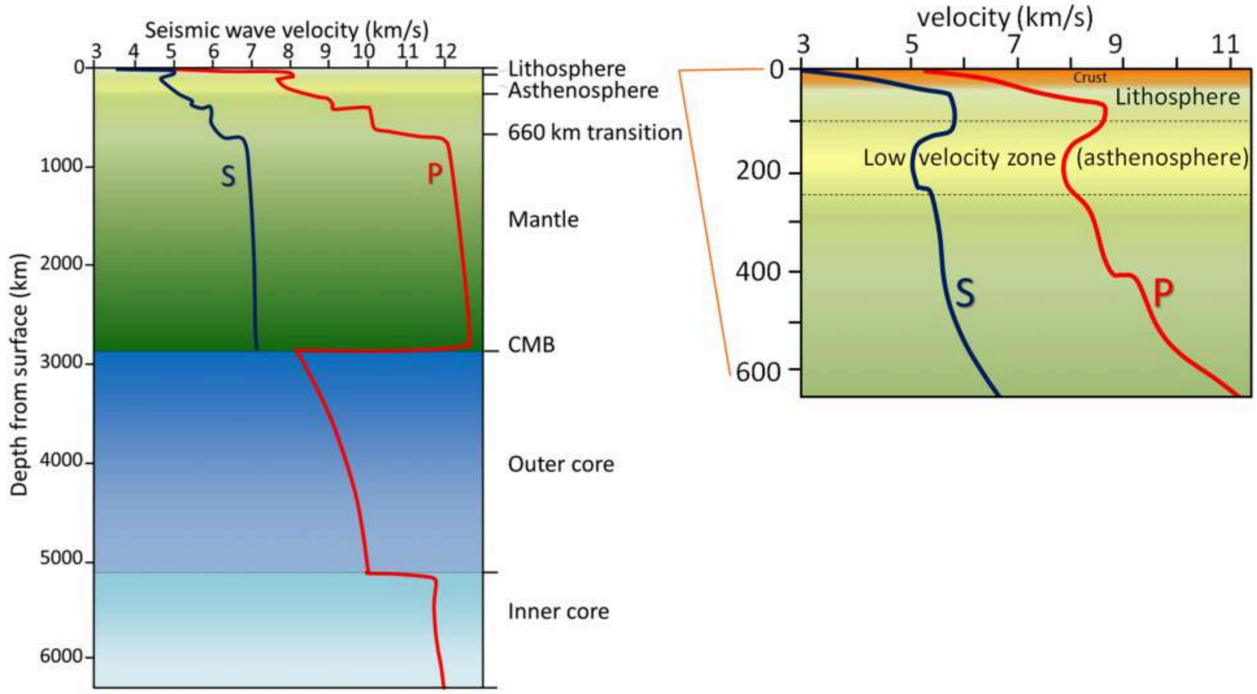


Figure 3.3.3 *Wave velocities through the different layers of the Earth (left). Enhanced view of wave velocities in the crust and upper mantle (right) (Steven Earle, "Physical Geology").*

Accurate seismometers have been used for earthquake studies since the late 1800s, and systematic use of seismic data to understand Earth's interior started in the early 1900s. The rate of change of seismic waves with depth in the Earth (Fig. 3.3.3) has been determined over the past several decades by analyzing seismic signals from large earthquakes at seismic stations around the world. Small differences in arrival time of signals at different locations have been interpreted to show that:

- Velocities are greater in mantle rock than in the crust.
- Velocities generally increase with pressure, and therefore with depth.
- Velocities slow in the area between 100 km and 250 km depth (called the "low-velocity zone"; equivalent to the asthenosphere).
- Velocities increase dramatically at 660 km depth (because of a mineralogical transition).
- Velocities slow in the region just above the core-mantle boundary (the D" layer or "ultra-low-velocity zone").
- S-waves do not pass through the outer part of the core.
- P-wave velocities increase dramatically at the boundary between the liquid outer core and the solid inner core.

One of the first discoveries about Earth's interior made through seismology was in the early 1900s when Croatian seismologist Andrija Mohorovičić (pronounced *Moho-ro-vi-chich*) realized that at certain distances from an earthquake, two separate sets of seismic waves arrived at a seismic station within a few seconds of each other. He reasoned that the waves that went down into the mantle, traveled through the mantle, and then were bent upward back into the crust, reached the seismic station first because although they had farther to go, they traveled faster through mantle rock (as shown in Figure 3.3.4). The boundary between the crust and the mantle is known as the **Mohorovičić discontinuity** (or **Moho**). Its depth is between 60 km and 80 km beneath major mountain ranges, around 30 km to 50 km beneath most of the continental crust, and between 5 km and 10 km beneath the oceanic crust.

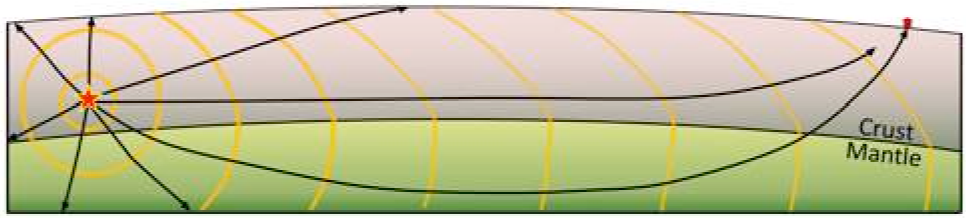


***Figure 3.3.4** Depiction of seismic waves emanating from an earthquake (red star). Some waves travel through the crust to the seismic station (at about 6 km/s), while others go down into the mantle (where they travel at around 8 km/s) and are bent upward toward the surface, reaching the station before the ones that traveled only through the crust (Steven Earle, "Physical Geology").*

Our current understanding of the patterns of seismic wave transmission through Earth is summarized in Figure 3.3.5. Because of the gradual increase in density with depth, all waves are refracted toward the lower density, slower velocity material as they travel through homogenous parts of Earth, and thus tend to curve outward toward the surface. Waves are also refracted at boundaries within Earth, such as at the Moho, at the core-mantle boundary (CMB), and at the outer-core/inner-core boundary. S-waves do not travel through liquids — they are stopped at the CMB — and there is an S-wave shadow on the side of Earth opposite a seismic source. The angular distance from the seismic source to the shadow zone is 103° on either side, so the total angular distance of the shadow zone is 154°. We can use this information to infer the depth to the CMB. P-waves do travel through liquids, so they can make it through the liquid part of the core. Because of the refraction that takes place at the CMB, waves that travel through the core are bent away from the surface, and this creates a P-wave shadow zone on either side, from 103° to 150°. This information can be used to discover the differences between the inner and outer parts of the core.

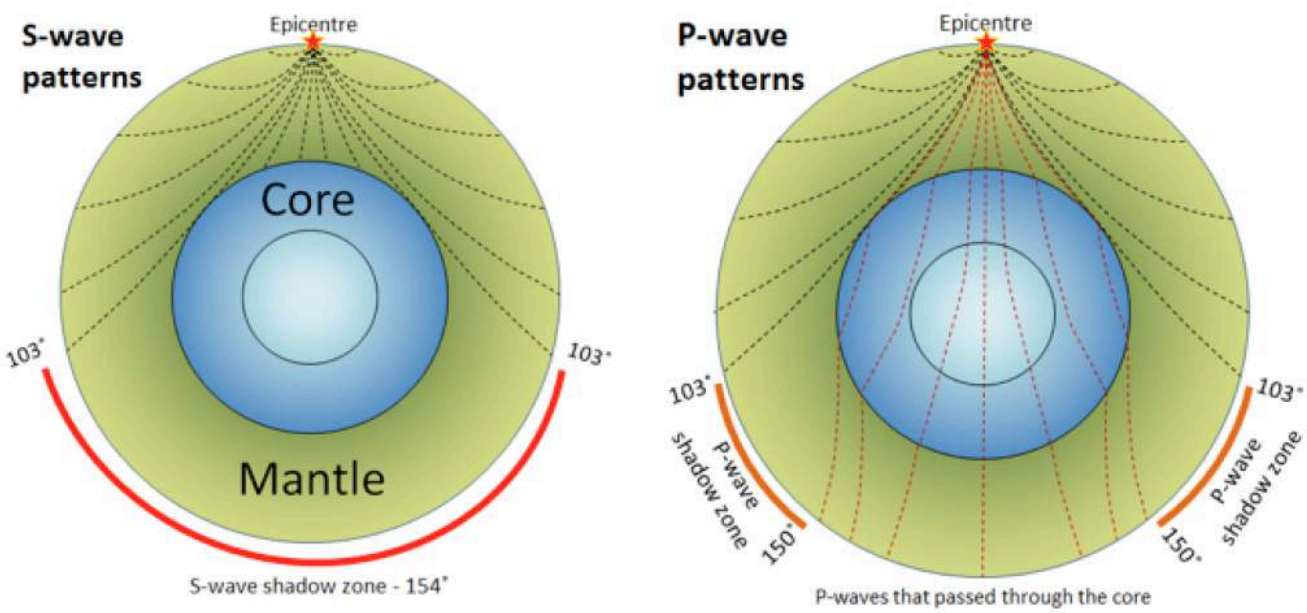


***Figure 3.3.5** Patterns of seismic waves moving through Earth's interior. Since S waves do not pass through the liquid outer core, a shadow zone is created on the opposite side from the original disturbance (Steven Earle, "Physical Geology").*

Using data from many seismometers and hundreds of earthquakes, it is possible to create a two- or three-dimensional image of the seismic properties of part of the mantle. This technique is known as seismic tomography, and an example of the result is shown in Figure 3.3.6.

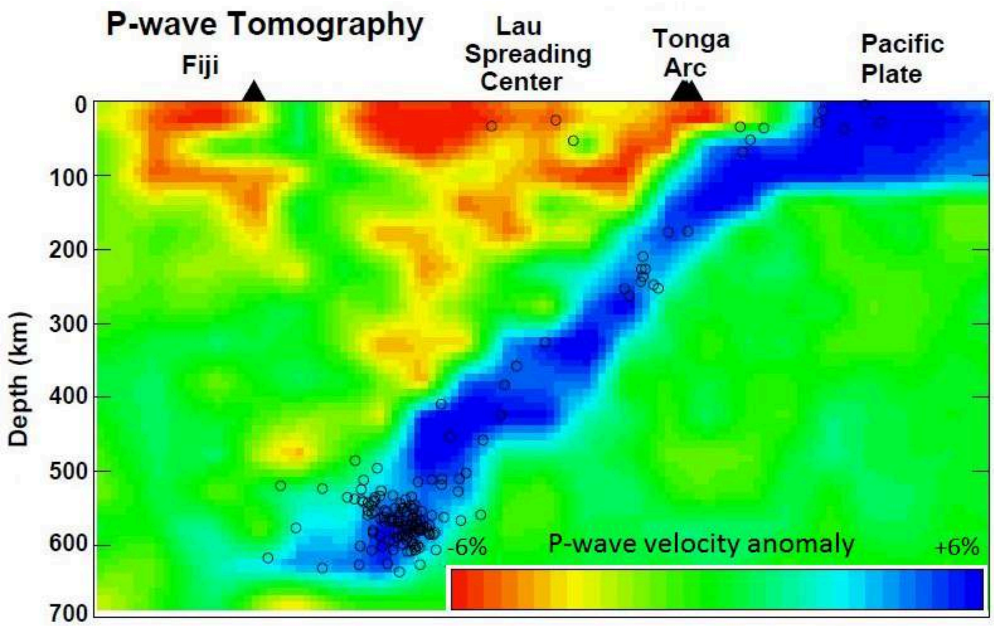


Figure 3.3.6 *Seismic tomography image showing the Pacific Plate (blue) subducting beneath Tonga (Steven Earle, "Physical Geology").*

The Pacific Plate subducts beneath Tonga and appears in Figure 3.3.6 as a 100 km thick slab of cold (blue-colored) oceanic crust that has pushed down into the surrounding hot mantle. The cold rock is more rigid than the surrounding hot mantle rock, so it is characterized by slightly faster seismic velocities. There is volcanism in the Lau spreading center and also in the Fiji area, and the warm rock in these areas has slower seismic velocities (yellow and red colors).

Seismic waves provide us with the structure of the inner Earth, but what about it's other properties? In terms of composition, there are several lines of evidence pointing to a core composed mostly of iron and nickel. Wave properties suggest the core is composed of an element with an atomic number around 25 (iron has an atomic number of 26). Aside from iron, all of the other elements with an atomic number close to 25 are too rare to make up the core. If the Earth was formed through the accretion of smaller bodies such as meteorites, we would expect the composition of Earth to be similar to the composition of meteorites. Meteorites are mostly iron and nickel, but in higher proportions than the Earth's crust. This suggests that most of this heavy iron and nickel from the meteorites must have sunk to the Earth's center as the planet was forming. However, the core is not dense enough to be pure iron and nickel; it it about 10% below the predicted density if that was the case. This is why scientists believe the core is composed of about 10% sulfur, oxygen, and hydrogen. Finally, if the Earth's magnetic field comes from the fluid outer core, the outer core must contain iron. In terms of the temperatures, we can calculate the melting points of these materials over the range of pressures that they would experience in the inner Earth, and then infer the temperatures that would allow these elements to exist in their solid or liquid forms.

CHAPTER 4: PLATE TECTONICS AND MARINE GEOLOGY

Chapter 4: Plate Tectonics and Marine Geology

*Modified from "Physical Geology" by Steven Earle**

Learning Objectives

After reading this chapter you should:

- know Wegener's original evidence for "continental drift"
- understand how paleomagnetic evidence supports the theory of plate movement
- understand mantle convection – the process that moves plates around
- know the different ways that tectonic plates can interact with each other, and the various geological features that result from these interactions.
- know the differences between passive and active continental margins
- be able to define geological features such as seamounts, guyots, hot spots
- understand why some island systems such as Hawaii are formed in chains
- understand how coral reefs evolve from fringing reefs to atolls
- understand the processes behind the formation of hydrothermal vents

In the previous chapter we learned about the crust as the solid outer layer of Earth. But the crust is not a single, solid piece; instead, it is broken up into about a dozen major plates that constantly move past each other, reshaping the surface of the Earth through the process of plate tectonics. Plate tectonics is a model that explains the origins of continents and oceans, folded rocks and mountain ranges, earthquakes and volcanoes, and continental drift. Plate tectonics was first proposed just over 100 years ago, but did not become an accepted part of geology until about 50 years ago. It took 50 years for this theory to become accepted for a few reasons. First, it was a true revolution in thinking about Earth, which was difficult for many established geologists. Second, there was a political gulf between the main proponent of the theory Alfred Wegener (from Germany) and the geological establishment of the day, which was mostly centered in Britain and the United States. Third, the evidence and understanding of Earth that would have supported plate tectonic theory simply didn't exist until the middle of the 20th century. This chapter will examine the evidence for plate tectonics and the mechanism by which it works. Following this, we will discuss the consequences of plate motion and the various geological features that can be explained through this revolutionary idea.

4.1 Alfred Wegener and the Theory of Plate Tectonics

*Modified from "Physical Geology" by Steven Earle**

If you look at a map of Earth, you may notice that some of the continents seem to fit together. An early reference to this phenomenon came from Francis Bacon in the 17th century, who noticed the similarities in the Atlantic coasts of Africa, and North and South America. This apparent fit is due to the fact the continents were once connected, and have since moved apart in what has been called **continental drift**. However, we now know that it is not just the continents that move, so a more correct term is **plate tectonics**. We can credit Alfred Wegener (Figure 4.1.1) as the originator of this idea.

Figure 4.1.1 *Prof. Dr. Alfred Wegener, ca. 1924-1930 (Public domain, via Wikimedia Commons).*

Alfred Wegener (1880-1930) earned a PhD in astronomy at the University of Berlin in 1904, but he had always been interested in geophysics and meteorology and spent most of his academic career working in meteorology. In 1911 he happened on a scientific publication that included a description of the existence of matching Permian-aged terrestrial fossils in various parts of South America, Africa, India, Antarctica, and Australia (Figure 4.1.2). Wegener concluded that this distribution of fossils could only exist if these continents were joined together. Furthermore, some of these transcontinental areas have similar fossils until around 150 million years ago, then they begin to diverge, suggesting that the areas eventually separated and speciation took different paths on the separate continents. Wegener coined the term **Pangaea** ("all land") for the supercontinent from which all of the present-day continents diverged.

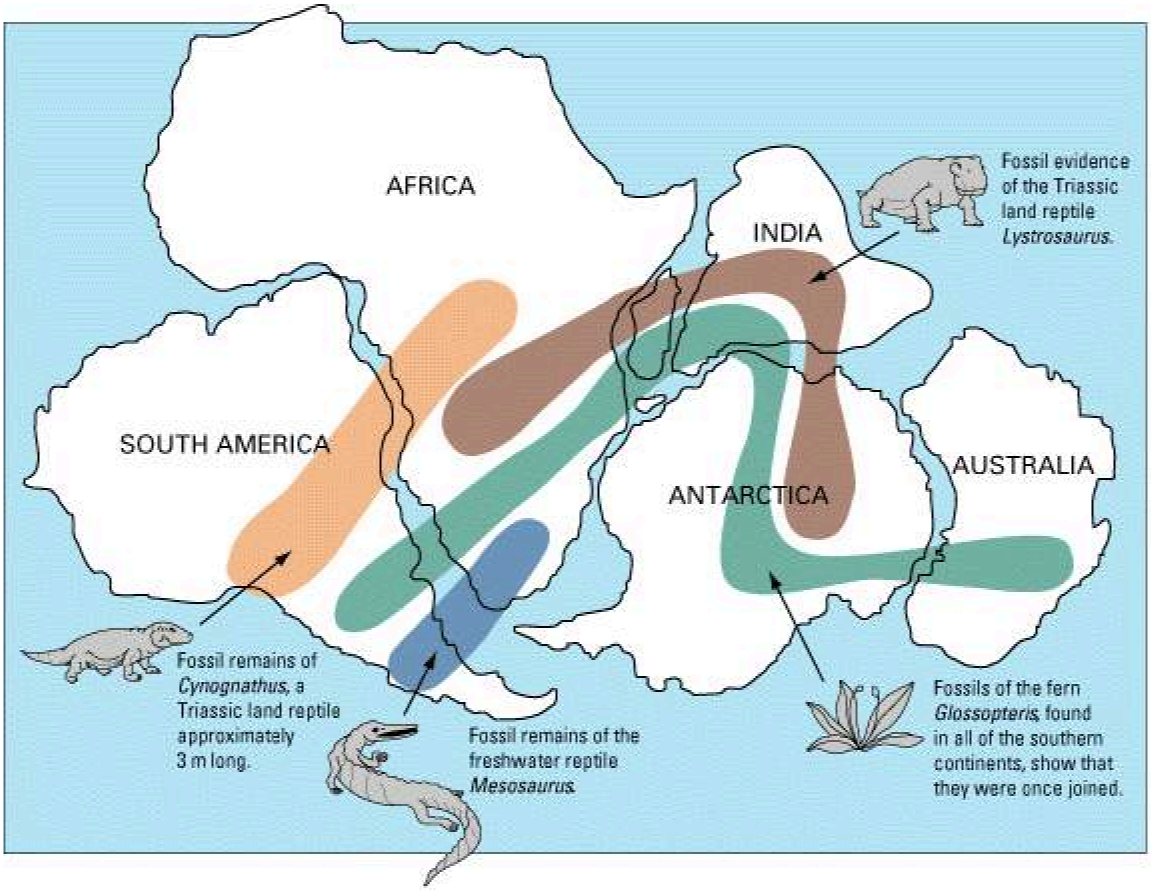


Figure 4.1.2 *Distribution of similar fossils across the continents, suggesting they were once connected into a single supercontinent (Steven Earle, "Physical Geology").*

Wegener pursued his theory with determination — combing the libraries, consulting with colleagues, and making observations — looking for evidence to support it. In addition to the fit of the continents and the fossil evidence, Wegener relied heavily on matching geological patterns across oceans, such as sedimentary strata in South America matching those in Africa (Fig. 4.1.3), North American coalfields matching those in Europe, and the mountains of Atlantic Canada matching those of northern Britain both in morphology and rock type.

Figure 4.1.3 *Matching geological formations spanning from South America to Africa (By Woudloper – Own work, [CC BY-SA 3.0], via Wikimedia Commons).*

Wegener also referred to the evidence for the Carboniferous and Permian (~300 Ma) Karoo Glaciation in South America, Africa, India, Antarctica, and Australia (Fig. 4.1.4). These areas contain evidence of past glacial deposits, including glacial scars oriented away from the poles, despite the fact that some of these locations are now tropical environments. This indicates that these continents were once closer to the south pole where the glaciers could have formed. Wegener argued that this could only have happened if these continents were once all connected as a single supercontinent. He also cited evidence (based on his own astronomical observations) that showed that the continents were moving with respect to each other, and determined a separation rate between Greenland and Scandinavia of 11 m per year, although he admitted that the measurements were not accurate. In fact they weren't even close — the separation rate is actually about 2.5 cm per year!

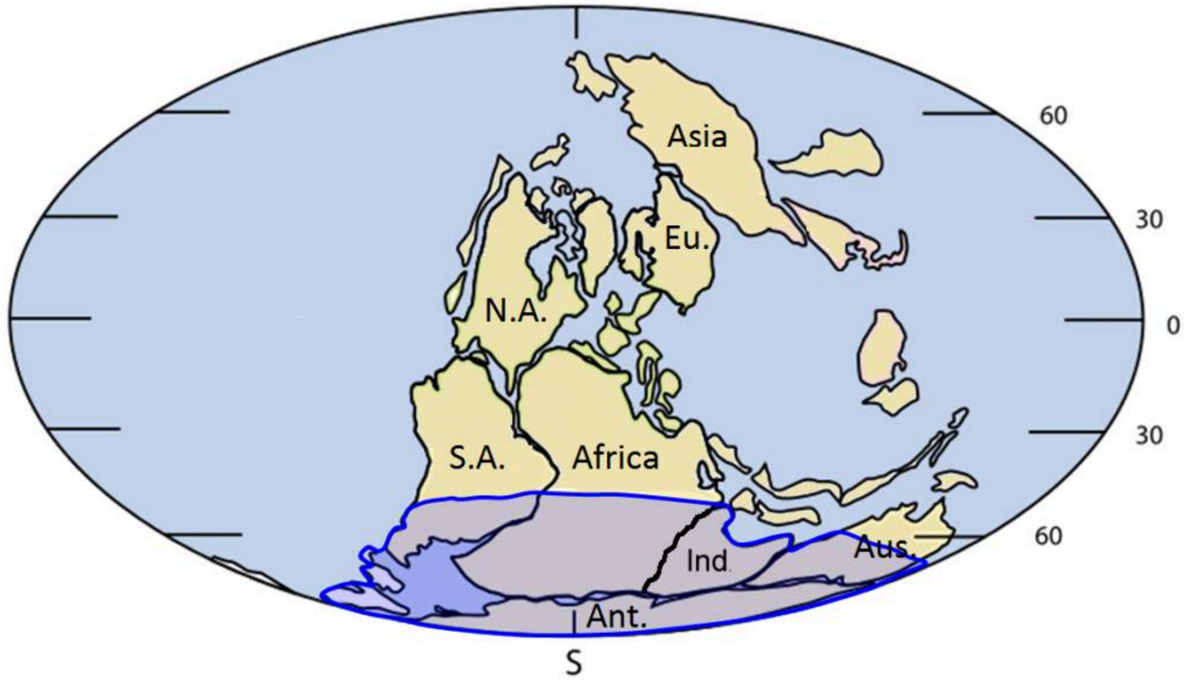


Figure 4.1.4 *Extent of glaciation (shaded area) on Pangaea (Modified by PW from Steven Earle, "Physical Geology").*

Wegener first published his ideas in 1912 in a short book called *Die Entstehung der Kontinente* (*The Origin of Continents*), and then in 1915 in *Die Entstehung der Kontinente und Ozeane* (*The Origin of Continents and Oceans*). He revised this book several times up to 1929, and it was translated into French, English, Spanish, and Russian. However, despite his range of evidence, the continental fits were not perfect and the geological match-ups were not always consistent (while the continental fit left some gaps when using the current coastline, it was demonstrated in the 1960s that using a 500 m depth contour gives a much tighter fit). But the most serious problem of all was that Wegener could not conceive of a good mechanism for moving the continents around. Wegener proposed that the continents were like icebergs floating on heavier crust, but the only forces that he could invoke to propel continents around were *poleflucht,* the effect of Earth's rotation pushing objects toward the equator, and the lunar and solar tidal forces, which tend to push objects toward the west. It was quickly shown that these forces were far too weak to move continents, and without any reasonable mechanism to make it work, Wegener's theory was quickly dismissed by most geologists of the day. Alfred Wegener died in Greenland in 1930 while carrying out studies related to glaciation and climate. At the time of his death, his ideas were tentatively accepted by only a small minority of geologists, and soundly rejected by most. However, within a few decades that was all to change.

Additional links for more information:

- For more about Wegener and the other pioneers of plate tectonics, visit The Geological Society's Plate Tectonics site: https://www.geolsoc.org.uk/Plate-Tectonics/Chap1-Pioneers-of-Plate-Tectonics

---

4.2 Paleomagnetic Evidence for Plate Tectonics

*Modified from "Physical Geology" by Steven Earle**

Although Alfred Wegener would not live to see it, his theory of plate tectonics would gradually gain acceptance within the scientific community as more evidence began to accumulate. Some of the most important evidence came from the study of **paleomagnetism**, or changes in Earth's magnetic field over millions of years.

Earth's magnetic field is defined by the North and South Poles that align generally with the axis of rotation (Figure 4.2.1). The lines of magnetic force flow into Earth in the Northern Hemisphere and out of Earth in the Southern Hemisphere. Because of the shape of the field lines, the magnetic force trends at different angles to the surface in different locations (red arrows of Figure 4.2.1). At the North and South Poles, the force is vertical. Anywhere on the equator the force is horizontal, and everywhere in between, the magnetic force is at some intermediate angle to the surface.

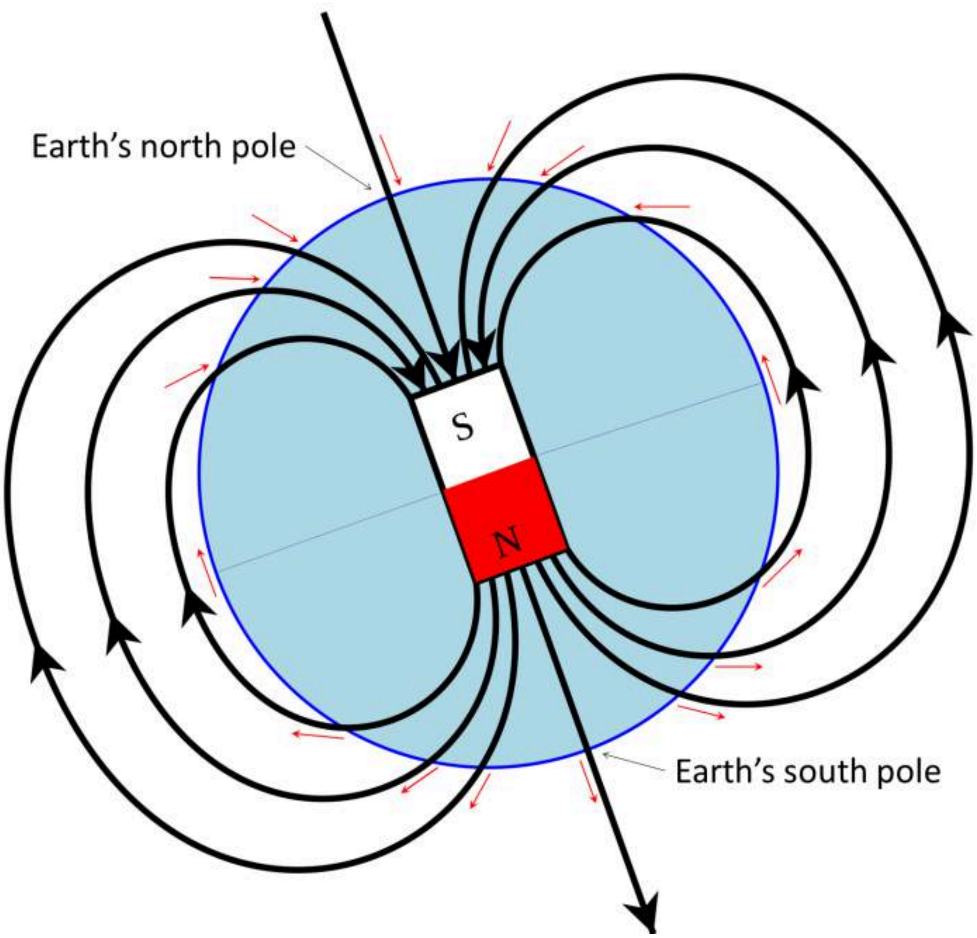


Figure 4.2.1 *Depiction of Earth's magnetic field as a bar magnet coinciding with the core. The south pole of such a magnet points to Earth's North Pole. The red arrows represent the orientation of the magnetic field at various locations on Earth's surface (Steven Earle after: http://upload.wikimedia.org/wikipedia/commons/ 1/17/Earths_Magnetic_Field_ Confusion.svg).*

In its fluid form, the minerals that make up magma are free to move in any direction and take on any orientation. But as the magma cools and solidifies, movement ceases and the mineral orientation and position become fixed. As the mineral magnetite (Fe_3O_4) crystallizes from magma, it becomes magnetized with an orientation parallel to that of Earth's magnetic field at that time, similar to the way a compass needle aligns with the magnetic field to point north. This magnetic record in the rock is called **remnant magnetism**. Rocks like basalt, which cool from a high temperature and commonly have relatively high levels of magnetite, are particularly susceptible to being magnetized in this way, but even sediments and sedimentary rocks, as long as they have small amounts of magnetite, will take on remnant magnetism because the magnetite grains gradually become reoriented following deposition. By studying both the horizontal and vertical components of the remnant magnetism, one can tell not only the direction to magnetic north at the time of the rock's formation, but also the latitude where the rock formed relative to magnetic north.

In the early 1950s, a group of geologists from Cambridge University, including Keith Runcorn, Edward Irving and several others, started looking at the remnant magnetism of Phanerozoic British and European volcanic rocks, and collecting paleomagnetic data. They found that rocks of different ages sampled from generally the same area showed quite different apparent magnetic pole positions (green line, Figure 4.2.2). They initially assumed that this meant that Earth's magnetic field had, over time, departed significantly from its present position, which is close to the rotational pole.

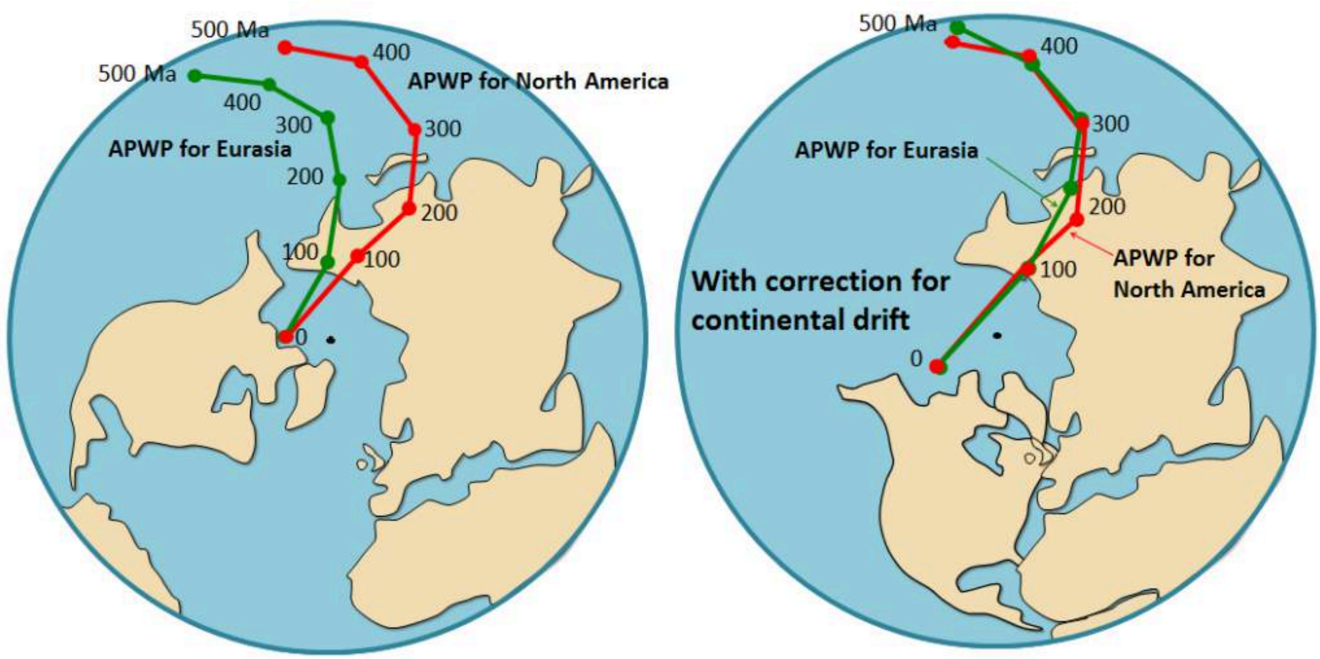


Figure 4.2.2 *Polar wandering curves. Curves from Eurasia and North America seem to show that the north magnetic pole was located in two places simultaneously throughout history (left). However, if the continents are rearranged into Pangaea, the two curves overlap, showing that it is the continents than have moved, not the pole (right) (Steven Earle, "Physical Geology").*

The curve defined by the paleomagnetic data was called a **polar wandering path** because Runcorn and his colleagues initially thought that their data represented actual movement of the magnetic poles (since geophysical models of the time suggested that the magnetic poles did not need to be aligned with the rotational poles). We now know that the magnetic data define movement of continents, and *not* of the magnetic poles, so we call it an ***apparent*** **polar wandering path** (APWP). Runcorn and colleagues soon extended their work to North America, and this also showed *apparent* polar wandering, but the results were not consistent with those from Europe (Figure 4.2.2). For example, the 200 Ma pole for North America placed somewhere in China, while the 200 Ma pole for Europe placed in the Pacific Ocean. Since there could only have been one pole position at 200 Ma, this evidence strongly supported the idea that North America and Europe had moved relative to each other since 200 Ma. Subsequent paleomagnetic work showed that South America, Africa, India, and Australia also have unique polar wandering curves. Rearranging the continents based on their positions in Pangaea caused these wandering curves to overlap, showing that the continents had moved over time.

Additional evidence for movement of the continents came from analysis of **magnetic dip**. Recall from Figure 4.2.1 that the angle of the magnetic field changes as a function of latitude, with the field directed vertically downwards at the north pole, upwards at the south pole, and horizontal at the equator. Every latitude between the equator and the poles will have a corresponding angle between horizontal and vertical (red arrows, Figure 4.2.1). By looking at the dip angle in rocks, we can determine the latitude at which those rocks were formed. Combining that with the age of the rocks, we can trace the movements of the continents over time. For example, at around 500 Ma, what we now call Europe was south of the equator, and so European rocks formed then would have acquired an upward-pointing magnetic field orientation (Figure 4.2.3). Between then and now, Europe gradually moved north, and the rocks forming at various times acquired steeper and steeper *downward-pointing* magnetic orientations.

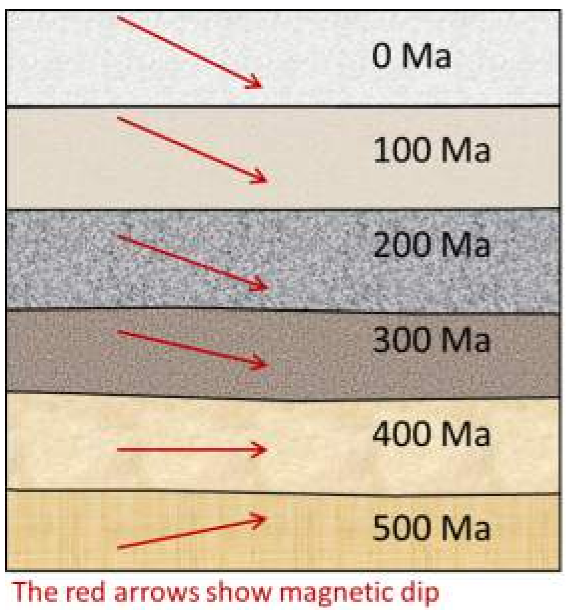


Figure 4.2.3 *Hypothetical magnetic dip angles from layers of rock. This rock would have been south of the equator 500 Ma, at the equator 400 Ma, and since then has been moving further north (Steven Earle, "Physical Geology").*

This paleomagnetic work of the 1950s was the first new evidence in favor of continental drift, and it led a number of geologists to start thinking that the idea might have some merit.

4.3 Mechanisms for Plate Motion

*Modified from "Physical Geology" by Steven Earle**

In section 4.1 we learned that one of the reasons that Wegener's ideas of continental drift were initially rejected by the scientific community was that he could not provide a plausible mechanism for plate motion. However, with all that we have learned about the processes occurring in the Earth's interior since then, there is still some debate about the actual forces that make the plates move. One side in the argument holds that the plates are only moved by the traction caused by mantle convection. The other side holds that traction plays only a minor role and that two other forces, ridge push and slab pull, are more important. Some argue that the real answer lies somewhere in between.

To understand **mantle convection**, imagine a pot of water on a hot stove. The water at the bottom of the pot near the heat source becomes hot and expands, making it lighter (less dense) than the water above. The hot, low density water rises, and cooler, denser water sinks and flows in from the sides. This water then gets heated and rises, and the cycle continues. This creates a circular pattern of rising and sinking water called a **convection cell**. (To test this, try sprinkling a few flakes of spice in the center of a rapidly boiling pot of water. The flakes will move outwards to the edge of the pot as warmer water rises and pushes them aside).

Heat is continuously flowing outward from Earth's interior, and the transfer of heat from the core to the mantle causes convection in the mantle (Figure 4.3.1). Even though the mantle material is essentially solid rock, it is sufficiently plastic (fluid) to slowly flow (at rates of centimeters per year) as long as a steady force is applied to it. This convection is a driving force for the movement of tectonic plates, as the horizontal movements of mantle under the crust drag the plates with them. At places where convection currents in the mantle are moving upward, new lithosphere forms and the plates move apart (diverge). Where two plates are converging (and the convective flow is downward), one plate will be **subducted** (pushed down) into the mantle beneath the other.

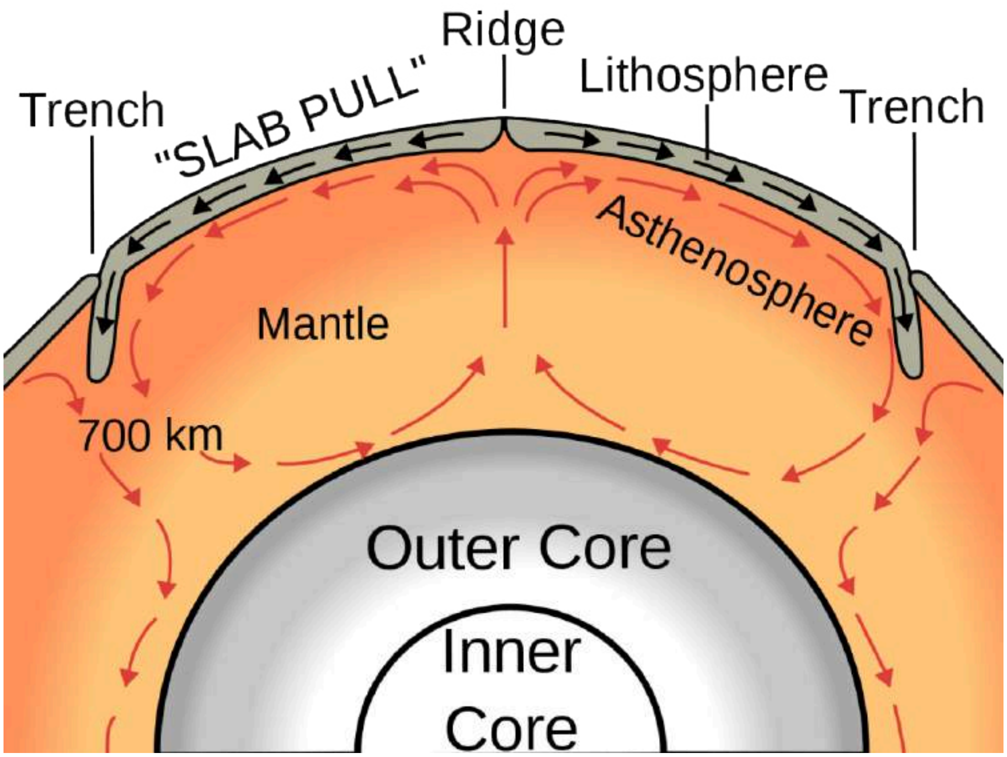


Figure 4.3.1 *Convection cells in the mantle (By Surachit [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 3.0], via Wikimedia Commons).*

The **ridge push/slab pull** model also relies on mantle convection, but in this case it is not simply the traction from the convection cell that moves the plates. In this model, plates move through a combination of pull from the weight of the subducting edge of the plates, and through the outward pushing of an ocean ridge where magma is rising and forming new crust (Figure 4.3.2).

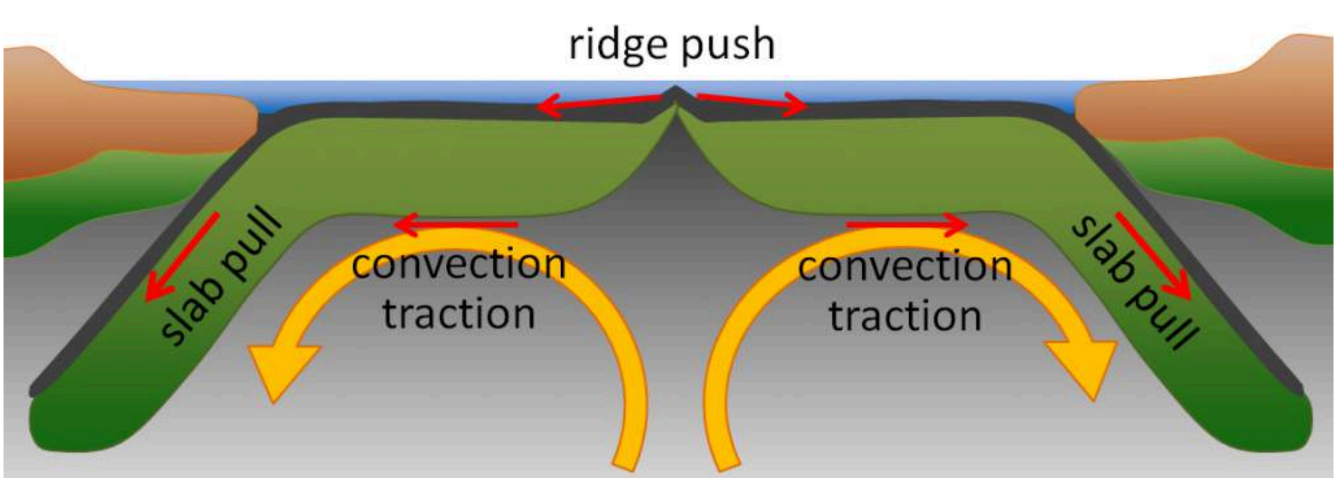


Figure 4.3.2 *Models for plate motion mechanisms (Steven Earle, "Physical Geology").*

Some compelling arguments in favor of the ridge-push/slab-pull model are as follows: (a) plates that are attached to subducting slabs (e.g., Pacific, Australian, and Nazca Plates) move the fastest, and plates that are not (e.g., North American, South American, Eurasian, and African Plates) move significantly slower; (b) in order for the traction model to apply, the mantle would have to be moving about five times faster than the plates are moving (because the coupling between the partially liquid asthenosphere and the plates is not strong), and such high rates of convection are not supported by geophysical models; and (c) although large plates have potential for much higher convection traction, plate velocity is not related to plate area. Although ridge-push/slab-pull is the favored mechanism for plate motion, it's important not to underestimate the role of mantle convection. Without convection, there would be no ridges to push from because upward convection brings hot buoyant rock to surface. Furthermore, many plates, including our own North American Plate, move along nicely — albeit slowly — without any slab-pull happening.

Additional links for more information:

- Animation speculating on the movements of the continents over the last 3 billion years…. https://www.youtube.com/watch?v=UwWWuttntio
- And what might happen over the next 300 million years: https://www.youtube.com/watch?v=bQywDr-btz4

4.4 Plates and Plate Motions

*Modified from "Physical Geology" by Steven Earle**

The idea of plate tectonics became widely accepted around 1965 as more and more geologists started thinking in these terms. By the end of 1967, Earth's surface had been mapped into a series of plates (Figure 4.4.1). The major plates are Eurasia, Pacific, India, Australia, North America, South America, Africa, and Antarctic. There are also numerous small plates (e.g., Juan de Fuca, Nazca, Scotia, Philippine, Caribbean), and many very small plates or sub-plates. For example the Juan de Fuca Plate is actually three separate plates (Gorda, Juan de Fuca, and Explorer) that all move in the same general direction but at slightly different rates.

The fact that the plates include both crustal material and lithospheric mantle material makes it possible for a single plate to be made up of both oceanic and continental crust. For example, the North American Plate includes most of North America, plus half of the northern Atlantic Ocean. Similarly the South American Plate extends across the western part of the southern Atlantic Ocean, while the European and African plates each include part of the eastern Atlantic Ocean. The Pacific Plate is almost entirely oceanic, but it does include the part of California west of the San Andreas Fault.

Rates of motions of the major plates range from less than 1 cm/year to over 10 cm/year (for comparison, human fingernails grow at around 6 cm/year). The Pacific Plate is the fastest at over 10 cm/year in some areas, followed by the Australian and Nazca Plates. The North American Plate is one of the slowest, averaging around 1 cm/year in the south up to almost 4 cm/year in the north. Plates move as rigid bodies, so it may seem surprising that the North American Plate can be moving at different rates in different places. The explanation is that plates move in a rotational manner. The North American Plate, for example, rotates counter-clockwise; the Eurasian Plate rotates clockwise.

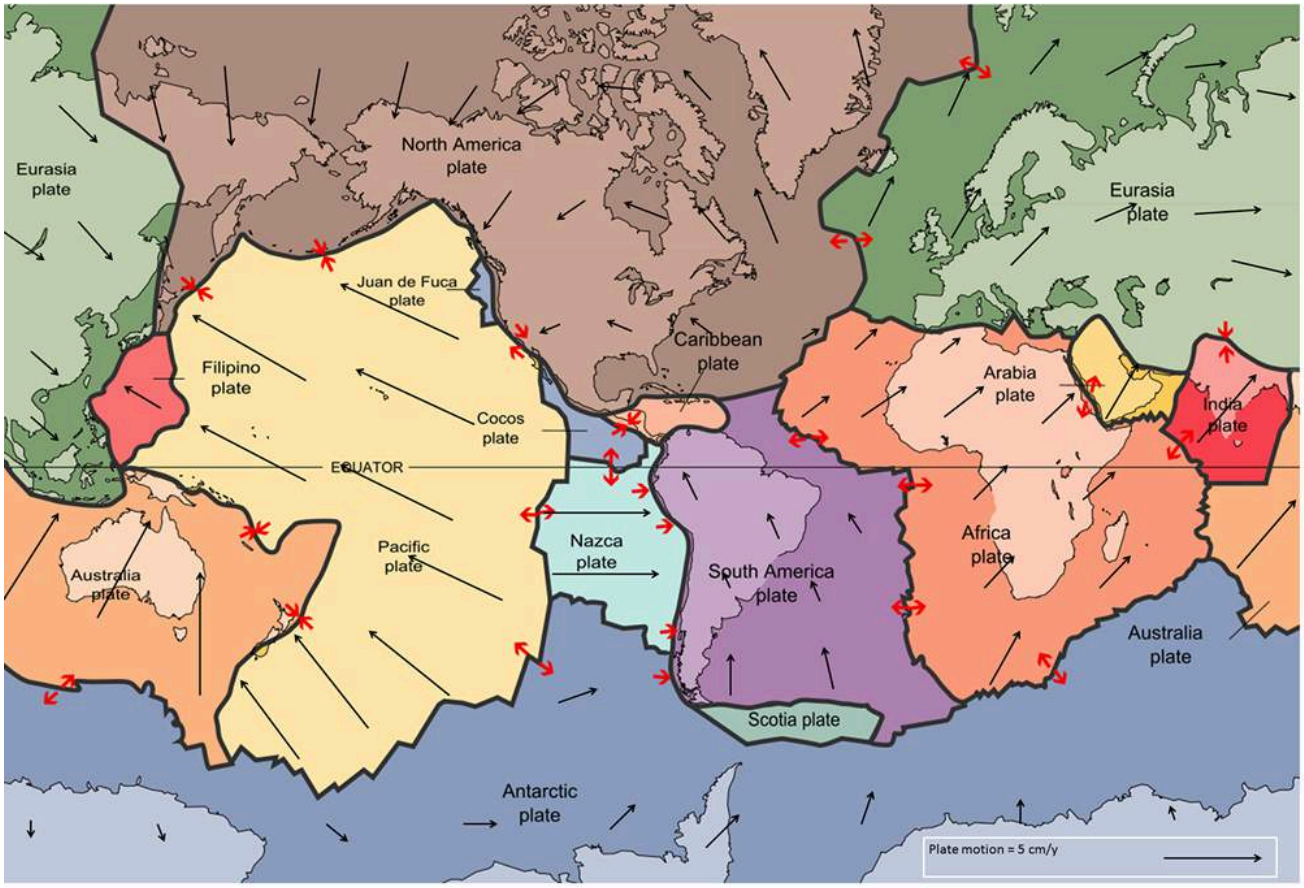


Figure 4.4.1 *The major lithospheric plates of Earth. Arrows indicate direction of plate movement, and the length of the arrows represent the speed of plate motion (Steven Earle, "Physical Geology").*

As originally described by Wegener in 1915, the present continents were once all part of the supercontinent Pangaea. More recent studies of continental match-ups and the magnetic ages of ocean-floor rocks have enabled us to reconstruct the history of the break-up of Pangaea.

Pangaea began to rift apart along a line between Africa and Asia and between North America and South America at around 200 Ma (Figure 4.4.2). During the same period, the Atlantic Ocean began to open up between northern Africa and North America, and India broke away from Antarctica. At this stage, Pangaea was divided into Laurasia (now Europe, Asia and North America) and Gondwanaland (the southern continents; South America, Africa, India, Australia, and Antarctica). Between 200 and 150 Ma, rifting started between South America and Africa and between North America and Europe, and India separated from Antarctica and moved north toward Asia. By 80 Ma, Africa had separated from South America, and most of Europe had separated from North America. By 50 Ma, Australia had separated from Antarctica, and shortly after that, India collided with Asia.

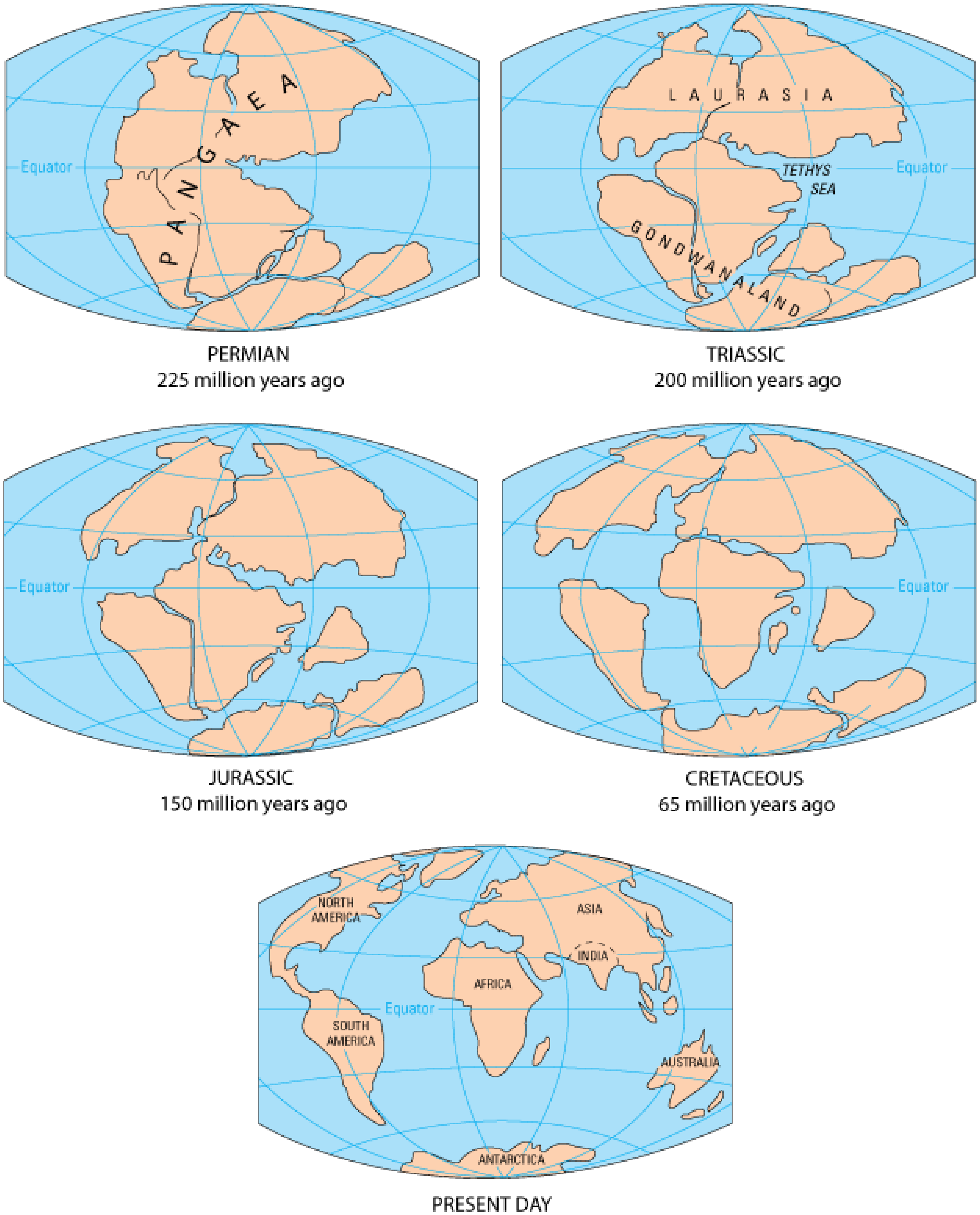


Figure 4.4.2 *Movement of the continents over the past 225 million years (USGS, https://pubs.usgs.gov/gip/dynamic/historical.html).*

Within the past few million years, rifting has taken place in the Gulf of Aden and the Red Sea, and also within the Gulf of California. Incipient rifting has begun along the Great Rift Valley of eastern Africa, extending from Ethiopia and Djibouti on the Gulf of Aden (Red Sea) all the way south to Malawi.

Over the next 50 million years, it is likely that there will be full development of the east African rift and creation of new ocean floor. Eventually Africa will split apart. There will also be continued northerly movement of Australia and Indonesia. The western part of California (including Los Angeles and part of San Francisco) will split away from the rest of North America, and eventually sail right by the west coast of Vancouver Island, en route to Alaska. Because the oceanic crust formed by spreading on the mid-Atlantic ridge is not currently being subducted (except in the Caribbean), the Atlantic Ocean is slowly getting bigger, and the Pacific Ocean is getting smaller. If this continues without changing for another couple hundred million years, we will be back to where we started, with one supercontinent.

Pangaea, which existed from about 350 to 200 Ma, was not the first supercontinent. In 1966, Tuzo Wilson proposed that there has been a continuous series of cycles of continental rifting and collision; that is, break-up of supercontinents, drifting, collision, and formation of other supercontinents. Pangaea was preceded by Pannotia (600 to 540 Ma), by Rodinia (1,100 to 750 Ma), and by other supercontinents before that.

With all of these plates constantly on the move, they inevitably end up interacting with each other at their plate boundaries. Plates can interact in three ways: they can move apart (divergent boundary), they can move towards each other (convergent boundary), or they can slide past each other (transform boundary). The following sections will examine each of these types of plate boundaries, and the geological features they create.

Additional links for more information:

- An interactive animation of plate motion over the past 550 million years: http://barabus.tru.ca/geol1031/plates.html

---

4.5 Divergent Plate Boundaries

*Modified from "Physical Geology" by Steven Earle**

Divergent boundaries are spreading boundaries, where new oceanic crust is created to fill in the space as the plates move apart. Most divergent boundaries are located along mid-ocean oceanic ridges (although some are on land). The **mid-ocean ridge** system is a giant undersea mountain range, and is the largest geological feature on Earth; at 65,000 km long and about 1000 km wide, it covers 23% of Earth's surface (Figure 4.5.1). Because the new crust formed at the plate boundary is warmer than the surrounding crust, it has a lower density so it sits higher on the mantle, creating the mountain chain. Running down the middle of the mid-ocean ridge is a **rift valley** 25-50 km wide and 1 km deep. Although oceanic spreading ridges appear to be curved features on Earth's surface, in fact the ridges are composed of a series of straight-line segments, offset at intervals by faults perpendicular to the ridge, called **transform faults**. These transform faults make the mid-ocean ridge system look like a giant zipper on the seafloor (Figure 4.5.2). As we will see in section 4.7, movements along transform faults between two adjacent ridge segments are responsible for many earthquakes.

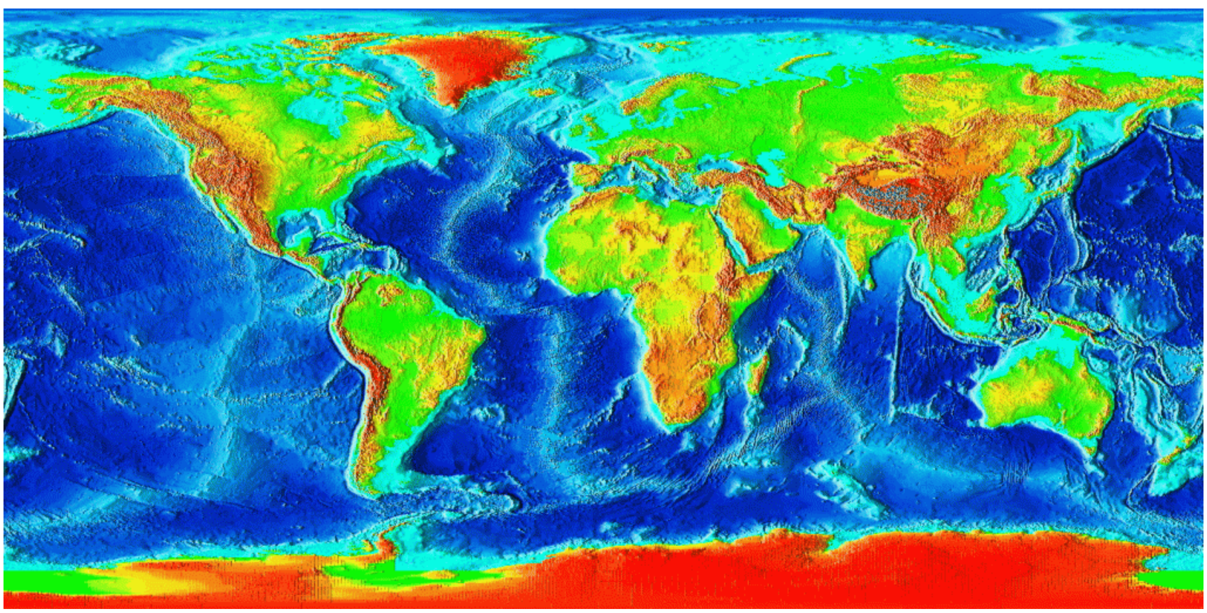


Figure 4.5.1 *Ocean floor topography. The mid-ocean ridge system can be seen as the light blue chain of mountains running throughout the oceans (http://www.ngdc.noaa.gov/mgg/image/mggd.gif).*

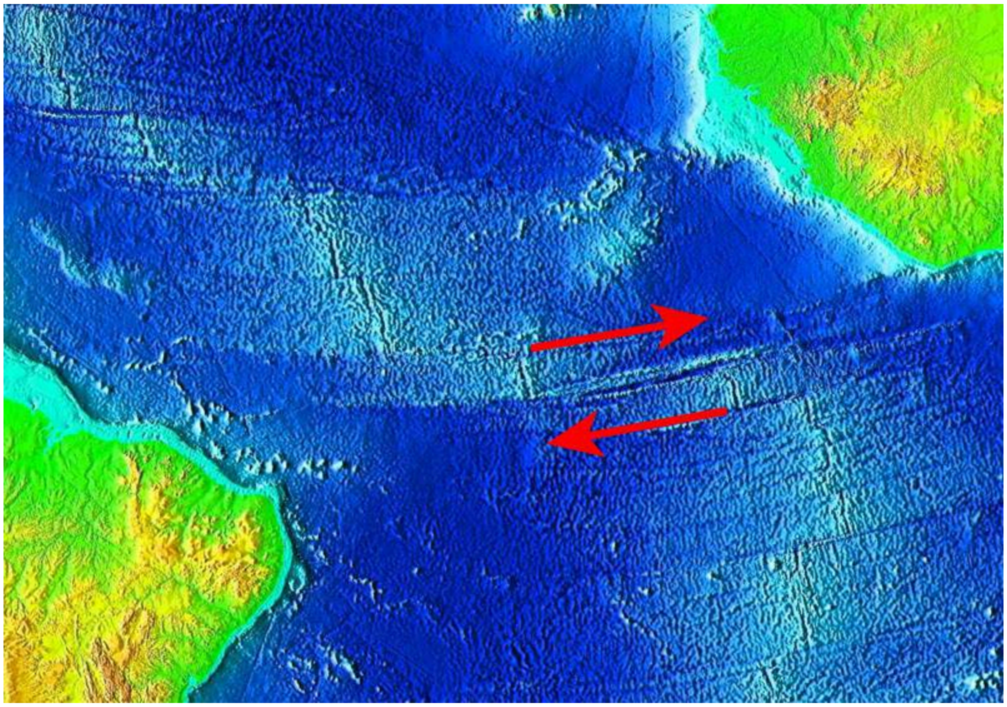


Figure 4.5.2 *Closeup of the mid-Atlantic ridge system, showing transform faults perpendicular to the ridge axis. Arrows indicate the direction of plate motion on either side of the fault (USGS, Public domain, via Wikimedia Commons).*

The crustal material created at a spreading boundary is always oceanic in character; in other words, it is igneous rock (e.g., basalt or gabbro, rich in ferromagnesian minerals), forming from magma derived from partial melting of the mantle caused by decompression as hot mantle rock from depth is moved toward the surface (Figure 4.5.3). The triangular zone of partial melting near the ridge crest is approximately 60 km thick and the proportion of magma is about 10% of the rock volume, thus producing crust that is about 6 km thick. This magma oozes out onto the seafloor to form pillow basalts, breccias (fragmented basaltic rock), and flows, interbedded in some cases with limestone or chert. Over time, the igneous rock of the oceanic crust gets covered with layers of sediment, which eventually become sedimentary rock.

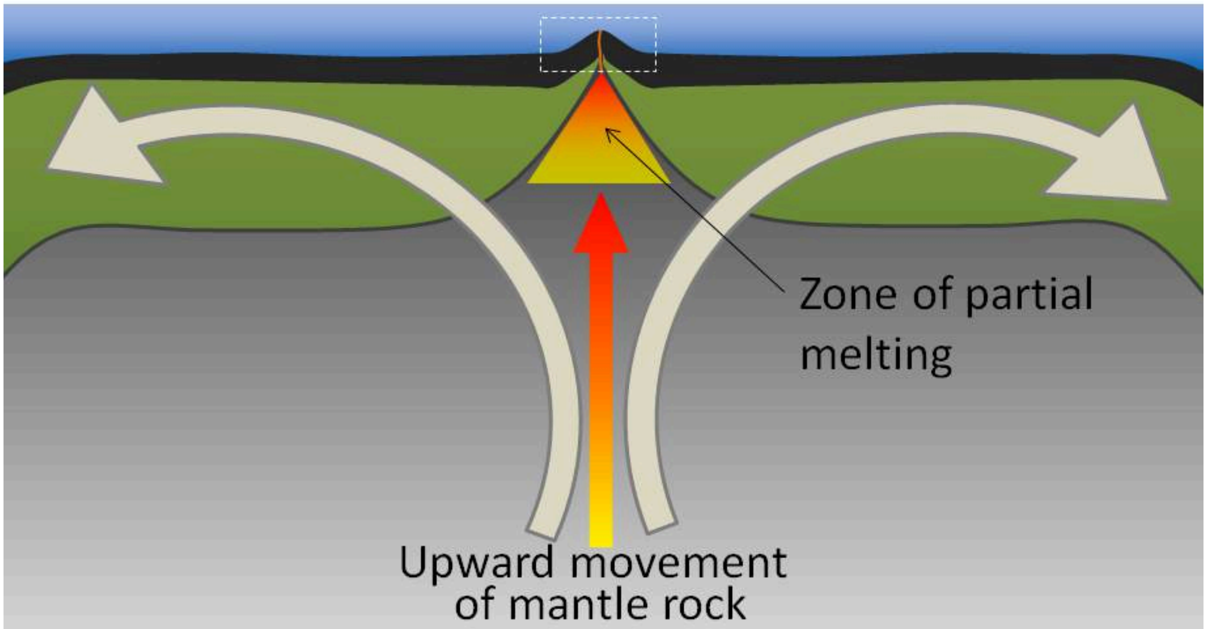


Figure 4.5.3 *Mechanism for divergent plate boundaries. The region in the outlined rectangle represent the mid-ocean ridge (Steven Earle, "Physical Geology").*

Spreading is hypothesized to start within a continental area with up-warping or doming of crust related to an underlying mantle plume or series of mantle plumes. The buoyancy of the mantle plume material creates a dome within the crust, causing it to fracture. When a series of mantle plumes exists beneath a large continent, the resulting rifts may align and lead to the formation of a rift valley (such as the present-day Great Rift Valley in eastern Africa). It is suggested that this type of valley eventually develops into a linear sea (such as the present-day Red Sea), and finally into an ocean (such as the Atlantic). It is likely that as many as 20 mantle plumes, many of which still exist, were responsible for the initiation of the rifting of Pangaea along what is now the mid-Atlantic ridge.

There are multiple lines of evidence demonstrating that new oceanic crust is forming at these seafloor spreading centers:

1. Age of the crust:

Comparing the ages of the oceanic crust near a mid-ocean ridge shows that the crust is youngest right at the spreading center, and gets progressively older as you move away from the divergent boundary in either direction, aging approximately 1 million years for every 20-40 km from the ridge. Furthermore, the pattern of crust age is fairly symmetrical on either side of the ridge (Figure 4.5.4).

The oldest oceanic crust is around 280 Ma in the eastern Mediterranean, and the oldest parts of the open ocean are around 180 Ma on either side of the north Atlantic. It may be surprising, considering that parts of the continental crust are close to 4,000 Ma old, that the oldest seafloor is less than 300 Ma. Of course, the reason for this is that all seafloor older than that has been either subducted (see section 4.6) or pushed up to become part of the continental crust. As one would expect, the oceanic crust is very young near the spreading ridges (Figure 4.5.4), and there are obvious differences in the rate of sea-floor spreading along different ridges. The ridges in the Pacific and southeastern Indian Oceans have wide age bands, indicating rapid spreading (approaching 10 cm/year on each side in some areas), while those in the Atlantic and western Indian Oceans are spreading much more slowly (less than 2 cm/year on each side in some areas).

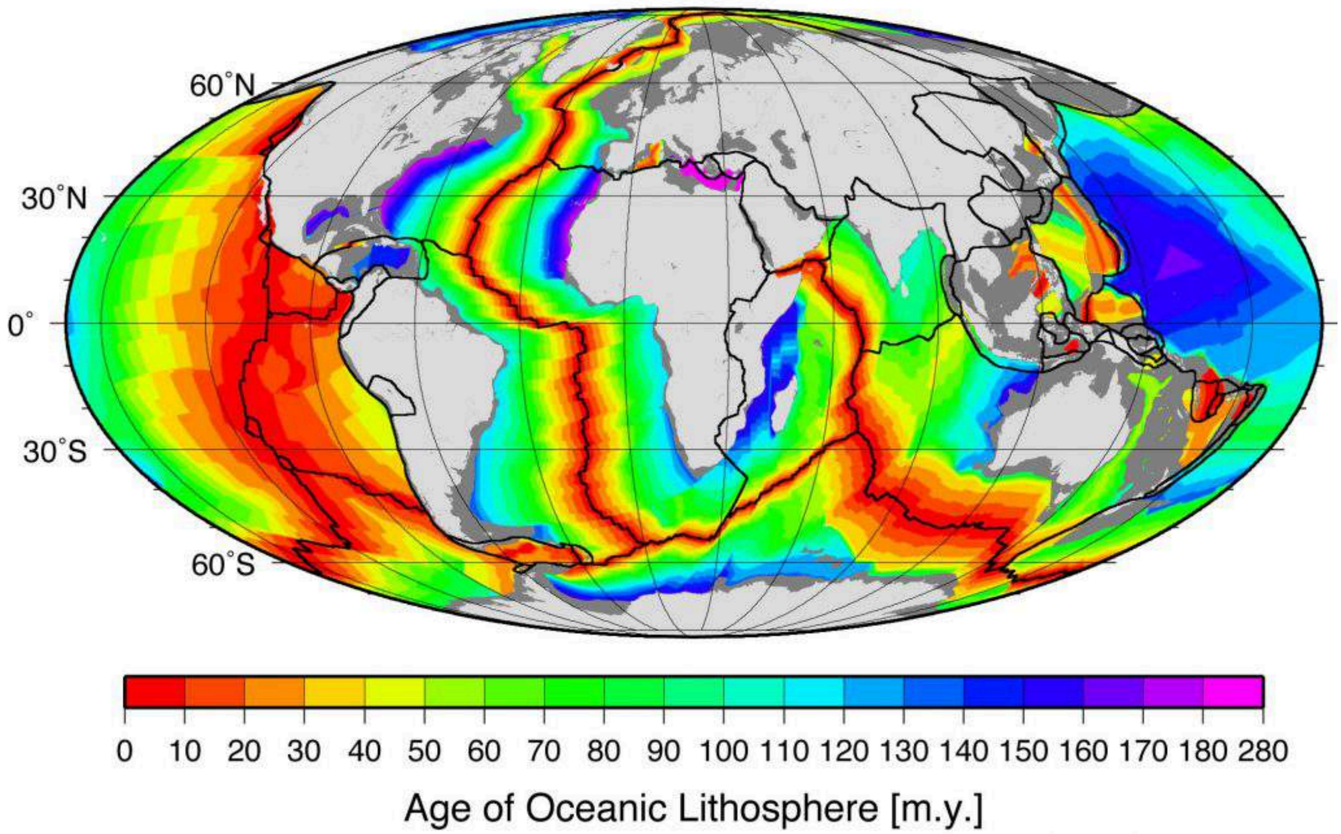


Figure 4.5.4 *Age of the oceanic crust (http://www.ngdc.noaa.gov/mgg/ocean_age/data/2008/image/age_oceanic_lith.jpg).*

2. Sediment thickness:

With the development of seismic reflection sounding (similar to echo sounding described in section 1.4) it became possible to *see through* the seafloor sediments and map the bedrock topography and crustal thickness. Hence sediment thicknesses could be mapped, and it was soon discovered that although the sediments were up to several thousands of meters thick near the continents, they were relatively thin — or even non-existent — in the ocean ridge areas (Figure 4.5.5). This makes sense when combined with the data on the age of the oceanic crust; the farther from the spreading center the older the crust, the longer it has had to accumulate sediment, and the thicker the sediment layer. Additionally, the bottom layers of sediment are older the farther you get from the ridge, indicating that they were deposited on the crust long ago when the crust was first formed at the ridge.

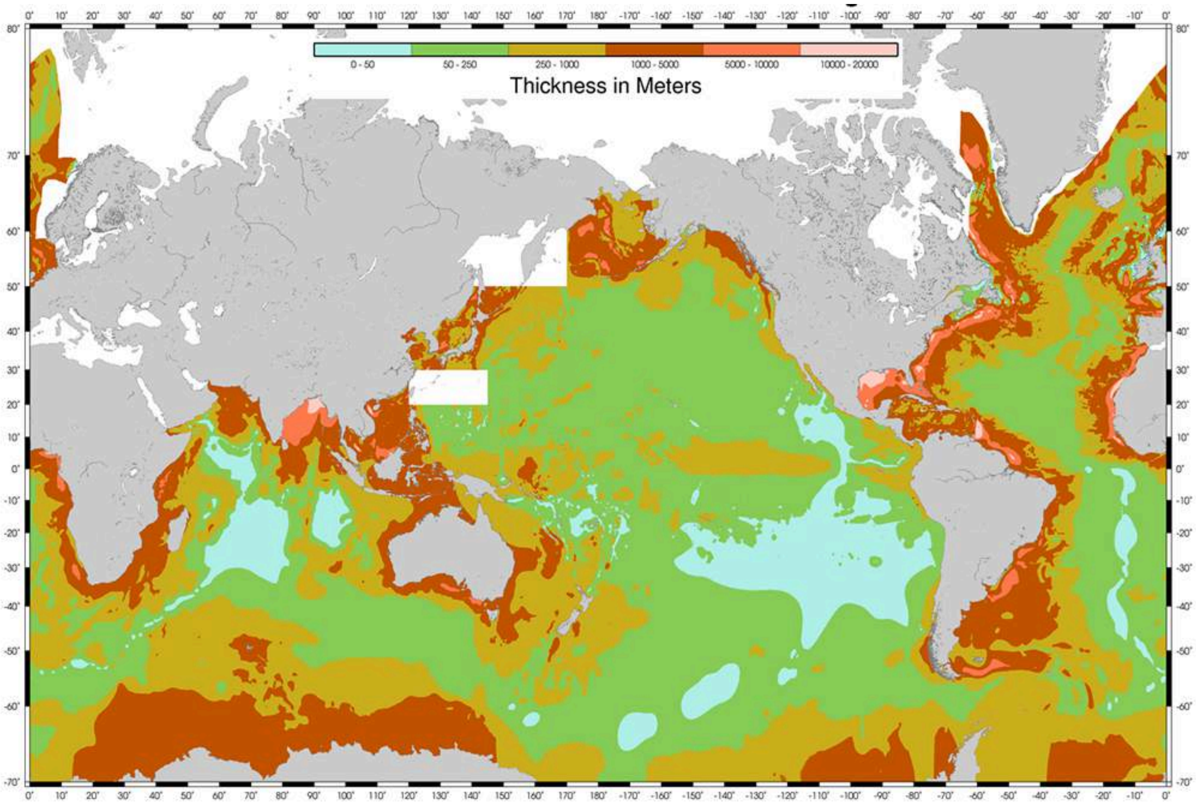


Figure 4.5.5 *Seafloor sediment thickness (Modified from https://www.ngdc.noaa.gov/mgg/sedthick/).*

3. Heat flow:

Measurements of rates of heat flow through the ocean floor revealed that the rates are higher than average (about 8x higher) along the ridges, and lower than average in the trench areas (about 1/20th of the average). The areas of high heat flow are correlated with upward convection of hot mantle material as new crust is formed, and the areas of low heat flow are correlated with downward convection at subduction zones.

4. Magnetic reversals:

In section 4.2 we saw that rocks could retain magnetic information that they acquired when they were formed. However, Earth's magnetic field is not stable over geological time. For reasons that are not completely understood, the magnetic field decays periodically and then becomes re-established. When it does re-establish, it may be oriented the way it was before the decay, or it may be oriented with the reversed polarity. During periods of reversed polarity, a compass would point south instead of north. Over the past 250 Ma, there have a few hundred magnetic field reversals, and their timing has been anything but regular. The shortest ones that geologists have been able to define lasted only a few thousand years, and the longest one was more than 30 million years, during the Cretaceous (Figure 4.5.6). The present "normal" event has persisted for about 780,000 years.

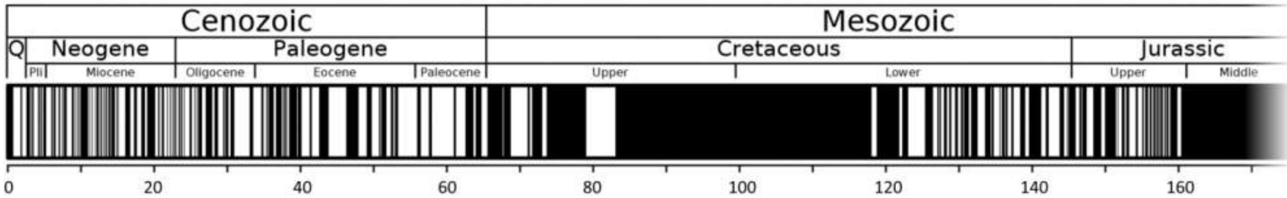


Figure 4.5.6 Magnetic field reversal chronology for the past 170 Ma (Steven Earle after: http://upload.wikimedia.org/ wikipedia/en/c/c0/Geomagnetic_polarity_0-169_Ma.svg).

Beginning in the 1950s, scientists started using magnetometer readings when studying ocean floor topography. The first comprehensive magnetic data set was compiled in 1958 for an area off the coast of British Columbia and Washington State. This survey revealed a mysterious pattern of alternating stripes of low and high magnetic intensity in sea-floor rocks (Figure 4.5.7). Subsequent studies elsewhere in the ocean also observed these magnetic anomalies, and most importantly, the fact that the magnetic patterns are symmetrical with respect to ocean ridges. In the 1960s, in what would become known as the Vine-Matthews-Morley (VMM) hypothesis, it was proposed that the patterns associated with ridges were related to the magnetic reversals, and that oceanic crust created from cooling basalt during a *normal* event would have polarity aligned with the present magnetic field, and thus would produce a positive anomaly (a black stripe on the sea-floor magnetic map), whereas oceanic crust created during a *reversed* event would have polarity opposite to the present field and thus would produce a negative magnetic anomaly (a white stripe). The widths of the anomalies varied according to the spreading rates characteristic of the different ridges. This process is illustrated in Figure 4.5.8. New crust is formed (panel a) and takes on the existing normal magnetic polarity. Over time, as the plates continue to diverge, the magnetic polarity reverses, and new crust formed at the ridge now takes on the reversed polarity (white stripes in Figure 4.5.8). In panel b, the poles have reverted to normal, so once again the new crust shows normal polarity before moving away from the ridge. Eventually, this creates a series of parallel, alternating bands of reversals, symmetrical around the spreading center (panel c).

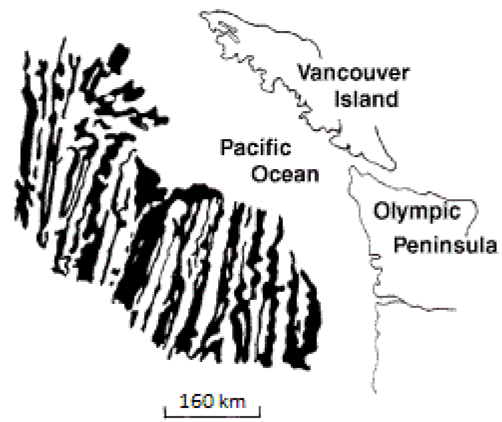


***Figure 4.5.7** Pattern of magnetic anomalies in oceanic crust in the Pacific northwest (Steven Earle, "Physical Geology").*

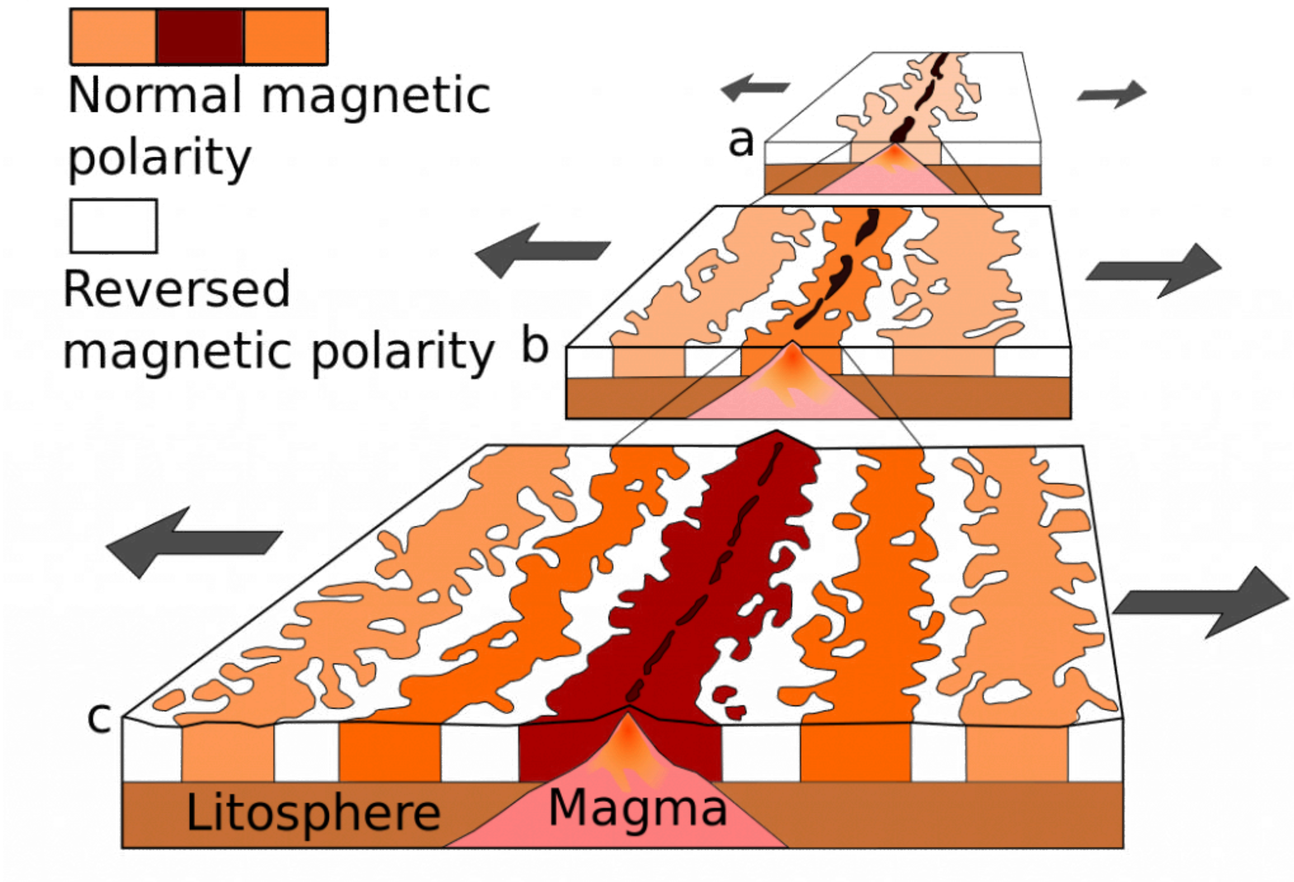


Figure 4.5.8 *Formation of alternating patterns of magnetic polarity along a mid-ocean ridge (Steven Earle, "Physical Geology").*

4.6 Convergent Plate Boundaries

*Modified from "Physical Geology" by Steven Earle**

Convergent boundaries, where two plates are moving toward each other, are of three types, depending on the type of crust present on either side of the boundary — oceanic or continental. The types are ocean-ocean, ocean-continent, and continent-continent.

At an ocean-ocean convergent boundary, one of the plates (oceanic crust and lithospheric mantle) is pushed, or subducted**,** under the other (Figure 4.6.1). Often it is the older and colder plate that is denser and subducts beneath the younger and warmer plate. There is commonly an ocean trench along the boundary as the crust bends downwards. The subducted lithosphere descends into the hot mantle at a relatively shallow angle close to the subduction zone, but at steeper angles farther down (up to about 45°). The significant volume of water within the subducting material is released as the subducting crust is heated. It mixes with the overlying mantle, and the addition of water to the hot mantle lowers the crust's melting point and leads to the formation of magma (flux melting). The magma, which is lighter than the surrounding mantle material, rises through the mantle and the overlying oceanic crust to the ocean floor where it creates a chain of volcanic islands known as an **island arc**. A mature island arc develops into a chain of relatively large islands (such as Japan or Indonesia) as more and more volcanic material is extruded and sedimentary rocks accumulate around the islands. Earthquakes occur relatively deep below the seafloor, where the subducting crust moves against the overriding crust.

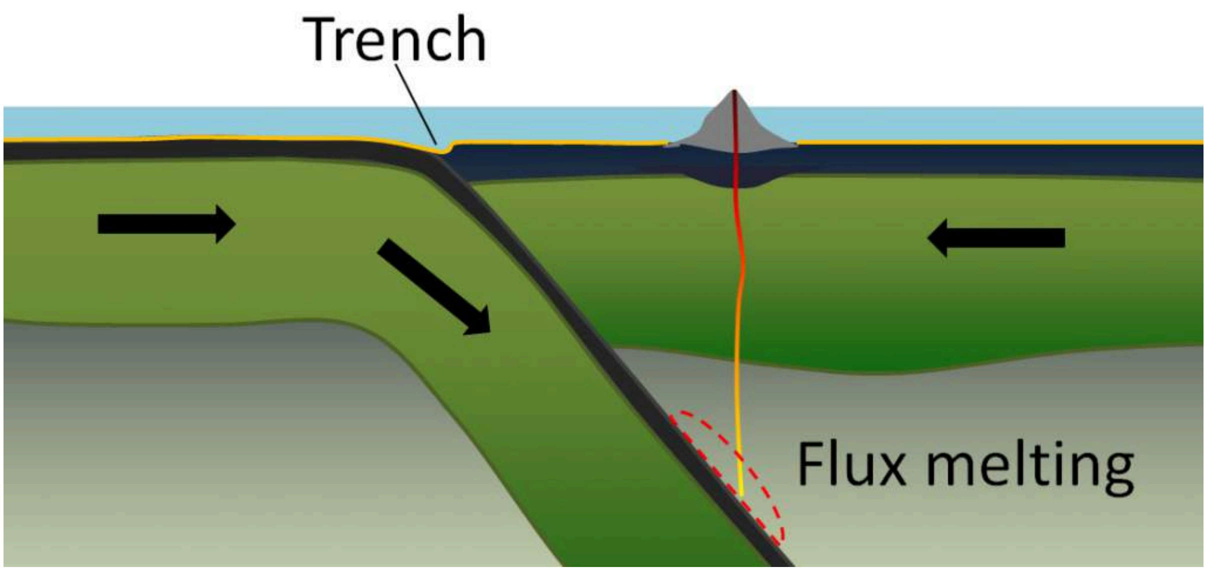


***Figure 4.6.1** A trench and volcanic island formed from an ocean-ocean convergent zone (Steven Earle, "Physical Geology").*

Examples of ocean-ocean convergent zones are subduction of the Pacific Plate south of Alaska (creating the Aleutian Islands) and under the Philippine Plate, where it creates the Marianas Trench, the deepest part of the ocean.

At an ocean-continent convergent boundary, the denser oceanic plate is pushed under the less dense continental plate in the same manner as at an ocean-ocean boundary. Sediment that has accumulated on the seafloor is thrust up into an accretionary wedge, and compression leads to thrusting within the continental plate (Figure 4.6.2). The magma produced adjacent to the subduction zone rises to the base of the continental crust and leads to partial melting of the crustal rock. The resulting magma ascends through the crust, producing a mountain chain with many volcanoes. As with an ocean-ocean boundary, the subducting crust can produce a deep trench running parallel to the coastline.

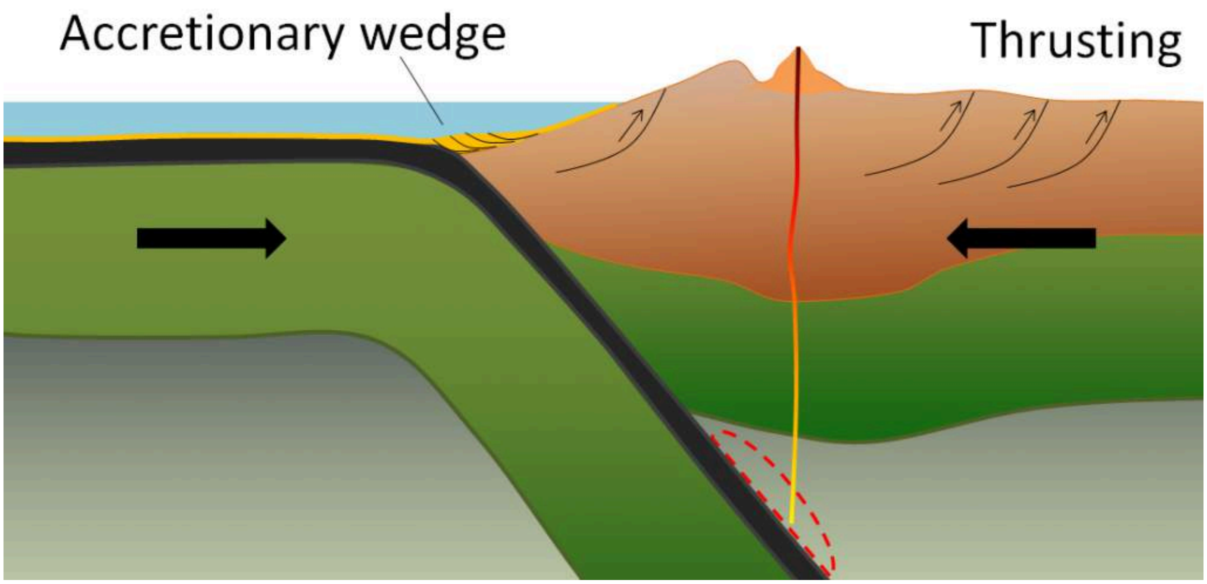


Figure 4.6.2 *A trench and volcanic mountains formed from an ocean-continent convergent zone (Steven Earle, "Physical Geology").*

Examples of ocean-continent convergent boundaries are subduction of the Nazca Plate under South America (which has created the Andes Mountains and the Peru Trench) and subduction of the Juan de Fuca Plate under North America (creating the Cascade Range).

A continent-continent collision occurs when a continent or large island that has been moved along with subducting oceanic crust collides with another continent (Figure 4.6.3). The colliding continental material will not be subducted because it is too light (i.e., because it is composed largely of light continental rocks), but the root of the oceanic plate will eventually break off and sink into the mantle. There is tremendous deformation of the pre-existing continental rocks, forcing the material upwards and creating mountains.

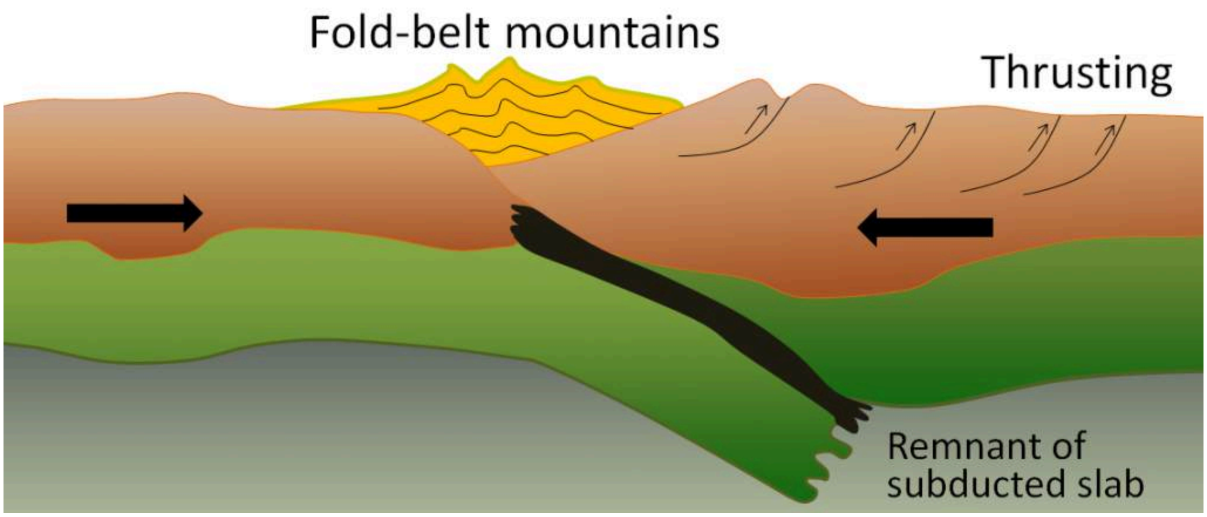


Figure 4.6.3 *Mountains formed from a continent-continent convergent zone (Steven Earle, "Physical Geology").*

Examples of continent-continent convergent boundaries are the collision of the India Plate with the Eurasian Plate, creating the Himalaya Mountains, and the collision of the African Plate with the Eurasian Plate, creating the series of ranges extending from the Alps in Europe to the Zagros Mountains in Iran. The Rocky Mountains in North America are also a result of continent-continent collisions.

4.7 Transform Plate Boundaries

*Modified from "Physical Geology" by Steven Earle**

Transform boundaries exist where one plate slides past another without production or destruction of crustal material. As explained in [section 4.5](), most transform faults connect segments of mid-ocean ridges and are thus ocean-ocean plate boundaries. Some transform faults connect continental parts of plates. An example is the San Andreas Fault, which connects the southern end of the Juan de Fuca Ridge with the northern end of the East Pacific Rise (ridge) in the Gulf of California (Figure 4.7.1). The part of California west of the San Andreas Fault and all of Baja California are on the Pacific Plate. Transform faults do not just connect divergent boundaries. For example, the Queen Charlotte Fault connects the north end of the Juan de Fuca Ridge, starting at the north end of Vancouver Island, to the Aleutian subduction zone.

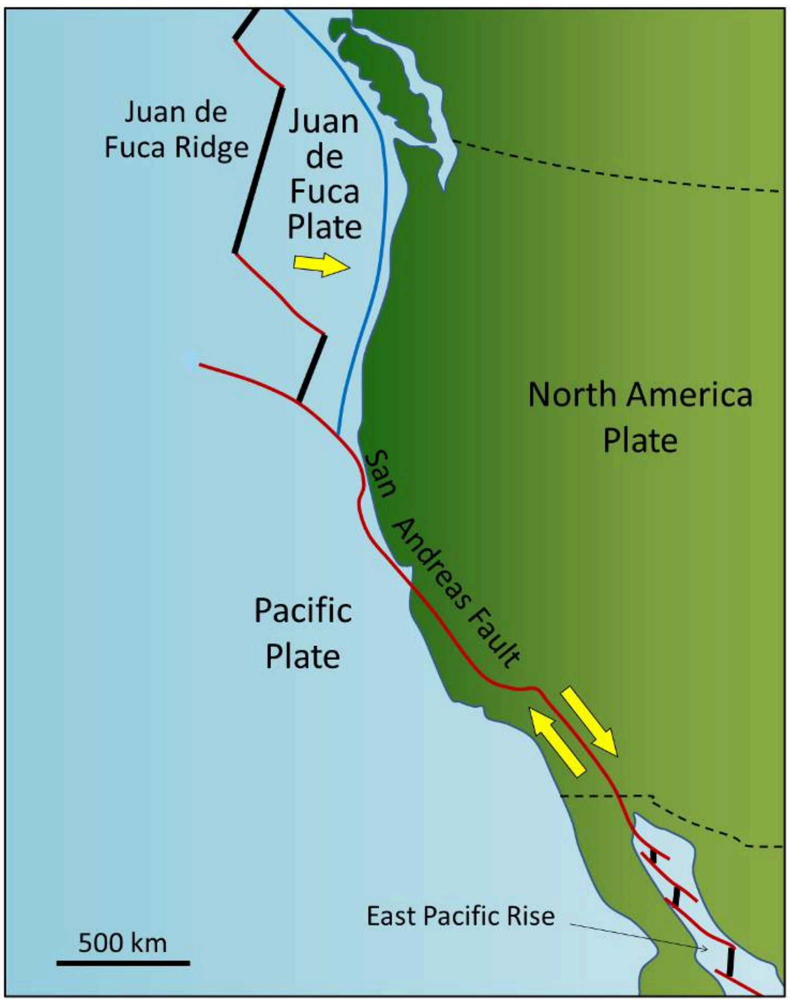


Figure 4.7.1 *Transform faults along the U.S. west coast (Steven Earle, "Physical Geology").*

As we will see in the next section, earthquakes are common along transform faults, as the two plates slide past each other.

4.8 Earthquakes and Plate Tectonics

*Modified from "Physical Geology" by Steven Earle**

An earthquake is the shaking caused by the rupture (breaking) and subsequent displacement of rocks (one body of rock moving with respect to another) beneath Earth's surface.

A body of rock that is under stress becomes deformed. When the rock can no longer withstand the deformation, it breaks and the two sides slide past each other. Because most rock is strong (unlike loose sand, for example), it can withstand a significant amount of deformation without breaking. But every rock has a deformation limit and will rupture (break) once that limit is reached. At that point, in the case of rocks within the crust, the rock breaks and there is displacement along the rupture surface. The magnitude of the earthquake depends on the extent of the area that breaks (the area of the rupture surface) and the average amount of displacement (sliding).

Most earthquakes take place near plate boundaries, but not necessarily right on a boundary, and not necessarily even on a pre-existing fault. The distribution of earthquakes across the globe is shown in Figure 4.8.1. It is relatively easy to see the relationships between earthquakes and the plate boundaries. Along divergent boundaries like the mid-Atlantic ridge and the East Pacific Rise, earthquakes are common, but restricted to a narrow zone close to the ridge, and consistently at less than 30 km depth. Shallow earthquakes are also common along transform faults, such as the San Andreas Fault. Along subduction zones earthquakes are very abundant, and they are increasingly deep on the landward side of the subduction zone.

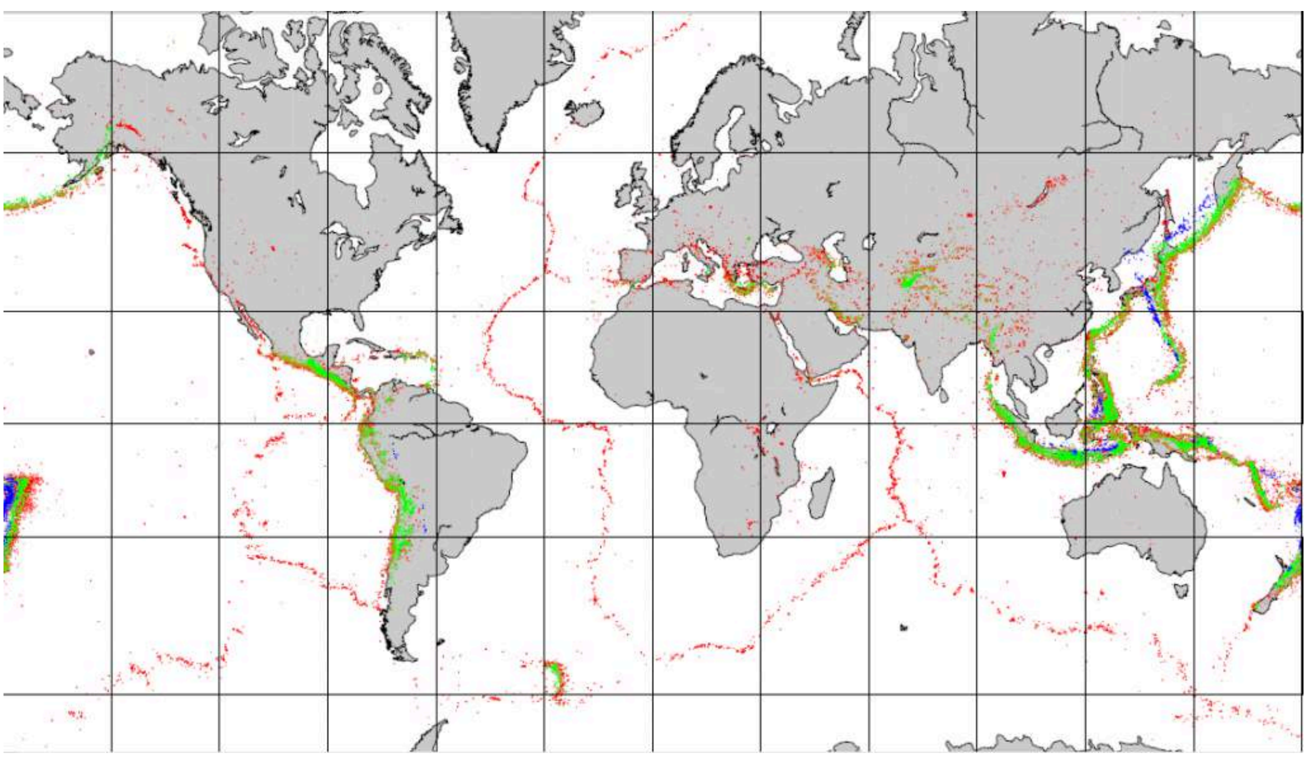


Figure 4.8.1 *Global distribution of earthquakes. Red dots indicate shallow earthquakes (<33 km deep), green and blue indicate deep earthquakes (Steven Earle, "Physical Geology").*

Earthquakes are also relatively common at a few intraplate locations. Some are related to the buildup of stress due to continental rifting or the transfer of stress from other regions, and some are not well understood. Examples of intraplate earthquake regions include the Great Rift Valley area of Africa, the Tibet region of China, and the Lake Baikal area of Russia.

Earthquakes at Divergent and Transform Boundaries

Figure 4.8.2 provides a closer look at magnitude (M) 4 and larger earthquakes in an area of divergent boundaries in the mid-Atlantic region near the equator. Here, as we saw in section 4.5, the segments of the mid-Atlantic ridge are offset by some long transform faults. Most of the earthquakes are located along the transform faults, rather than along the spreading segments, although there are clusters of earthquakes at some of the ridge-transform boundaries. Some earthquakes do occur on spreading ridges, but they tend to be small and infrequent because of the relatively high rock temperatures in the areas where spreading is taking place. Earthquakes along divergent and transform boundaries tend to be shallow, as the crust is not very thick.

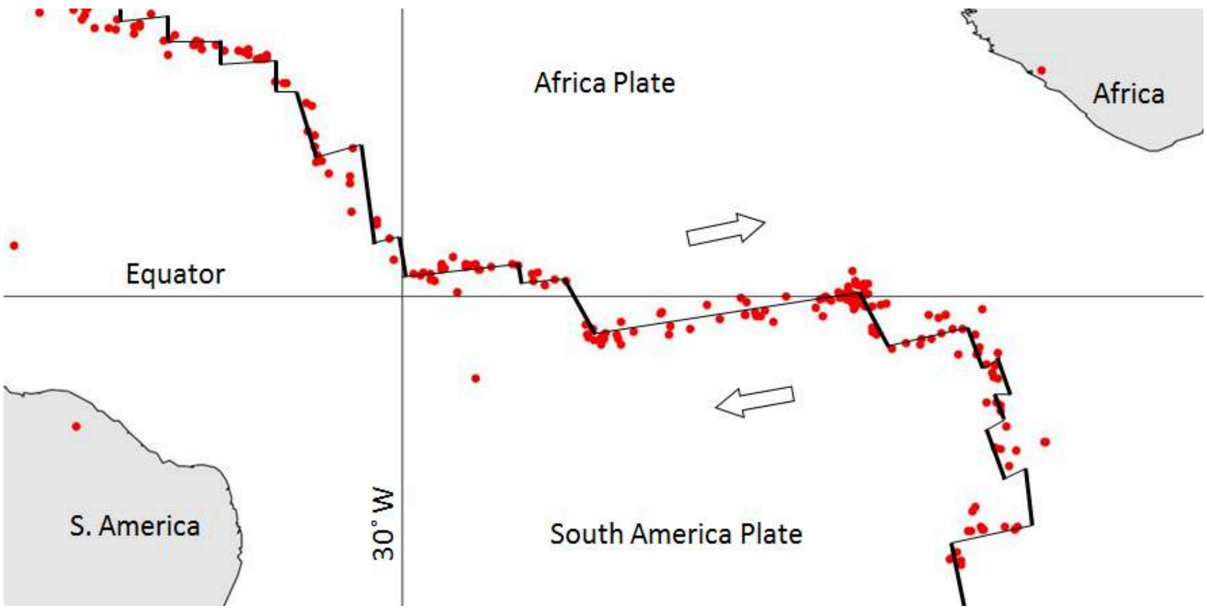


Figure 4.8.2 *Earthquake activity along the mid Atlantic ridge (Steven Earle, "Physical Geology").*

Earthquakes at Convergent Boundaries

The distribution and depths of earthquakes in the North Pacific are shown in Figure 4.8.3. In this region, the Pacific Plate is subducting beneath the North America Plate, creating the Aleutian Trench and the Aleutian Islands. Shallow earthquakes are common along the trench, but there is also significant earthquake activity extending down several hundred kilometers, as the subducting plate continues to interact at depth with the overriding plate. The earthquakes get deeper with distance from the trench; note in the left panel in Figure 4.8.3 that as you move along the transect from point a to point b, there is a trend of increasing earthquake depth. This reveals that it is the Pacific Plate that is moving northwards and being subducted.

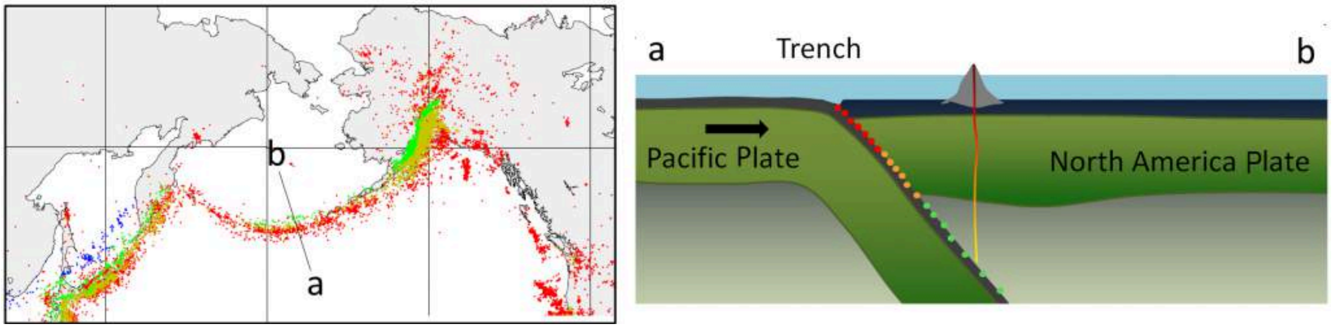


Figure 4.8.3 *Earthquake activity along a convergent boundary at the Aleutian Islands. Red dots indicate shallow earthquakes, green and blue indicate deeper earthquakes (Steven Earle, "Physical Geology").*

The distribution of earthquakes in the area of the India-Eurasia plate boundary is shown in Figure 4.8.4. This is a continent-continent convergent boundary, and it is generally assumed that although the India Plate continues to move north toward the Asia Plate, there is no actual subduction taking place. There are transform faults on either side of the India Plate in this area.

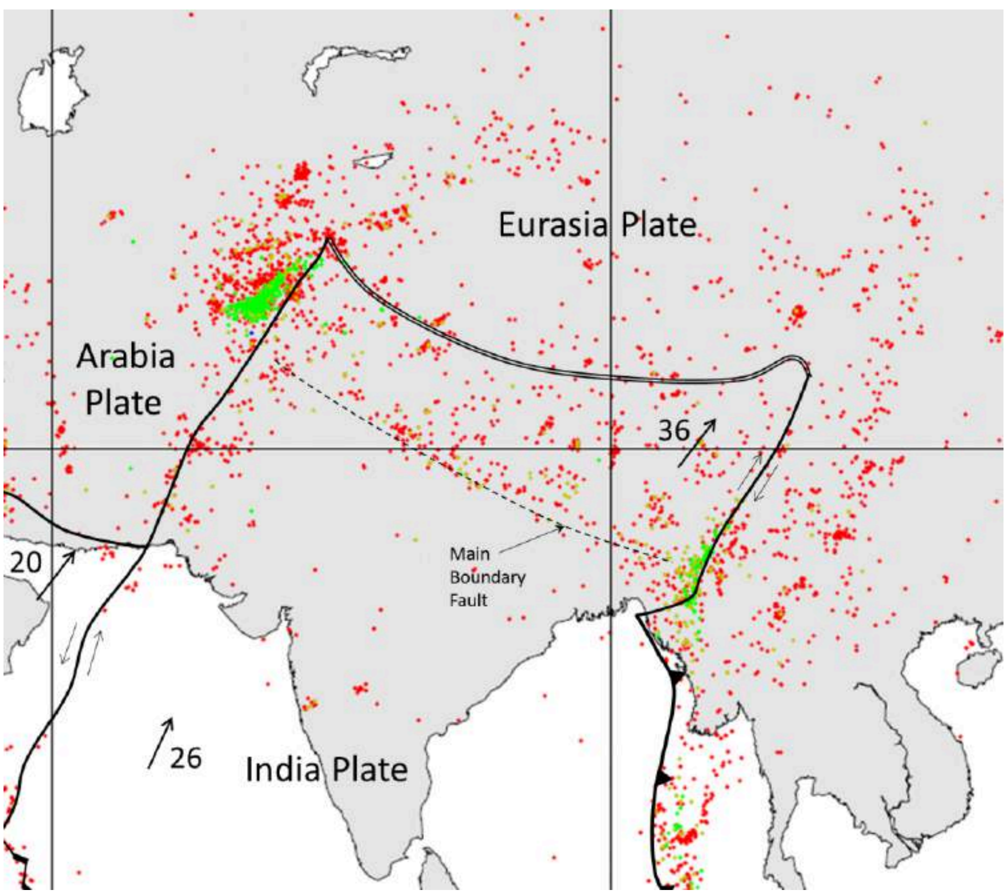


Figure 4.8.4 *The distribution of earthquakes in the area of the India-Eurasia plate boundary (Steven Earle, "Physical Geology").*

The entire northern India and southern Asia region is very seismically active. Earthquakes are common in northern India, Nepal, Bhutan, Bangladesh and adjacent parts of China, and throughout Pakistan and Afghanistan. Many of the earthquakes are related to the transform faults on either side of the India Plate, and most of the others are related to the significant tectonic squeezing caused by the continued convergence of the India and Asia Plates. That squeezing has caused the Asia Plate to be thrust over top of the India Plate, building the Himalayas and the Tibet Plateau to enormous heights.

4.9 Seamounts and Hot Spots

*Modified from "Physical Geology" by Steven Earle**

The ocean floor is dotted with **seamounts**, some isolated and some in chains. Seamounts are underwater volcanoes, and most are much younger than the oceanic crust on which they formed. If a seamount gets large enough to break the ocean surface, it becomes a volcanic island. Some seamounts are formed from magma rising at a divergent boundary, and as the plates move apart, the seamounts move with them, which can result in a seamount chain. Other seamounts form from the rising magma at an ocean-ocean subduction zone; these include the Aleutians, extending from Alaska to Russia, and the Lesser Antilles in the eastern part of the Caribbean. Sometimes the crust on which an island or seamount sits will subside, taking the seamount with it. As this happens, the top of the seamount can become eroded flat, and these flat-topped seamounts are then called **tablemounts** or **guyots**.

However, some seamounts are formed far away from plate boundaries, in places where we would not usually expect much volcanic activity. Some seamounts and ocean islands are formed above a **mantle plume** or **hot spot** — a place where hot mantle material rises in a stationary and semi-permanent plume, and affects the overlying crust. Mantle plumes are thought to rise at approximately 10 times the rate of mantle convection. The ascending column may be on the order of kilometers to tens of kilometers across, but near the surface it spreads out to create a mushroom-style head that is several tens to over 100 kilometers across. Near the base of the lithosphere (the rigid part of the mantle), the mantle plume (and possibly some of the surrounding mantle material) partially melts to form magma that rises to feed volcanoes.

A great example of seamounts created from a hot spot includes the Hawaiian and Emperor Seamount island chains in the Pacific Ocean (Figure 4.9.1). The oldest of the Hawaiian/Emperor seamounts is dated at around 80 Ma, and it is situated on oceanic crust aged around 90 to 100 Ma. The volcanic rock making up these islands gets progressively younger toward the southeast, culminating with the island of Hawaii itself, which consists of rock that is almost all younger than 1 Ma. It appears that a stationary plume of hot upwelling mantle material is the source of the Hawaiian volcanism, and that the ocean crust of the Pacific Plate is moving toward the northwest over this hot spot. A seamount will be formed through volcanic activity over the hot spot, then the plate will move and displace the seamount before the hot spot produces the next seamount, and so on. In this way, over time, the seamounts are formed in chains. Near the Midway Islands, the chain takes a pronounced change in direction, from northwest-southeast for the Hawaiian Islands to nearly north-south for the Emperor Seamounts. This change is widely ascribed to a change in direction of the Pacific Plate moving over the stationary mantle plume, but it is also possible that the Hawaiian mantle plume has not actually been stationary throughout its history, and in fact moved at least 2,000 km south over the period between 81 and 45 Ma.

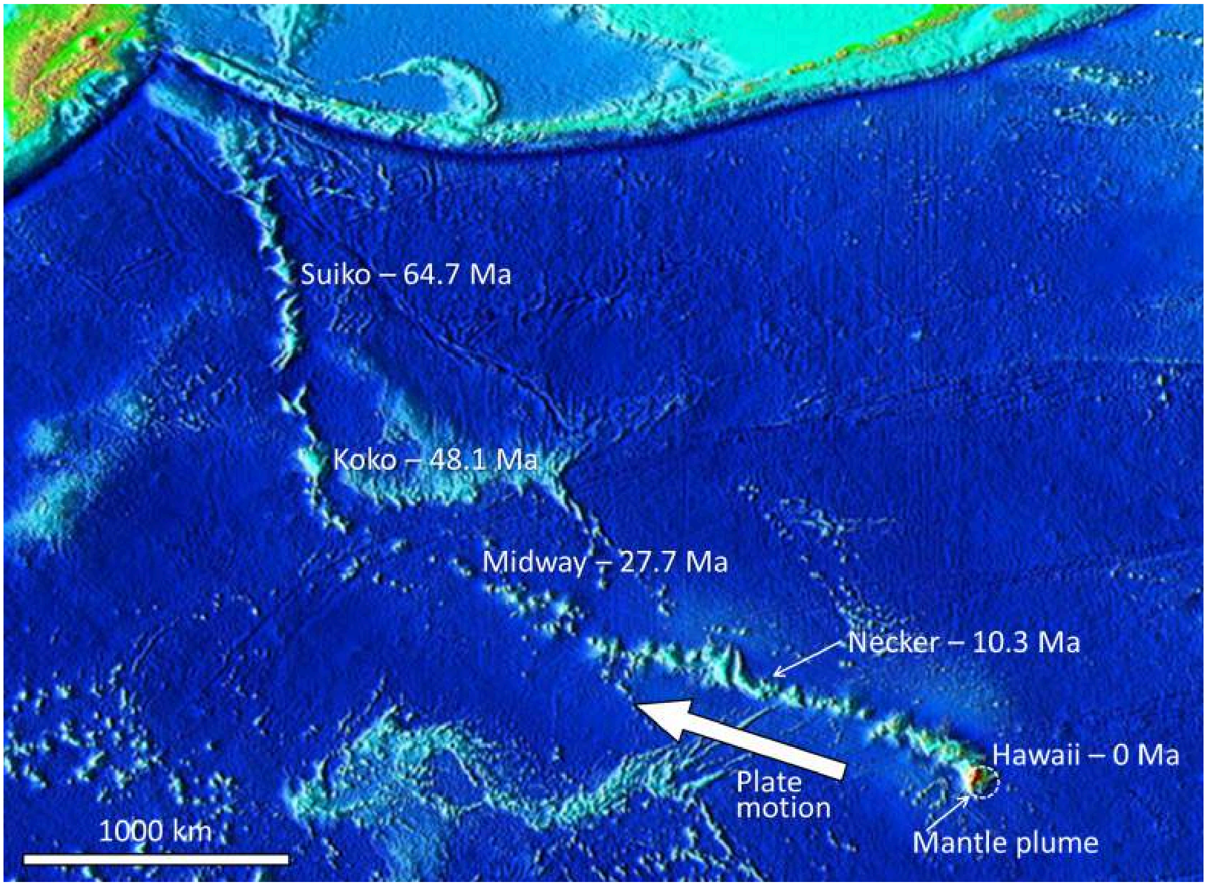


***Figure 4.9.1** The Hawaiian Islands/Emperor Seamount chain, with ages of selected structures. This chain has formed as the Pacific Plate moved northwest over a hot spot (Steven Earle, "Physical Geology").*

Since most mantle plumes are beneath the oceans, the early stages of volcanism typically take place on the seafloor. Over time, very large islands may form like those in Hawaii. In fact, if you measure it from its base on the seafloor to its summit, Mauna Loa on the island of Hawaii is the largest mountain on Earth, rising 9700 m (in comparison, the elevation of the summit of Mt. Everest is 8848 m). While the island of Hawaii is the youngest in the chain, there is actually a new volcano named Loihi, that is still submerged at a depth of 980 m SE of Hawaii, and may one day become a new Hawaiian island when it emerges 10,000 – 100,000 years from now.

There is evidence of many such mantle plumes around the world. Most are within the ocean basins, including places like Hawaii, Iceland, and the Galapagos Islands, but some are under continents. One example is the Yellowstone hot spot in the west-central United States, and another is the one responsible for the Anahim Volcanic Belt in central British Columbia. It is evident that mantle plumes are very long-lived phenomena, lasting for at least tens of millions of years, possibly for hundreds of millions of years in some cases.

4.10 Coral Reefs

*Modified from "Physical Geology" by Steven Earle**

It may seem odd to be discussing coral reefs in a section about geology, but due to the stony calcium carbonate skeletons secreted by many coral species, coral reefs are as interesting as geological features as they are biological ones. Corals grow best in warm, clear, tropical water, that is close enough to the surface for light to support photosynthesis by the algae living in the coral tissues. Because of this need for light, new coral will often grown on top of the stony skeletons of older corals.

In the 1830s Charles Darwin made some observations about different types of coral reefs, and hypothesized that they represent a progression from one form to the next. The types of reefs he examined were fringing reefs, barrier reefs, and atolls, which are associated with oceanic islands (Figure 4.10.1). **Fringing reefs** are reefs that are close to or are connected to shore. **Barrier reefs** are offshore reefs that are separated from the land by an expanse of water, such as a lagoon. **Atolls** are circular or oval reefs surrounding a lagoon, without any central land mass in the lagoon. Darwin speculated that reefs progressed from fringing, to barrier, to atolls as the land mass subsided. However, he had no explanation for how volcanic islands could sink. Today we know that Darwin was correct, and that islands can sink as oceanic crust subsides as it moves away from a spreading center, or as sea level rises as glaciers melt.

Figure 4.10.1 *A fringing reef (left), barrier reef (center), and atoll (right) (Public domain, via Wikimedia Commons).*

The progression starts with a fringing reef built against the shores of island (Figure 4.10.2). If sea level doesn't change or the land doesn't sink, many reefs will not progress beyond this stage. But if the land does subside, the corals would eventually sink too deep for light penetration (see section 6.5) and they would die. So as the reef gets deeper, the corals continue to grow upwards, at a rate of about 3-5 m per 1000 years, and eventually a lagoon develops between the reef and the island; the reef is now a barrier reef. If the land continues to subside until it is completely submerged, all that is left is a ring of coral that has been growing upwards around the central lagoon; an atoll.

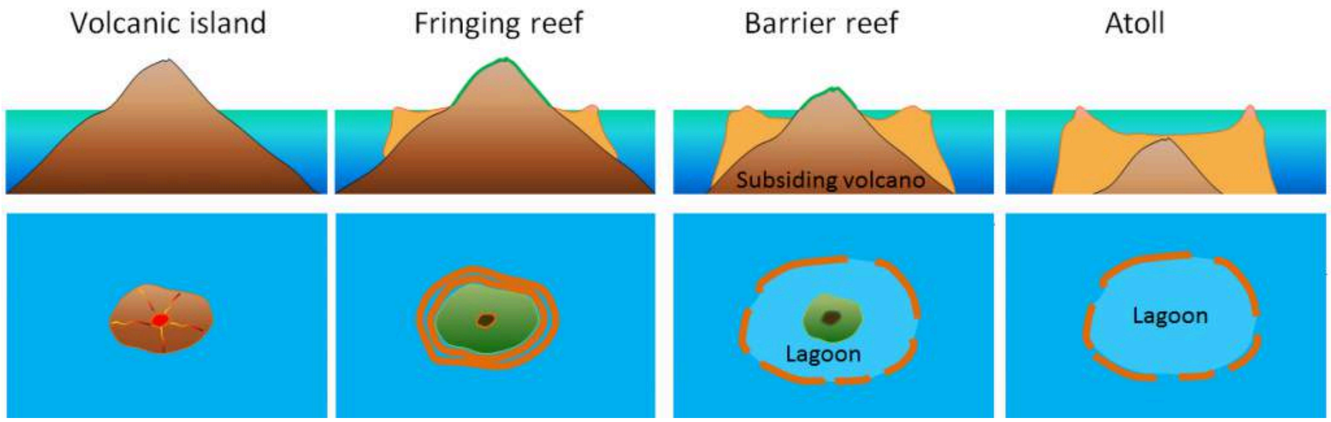


Figure 4.10.2 *Steps in the development of coral reefs (Steven Earle, "Physical Geology").*

4.11 Hydrothermal Vents

A whole new ecosystem reliant on the processes of plate tectonics was discovered on the deep seafloor of the Galapagos Rift in 1977. The deep sea submersible *Alvin* was exploring in 2500 m of water when it encountered unusually warm water. Following the temperature gradient, *Alvin* eventually discovered jets of superheated water coming from out of the seafloor at temperatures up to 350^{o} C (the normal temperature for water at this depth would be 2-4^{o} C). The water poured out of cracks in the crust, as well as through tall chimneys up to 20 m high and 1 m wide, and as it emerged it took on the appearance of thick black smoke, These fissures were named **hydrothermal vents**, and the chimneys “black smokers”.

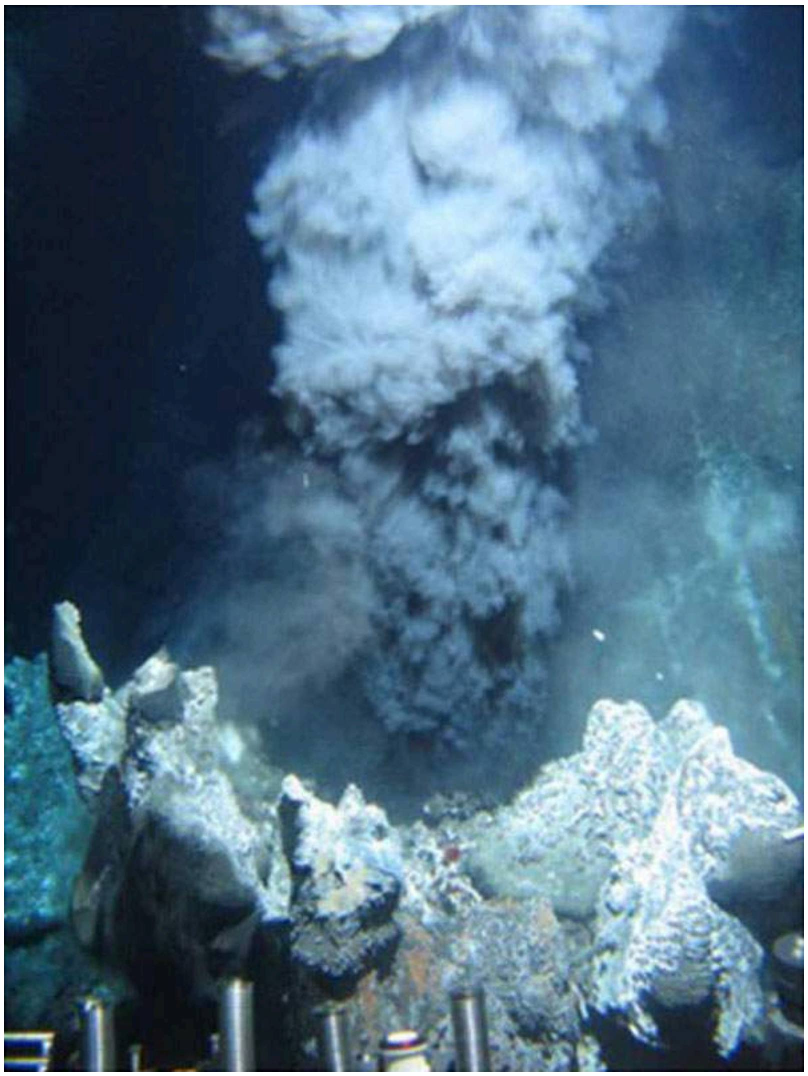


Figure 4.11.1 *A black smoker in the High Rise portion of the Endeavour hydrothermal vents (NOAA).*

To create these vents, water percolates into the crust where there are plumes of magma close to the surface. The water gets superheated by the magma, then moves back to the surface through convection and is released through the vents. The hot water dissolves minerals from the surrounding rock, and as the water emerges and cools, the dissolved minerals and inorganic sulfides precipitate out as small particles and turn the water black, leading to the black "smoke" coming from the vents. Precipitation of these minerals also create the tall chimneys characteristic of many hydrothermal vents.

Since their original discovery in the Galapagos Rift, hydrothermal vents have been located across the globe along oceanic ridges where there is shallow crust and a lot of tectonic activity (Figure 4.11.2).

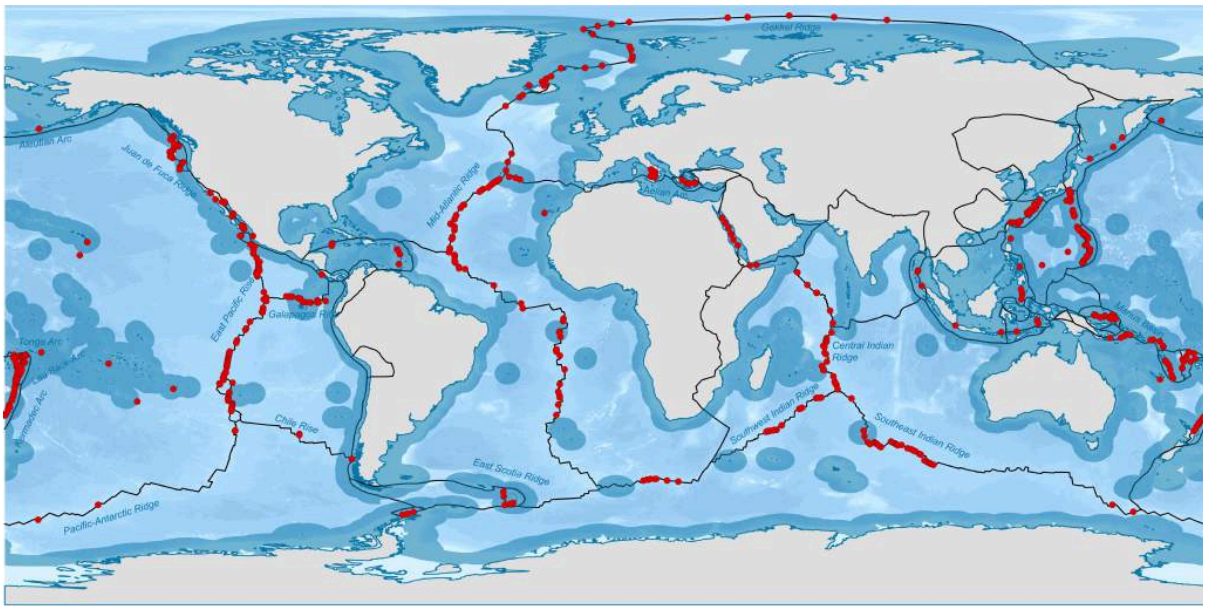


Figure 4.11.2 *Distribution of hydrothermal vents (red dots) and their association with plate boundaries (By DeDuijn (Own work) [CC BY-SA 4.0], via Wikimedia Commons).*

As unexpected as it was to discover these vent systems, even more surprising was the fact that they were teeming with life. The vents are surrounded by a diverse range of previously unknown organisms, including giant tube worms over 2 m long, crabs, shrimp, giant mussels, and mats of bacteria. How is it that such a diverse community can exist in the ocean depths, far removed from the sunlight that supports photosynthesis and primary production in most other ecosystems? The answer is that the water exiting the vents is rich in hydrogen sulfide (H_2S), oxygen and CO_2. The bacteria surrounding the vents use energy from the oxidation of sulfur compounds like H_2S to form carbohydrates from CO_2 and water. This is the process of **chemosynthesis**, and the bacteria are very productive as these reactions occur faster at high temperatures. The bacteria then represent the base of the food web, as other organisms eat the bacteria, or derive their energy from bacteria living symbiotically within their tissues. Watch the video below for more about hydrothermal vents.

A YouTube element has been excluded from this version of the text. You can view it online here: https://rwu.pressbooks.pub/webboceanography/?p=134

CHAPTER 5: CHEMICAL OCEANOGRAPHY

Chapter 5: Chemical Oceanography

Learning Objectives

After reading this chapter you should:

- understand some of the unique properties of water
- understand hydrogen bonds and how they affect water's properties
- understand the concept of outgassing
- understand the concept of residence time
- understand how salinity is measured and calculated
- know the major ions in seawater
- understand the rule of constant proportions
- understand how and why salinity changes as a function of latitude and depth
- know the processes responsible for dissolved gases in the ocean
- be able to produce and describe depth profiles for oxygen, carbon dioxide, and nutrients
- understand the role of carbon dioxide as a buffer in the ocean
- understand the issue of ocean acidification

Ask anyone to describe the oceans, and it is very likely that among the first words you hear will be "water" and "salty". Despite being perhaps our most vital natural resource, water often does not get the credit it deserves. We describe food or drinks we don't like as "watery", and we deride a perceived lack of strength as "weak as water" or "watered down". But in reality, water is an amazing substance, characterized by a number of unique properties that make life possible (it has been remarked that life is just water's way of moving itself around). And it is the salty nature of the ocean that sets it apart from most other bodies of water on Earth, as anyone who accidentally swallowed a mouthful of seawater at the beach will tell you. But ocean water contains a vast array of substances besides salt, and the field of chemical oceanography examines how these various ions, elements, gases, and other substances interact with each other, with marine organisms and oceanographic processes, and with the seawater itself.

This chapter begins with an introduction to the properties of water, before examining the origin and distribution of salts in the oceans. It then discusses the fates of dissolved gases in the ocean, particularly those that are important for marine life. The chapter concludes with a look at the current environmental issues associated with ocean acidification.

5.1 Properties of Water

The most obvious feature of the oceans is that they contain water. Water is so ubiquitous that it may not seem like a very interesting substance, but it has many unique properties that impact global oceanographic and climatological processes. Many of these processes are due to **hydrogen bonds** forming between water molecules.

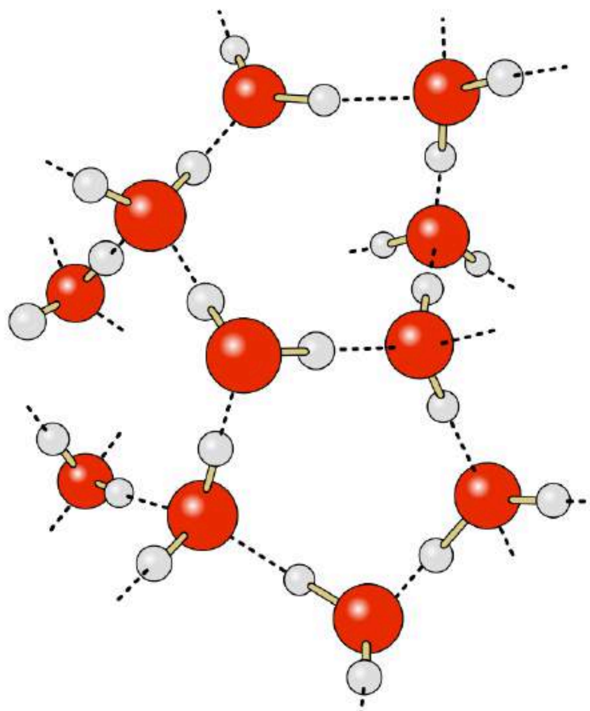


Figure 5.1.1 *Hydrogen bonds (dashed lines) between water molecules. Oxygen atoms are shown in red, hydrogen atoms in white (Public domain, via Wikimedia Commons).*

The water molecule consists of two hydrogen atoms and one oxygen atom. The electrons responsible for the bonds between the atoms are not distributed equally throughout the molecule, so that the hydrogen ends of water molecules have a slight positive charge, and the oxygen end has a slight negative charge, making water a **polar molecule**. The negative oxygen side of the molecule forms an attraction to the positive hydrogen end of a neighboring molecule. This rather weak force of attraction is called a hydrogen bond (Figure 5.1.1). If not for hydrogen bonds, water would vaporize at -68^{o} C, meaning liquid water (and thus life) could not exist on Earth. These hydrogen bonds are responsible for some of water's unique properties:

1. Water is the only substance to naturally exist in a solid, liquid, and gaseous form under the normal range of temperatures and pressures found on Earth. This is due to water's relatively high freezing and vaporizing points (see below).

2. Water has a **high heat capacity**, which is the amount of heat that must be added to raise its temperature. Specific heat is the heat required to raise the temperature of 1 g of a substance by 1^{o} C. Water has the highest specific heat of any liquid except ammonia (Table 5.1.1).

Table 5.1.1 Specific heat values for a number of common substances

	Specific Heat (calories/g/C^o)
Ammonia	1.13
Water	1.00
Acetone	0.51
Grain Alcohol	0.23
Aluminum	0.22
Copper	0.09
Silver	0.06

Water is therefore one of the most difficult liquids to heat or cool; it can absorb large amounts of heat without increasing its temperature. Remember that temperature reflects the average kinetic energy of the molecules within a substance; the more vigorous the motion, the higher the temperature. In water, the molecules are held together by hydrogen bonds, and these bonds must be overcome to allow the molecules to move freely. When heat is added to water the energy must first go to breaking the hydrogen bonds before the temperature can begin to rise. Therefore, much of the added heat is absorbed by breaking H bonds, not by increasing the temperature, giving water a high heat capacity.

Hydrogen bonds also give water a high latent heat; the heat required to undergo a phase change from solid to liquid, or liquid to gas. The **latent heat of fusion** is the heat required to go from solid to liquid; 80 cal/g in the case of ice melting to water. Ice is a solid because hydrogen bonds hold the water molecules into a solid crystal lattice (see below). As ice is heated, the temperature rises up to 0^o C. At that point, any additional heat goes to melting the ice by breaking the hydrogen bonds, not to increasing the temperature. So as long as ice is present, the water temperature will not increase. This is why your drink will remain cold as long as it contains ice; any heat absorbed goes to melting the ice, not to warming the drink.

When all of the ice is melted, additional heat will increase the temperature of the water 1^o C for each calorie of heat added, until it reaches 100^o C. At that point, any additional heat goes to overcoming the hydrogen bonds and turning the liquid water into water vapor, rather than increasing the water temperature. The heat required to evaporate liquid water into water vapor is the **latent heat of vaporization** which has a value of 540 cal/g (Figure 5.1.2).

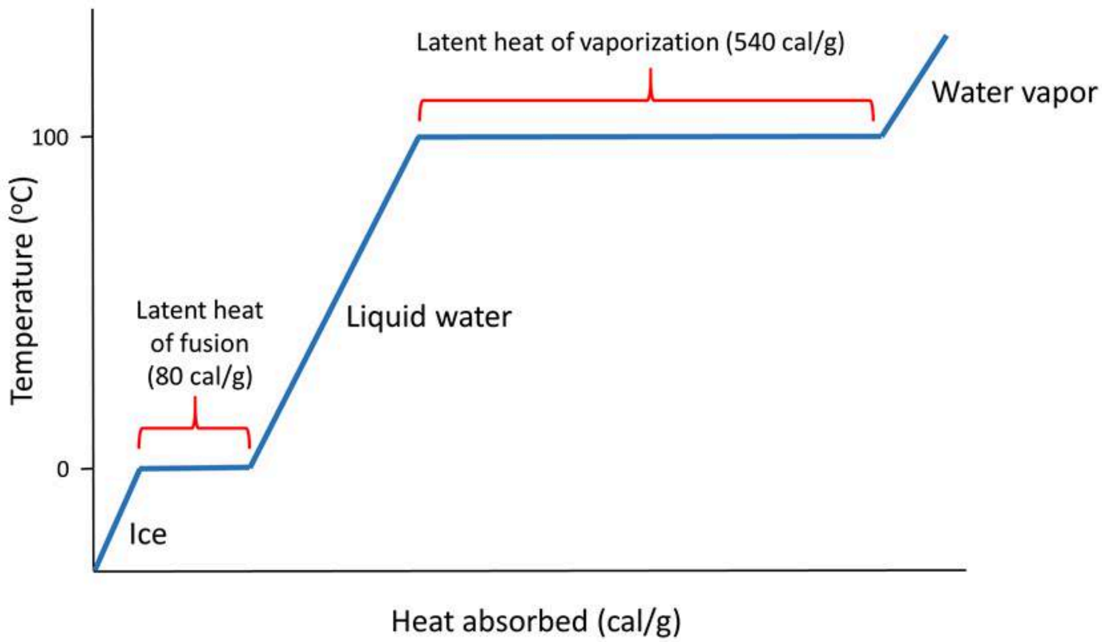


Figure 5.1.2 *Latent heat required for phase changes in water. Latent heat of fusion is the heat required to melt ice (80 cal/g), and latent heat of vaporization is the heat required to turn liquid water into water vapor (540 cal/g) (PW).*

The high heat capacity of water helps regulate global climate, as the oceans slowly absorb and release heat, preventing rapid swings in temperature (see section 8.1). It also means that aquatic organisms aren't as subjected to the same rapid temperature changes as terrestrial organisms. A deep ocean organism may not experience more than a 0.5^{o} C change in temperature over its entire life, while a terrestrial species may encounter changes of more than 20^{o} C in a single day!

3. Water dissolves more substances than any other liquid; it is a **"universal solvent"**, which is why so many substances are dissolved in the ocean. Water is especially good at dissolving ionic salts; molecules made from oppositely charged ions such as NaCl (Na^+ and Cl^-). In water, the charged ions attract the polar water molecules. The ions get surrounded by a layer of water molecules, weakening the bond between the ions by up to 80 times. With the bonds weakened between ions, the substance dissolves (Figure 5.1.3).

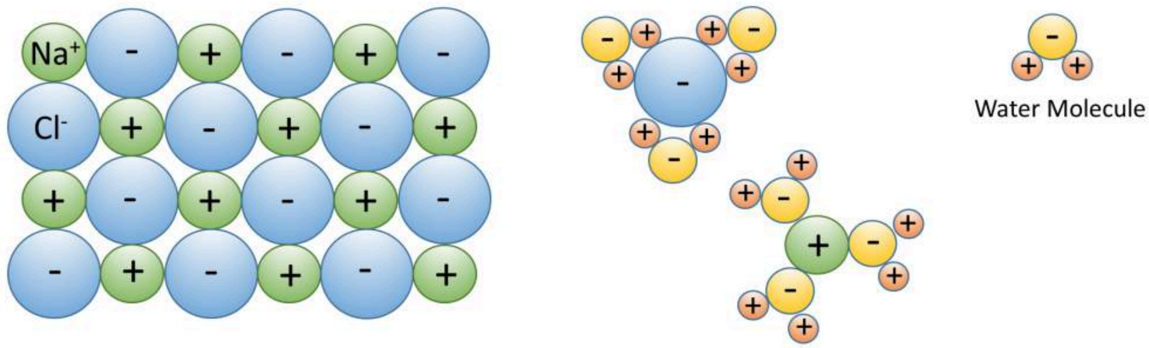


Figure 5.1.3 *Attraction between polar water molecules and charged ions (such as in NaCl) is greater than the attraction between the charged ions, causing the ions to dissociate and the salt to dissolve (PW).*

4. The **solid phase is less dense than the liquid phase**. In other words, ice floats. Most substances are denser in the solid form than in the liquid form, as their molecules are more closely packed together as a solid. Water is an exception: the density of fresh water is 1.0 g/cm^3, while the density of ice is 0.92 g/cm^3, and once again, this is due to the action of hydrogen bonds.

As water temperature cools the molecules slow down, eventually slowing enough that hydrogen bonds can form and hold the water molecules in a crystal lattice. The molecules in the lattice are spaced farther apart than the molecules in liquid water, which makes ice less dense than liquid water (Figure 5.1.4). This is familiar to anyone who has ever left a full water bottle in the freezer, only to have it burst as the water freezes and expands.

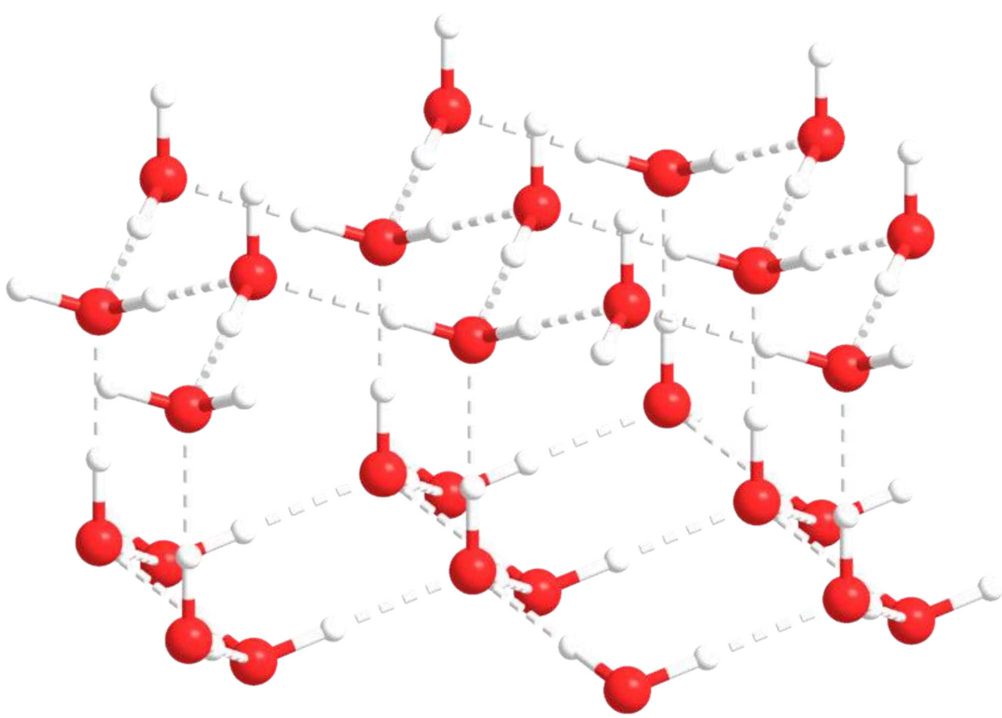


Figure 5.1.4 *Crystal lattice structure of ice, showing water molecules held together by hydrogen bonds (By Adam001d (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

But the relationship between temperature and water density is not a simple linear one. As water cools, its density increases as expected, as the water molecules slow down and get closer together. However, fresh water reaches its maximum density at a temperature of 4° C, and as it cools beyond that point its density declines as the hydrogen bonds begin to form and the intermolecular spacing increases (Figure 5.1.5 inset). The density continues to decline until the temperature reaches 0° C and ice crystals form, reducing the density dramatically (Figure 5.1.5).

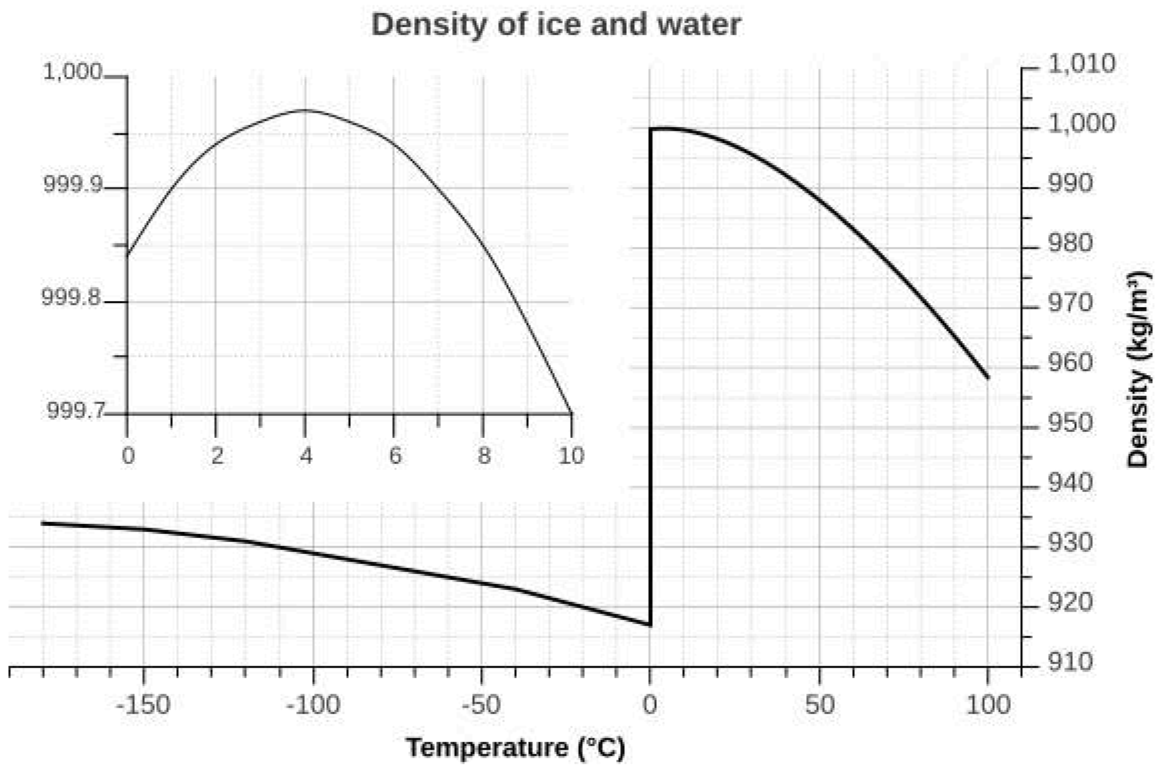


Figure 5.1.5 *As temperature declines, the density of water increases until it reaches maximum density at 4° C (inset). Density then declines slightly down to 0° C, where it declines dramatically as ice crystals form (Klaus-Dieter Keller, [CC BY-SA 3.0], via Wikimedia Commons).*

There are a number of important implications to ice being less dense than water. Ice floating on the surface of the ocean helps regulate ocean temperatures, and therefore global climate, by influencing the amount of sunlight that is reflected rather than absorbed (see section 8.1). On a smaller scale, surface ice can prevent lakes and ponds from freezing solid during the winter. As fresh surface water cools, the water gets denser, and sinks to the bottom. The new surface water then cools and sinks, and the process is repeated in what is referred to as **overturning**, with denser water sinking and less dense water moving to the surface only to be cooled and sink itself. In this way, the entire body of water is cooled somewhat evenly. This process continues until the surface water cools below 4° C. Below 4° C, the water becomes *less* dense as it cools, so it no longer sinks. Instead, it remains as the surface, getting colder and less dense, until it freezes at 0° C. Once fresh water freezes, the ice floats and insulates the rest of the water beneath it, reducing further cooling. The densest bottom water is still at 4° C, so it does not freeze, allowing the bottom of a lake or pond to remain unfrozen (which is good news for the animals living there) no matter how cold it gets outside.

The dissolved salts in seawater inhibit the formation of the crystal lattice, and therefore make it harder for ice to form. So seawater has a freezing point of about -2° C (depending on salinity), and freezes before a temperature of maximum density is reached. Thus seawater will continue to sink as it gets colder, until it finally freezes.

5. Water has a very **high surface tension**, the highest of any liquid except mercury (Table 5.1.2). Water molecules are attracted to each other by hydrogen bonds. For molecules not at the water surface, they are surrounded by other water molecules in all directions, so the attractive forces are evenly distributed in all directions. But for molecules at the surface there are few adjacent molecules above them, only below, so all of the attractive forces are directed inwards, away from the surface (Figure 5.1.6). This inwards force is what causes water droplets to take on a spherical shape, and water to bead up on a surface, as the spherical shape provides the minimum possible surface area. These attractive forces also cause the surface of the water to act like an elastic "skin" which allows things like insects to sit on the water's surface without sinking.

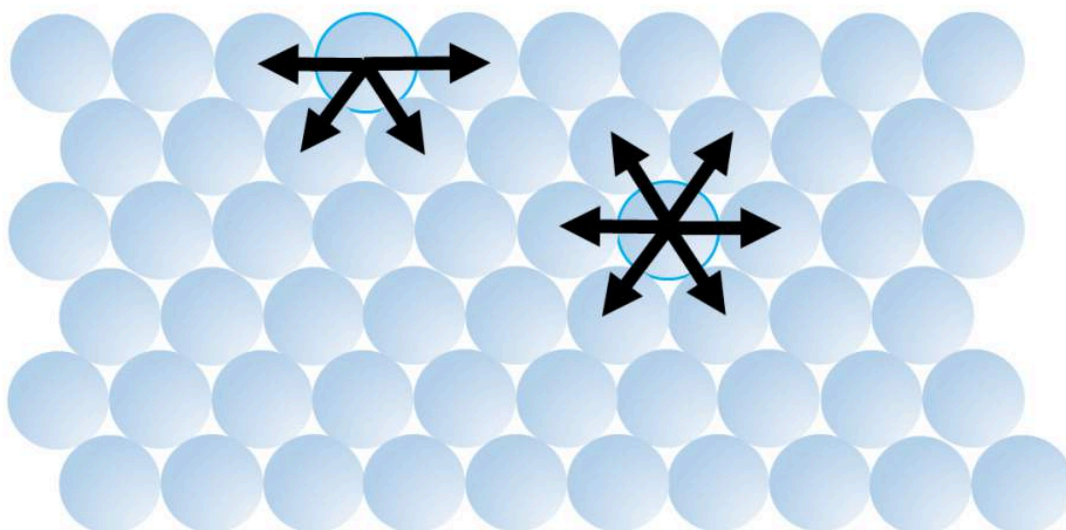


Figure 5.1.6 *The net attractive force between molecules at the surface is inwards, leading to surface tension. For molecules in the center, the force is equal in all directions (PW).*

Table 5.1.2 Surface tensions of various liquids

Liquid	**Surface Tension (millinewton/meter)**	**Temperature °C**
Mercury	487.00	15
Water	71.97	25
Glycerol	63.00	20
Acetone	23.70	20
Ethanol	22.27	20

5.2 Origin of the Oceans

*Portions modified from Karla Panchuk in "Physical Geology" by Steven Earle**

So how did the oceans form in the first place? Remember from section 3.1 that the early Earth was formed through the accretion of various materials, and that a period of melting and intense volcanic activity followed. The materials that accreted on the early Earth contained the components that would eventually become our oceans and atmosphere. Under the high pressures found in the Earth's interior, gases remain dissolved in magma. As these magmas rise to the surface through volcanic activity, the pressure is reduced and the gases are released through a process called **outgassing**. Volcanic activity releases many different gases, including water vapor, carbon dioxide (CO_2), sulfur dioxide (SO_2), carbon monoxide (CO), hydrogen sulfide (H_2S), hydrogen gas, nitrogen, and methane (CH_4). Lighter gases such as hydrogen and helium dissipated into space, but the heavier gases remained and formed Earth's early atmosphere.

The rise of atmospheric oxygen

It's worth noting that the early atmosphere lacked free oxygen (O_2), the form of oxygen that we breathe. We know this in part because prior to 2 billion years ago, there were no sedimentary beds stained red from oxidized iron minerals. Iron minerals were present, but not in oxidized form. At that time, O_2 was produced in the atmosphere when the Sun's ultraviolet rays split water molecules apart; however, chemical reactions removed the oxygen as quickly as it was produced. It wasn't until the appearance of life that Earth's atmosphere began to become oxygenated. Photosynthetic organisms used the abundant CO_2 in the atmosphere to manufacture their food, and released O_2 as a by-product. At first all of the oxygen was consumed by chemical reactions, but eventually the organisms released so much O_2 that it overwhelmed the chemical reactions and oxygen began to accumulate in the atmosphere, although present levels of 21% oxygen didn't occur until about 350 Ma. Today the part of our atmosphere that isn't oxygen consists largely of nitrogen (78%). The oxygen-rich atmosphere on our planet is life's signature. If geologic process were the only processes controlling our atmosphere, it would consist mostly of carbon dioxide, like the atmosphere of Venus.

As the early Earth cooled, the water vapor in the atmosphere condensed and fell as rain. By about 4 billion years ago, the first permanent accumulations of water were present on Earth, forming the oceans and other bodies of water. Water moves between these different reservoirs through the **hydrological cycle**. Water is evaporated from the oceans, lakes, streams, the surface of the land, and plants (transpiration) by solar energy (Figure 5.2.1). It is moved through the atmosphere by winds and condenses to form clouds of water droplets or ice crystals. It comes back down as rain or snow and then flows through streams and rivers, into lakes, and eventually back to the oceans. Water on the surface and in streams and lakes infiltrates the ground to become groundwater. Groundwater slowly moves through the rock and surface materials; some returns to other streams and lakes, and some goes directly back to the oceans.

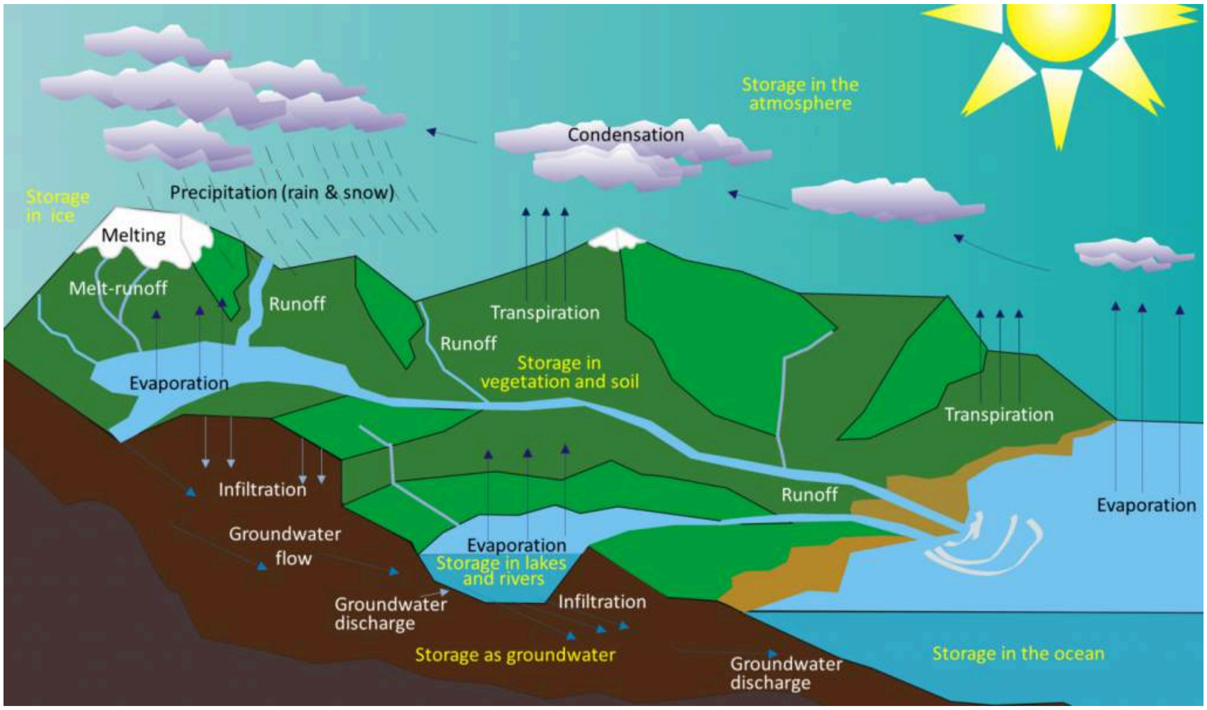


Figure 5.2.1 *Earth's hydrological cycle (Steven Earle, "Physical Geology").*

Water is stored in various reservoirs as it moves through this cycle. The largest, by far, is the oceans, accounting for 97% of the volume (Figure 5.2.2). Of course, that water is salty. The remaining 3% is fresh water. Two-thirds of our fresh water is stored in the ground and one-third is stored in ice. The remaining fresh water — about 0.03% of the total — is stored in lakes, streams, vegetation, and the atmosphere.

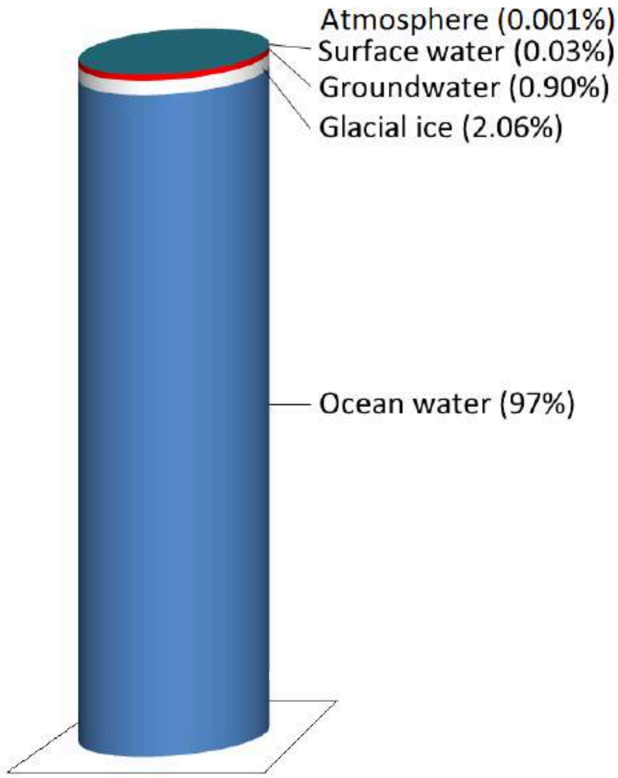


Figure 5.2.2 *Proportions of Earth's water found in the various reservoirs (Steven Earle, "Physical Geology").*

To put that in perspective, let's think about putting all of Earth's water into a 1 L jug. We start by almost filling the jug with 970 ml of water and 34 g of salt. Then we add one regular-sized (~20 mL) ice cube (representing glacial ice) and two teaspoons (~10 mL) of groundwater. All of the water that we see around us in lakes and streams and up in the sky can be represented by adding three more drops from an eyedropper.

Although the proportion of Earth's water that is in the atmosphere is tiny, the actual volume is huge. At any given time, there is the equivalent of approximately 13,000 km^3 of water in the air in the form of water vapor and water droplets in clouds. Water is evaporated from the oceans, vegetation, and lakes at a rate of 1,580 km^3 per day, and just about exactly the same volume falls as rain and snow every day, over both the oceans and land. The precipitation that falls on land goes back to the ocean in the form of stream flow (117 km^3/day) and groundwater flow (6 km^3/day).

How did the oceans get salty?

Outgassing was responsible for ocean formation, but how did the ocean water get salty? Most of the salts and dissolved elements in the ocean were probably outgassed along with the water vapor, so the ocean has probably always been about as salty as it is now. But we know that rainfall and other processes weather rocks on the Earth's surface, and runoff carries dissolved substances into the ocean, contributing to its salinity. Yet despite this constant input, the ocean's salt composition remains essentially the same. Therefore, the rate of input of new material must be balanced by the rate of removal; in other words, the oceans are in a **steady state** in regards to salinity.

There are multiple pathways through which dissolved ions enter the ocean; runoff from streams and rivers, volcanic activity, hydrothermal vents (see section 4.11), dissolution or decay of substances in the ocean, and groundwater input. Ions are removed from seawater as they are incorporated by living organisms (for example in shell production) or sediments, sea spray, percolation of water into the crust, or when sea water gets isolated from the ocean and evaporates.

The relationship between the input and removal of an ion can be examined through the concept of **residence time**, which is the average length of time a single atom of an element remains in the ocean before being removed. Residence time is calculated as:

$$\text{Residence time} = \frac{\text{amount of the substance in the ocean}}{\text{the rate at which the substance is added or removed}}$$

There is great variation in residence times for different substances (Table 5.2.1). Generally speaking, substances that are readily used in biological processes have short residence times, as they are used up as they become available. Substances with longer residence times are less reactive, and may be a part of long-scale geological cycles.

Table 5.2.1 Residence times for some constituents of sea water

Constituent	**Residence time (years)**
Chloride (Cl^-)	100,000,000
Sodium (Na^+)	68,000,000
Calcium (Ca^{2+})	1,000,000
Water	4100
Iron (Fe)	200

So what about lakes? They are subjected to runoff and river input, so why aren't they salty like the oceans? One reason is that compared to the oceans, lakes and ponds are relatively temporary phenomena, so they do not last long enough to accumulate the same levels of ions as the oceans. Furthermore, lakes often have rivers flowing both into and out of them, so many ions are removed through the outflow, eventually finding their way to the oceans. The oceans only receive river input; there are no rivers flowing out of the ocean to remove these materials, so they are found in greater abundance in sea water. It should be noted that there are some lakes that contain water whose salt content may rival or exceed that of the ocean; these lakes usually lack river outflow. The Great Salt Lake in the western United States is an example.

5.3 Salinity Patterns

All of the salts and ions that dissolve in seawater contribute to its overall **salinity**. Salinity of seawater is usually expressed as the grams of salt per kilogram (1000 g) of seawater. On average, about 35 g of salt is present in each 1 kg of seawater, so we say that the average salinity of the ocean salinity is 35 parts per thousand (ppt). Note that 35 ppt is equivalent to 3.5% (parts per hundred). Some sources now use practical salinity units (PSU) to express salinity values, where 1 PSU = 1 ppt. The units are not included, so we can refer simply to a salinity of 35.

Many different substances are dissolved in the ocean, but six ions comprise about 99.4% of all the dissolved ions in seawater. These six **major ions** are (Table 5.3.1):

Table 5.3.1 The six major ions in seawater

	g/kg in seawater	**% of ions by weight**
Chloride Cl^-	19.35	55.07%
Sodium Na^+	10.76	30.6%
Sulfate SO_4^{2-}	2.71	7.72%
Magnesium Mg^{2+}	1.29	3.68%
Calcium Ca^{2+}	0.41	1.17%
Potassium K^+	0.39	1.1%
		99.36%

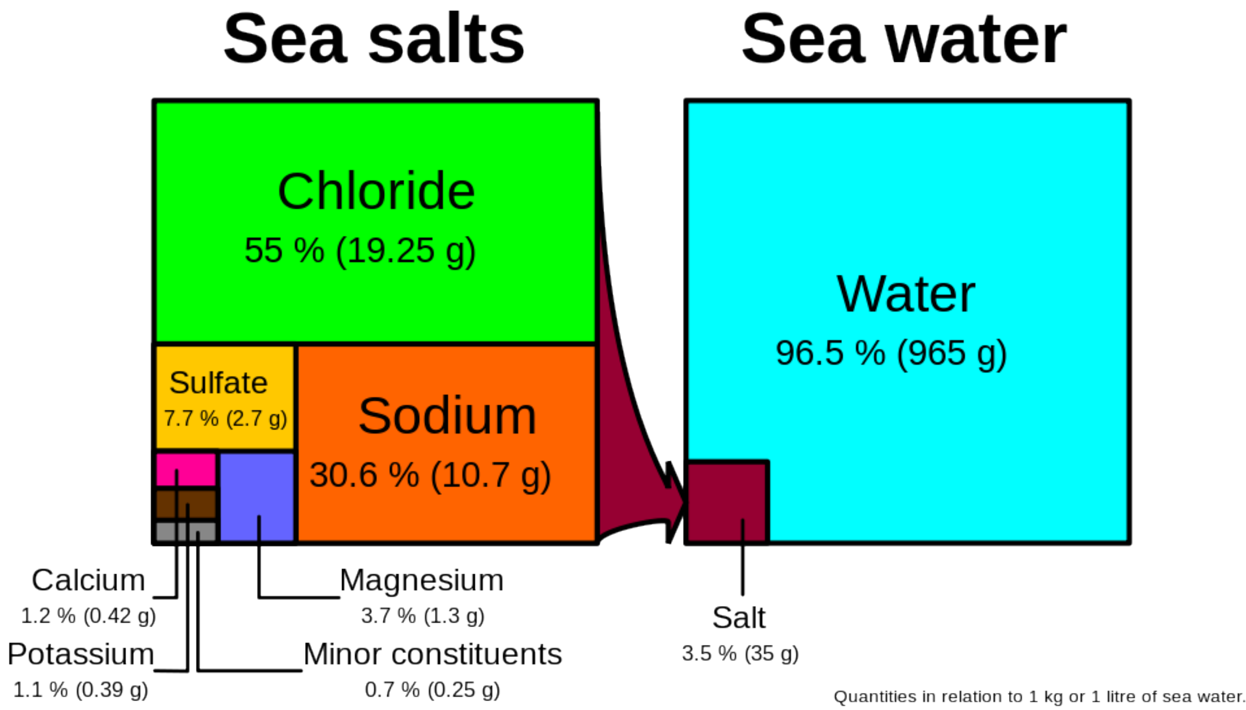


Figure 5.3.1 *The relative proportions of ions in seawater. (By derivative work: Tcncv (talk) Sea_salt-e_hg.svg: Hannes Grobe, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany; SVG version by Stefan Majewsky (Sea_salt-e_hg.svg) [CC BY-SA 2.5], via Wikimedia Commons)*

Chloride and sodium, the components of table salt (sodium chloride NaCl), make up over 85% of the ions in the ocean, which is why seawater tastes salty (Figure 5.3.1). In addition to the major constituents, there are numerous minor constituents; radionucleotides, organic compounds, metals etc. These minor constituents are found in concentrations of ppm (parts per million) or ppb (parts per billion), unlike the major ions that are far more abundant (ppt) (Table 5.3.2). To put this into perspective, 1 ppm = 1 mg/kg, or the equivalent of 1 teaspoon of sugar dissolved in 14,000 cans of soda. 1 ppb = 1 µg/kg, or the equivalent of 1 teaspoon of a substance dissolved in five Olympic-sized swimming pools! These minor constituents represent numerous substances, but together they make up less than 1% of the ions in the seawater. Some of these may be important as minerals and trace elements vital to living organisms, but they don't have much impact on the overall salinity. But given the vast size of the oceans, even materials found in trace abundance can represent fairly large reservoirs. For example gold is a trace element in seawater, found in concentrations of parts per trillion, yet if you could extract all of the gold in just one km^3 of seawater, it would be worth about $20 million!

Table 5.3.2 Concentrations of some minor elements in seawater

	g/kg in seawater		**g/kg in seawater**
Carbon	0.028	Iron	2×10^{-6}
Nitrogen	0.0115	Manganese	2×10^{-7}
Oxygen	0.006	Copper	1×10^{-7}
Silicon	0.002	Mercury	3×10^{-8}
Phosphorous	6×10^{-5}	Gold	4×10^{-9}
Uranium	3.2×10^{-6}	Lead	5×10^{-10}
Aluminum	2×10^{-6}	Radon	6×10^{-19}

Because the six major ions in seawater comprise over 99% of the total salinity, changes in abundance of the minor constituents have little effect on overall salinity. Furthermore, the **rule of constant proportions** states that even though the absolute salinity of ocean water might differ in different places, the relative *proportions* of the six *major ions* within that water are always constant. For example, no matter the total salinity of a seawater sample, 55% of the total salinity will be due to chloride, 30% due to sodium, and so on. Since the proportion of these major ions does not change, we call these **conservative ions**.

Given these constant proportions, in order to calculate total salinity you can simply measure the concentration of just one of the major ions and use that value to calculate the rest. Traditionally chloride has been the ion measured because it is the most abundant, and thus the simplest to measure accurately. Multiplying the concentration of chloride by 1.8 gives the total salinity. For example, looking at Figure 5.3.1, 19.25 g/kg (ppt) chloride x 1.8 = 35 ppt. Today, for rapid measurements of salinity, electrical conductivity is often used rather than determining chloride concentrations (see box below).

Measuring salinity

There are a number of methods available for measuring the salinity of water. The most precise measurements utilize direct chemical analysis of the seawater in a lab setting, but there are a number of ways to get immediate salinity measurements in the field. For a quick estimate of salinity, a hand-held **refractometer** can be used (right).

This instrument measures the degree of bending, or refraction, of light rays as they pass through a fluid. The greater the amount of dissolved salts in the sample, the greater the degree of light refraction. The observer traps a drop of water on the blue screen, and looks through the eyepiece. The dividing line between the blue and white sections of the scale (inset) can be used to read the salinity.

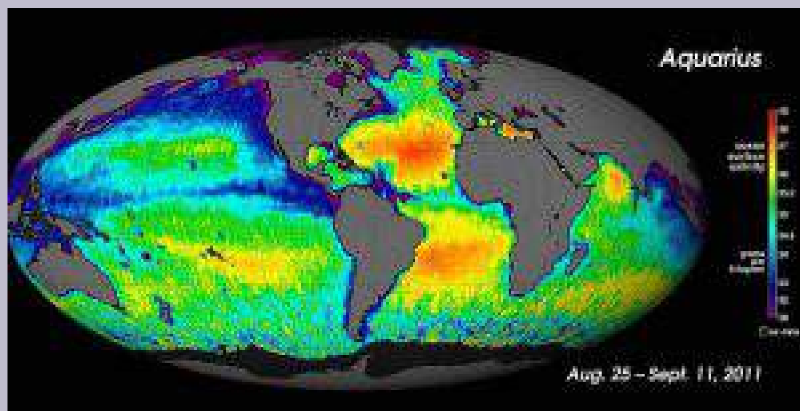
Handheld refractometer (Photo by CEphoto, Uwe Aranas / CC-BY-SA-3.0 via Wikimedia Commons)

For more accurate measurements, most oceanographers use an instrument that measures electrical **conductivity**. An electrical current is passed between two electrodes immersed in water, and the higher the salinity, the more readily the current will be conducted (the ions in seawater conduct electrical currents). Conductivity probes are often bundled into an instrument called a **CTD**, which stands for Conductivity, Temperature, and Depth, which are the most commonly-measured parameters. Modern CTDs can be outfitted with an array of probes measuring parameters like light, turbidity (water clarity), dissolved gases etc. CTDs can be large instruments (below), but small hand-held salinity probes are also widely available.

Left: *The inner workings of a CTD. An array of different probes are attached.* ***Right:*** *A CTD in a protective cage ready to be deployed at sea. (Left: Hannes Grobe [CC BY 3.0 (https://creativecommons.org/licenses/by/3.0)]; Right: © Hans Hillewaert (CC BY-SA 4.0) via Wikimedia Commons)*

For large-scale salinity measurements, oceanographers can use satellites, such as the

Aquarius satellite, which was able to measure surface salinity differences as small as 0.2 PSU as it mapped the ocean surface every seven days (below).



(NASA, Public Domain)

It is important to be aware that while the rule of constant proportions applies to most of the ocean, there may be certain coastal areas where lots of river discharge may alter these proportions slightly. Furthermore, it is important to remember that the rule of constant proportions only applies to the *major* ions. The proportions of the minor ions may fluctuate, but remember that they make a very minor contribution to overall salinity. Because the concentrations of the minor ions are not constant, these are referred to as **non-conservative ions**.

Why are the major ions found in constant proportions? There is constant input of ions from river runoff and other processes, usually in very different proportions from what is found in the ocean. So why don't the proportions in the oceans change? Most of the ions discharged by rivers have fairly low residence times (see section 5.2) compared to ions in seawater, usually because they are used in biological processes. These low residence times do not allow the ions to accumulate and alter salinity. Also, the mixing time of the world ocean is around 1000 years, which is very short compared to the residence times of the major ions, which may be tens of millions of years long. So during the residence time of a single ion the ocean has mixed numerous times, and the major ions have become evenly distributed throughout the ocean.

Variations in Salinity

Total salinity in the open ocean averages 33-37 ppt, but it can vary significantly in different locations. But since the major ion proportions are constant, the regional salinity differences must be due more to water input and removal rather than the addition or removal of ions. Fresh water input comes through processes like precipitation, runoff from land, and melting ice. Fresh water removal primarily comes from evaporation and freezing (when seawater freezes, the resulting ice is mostly fresh water and the salts are excluded, making the remaining water even saltier). So differences in rates of precipitation, evaporation, river discharge, and ice formation play a significant role in regional salinity variations. For example, the Baltic Sea has a very low surface salinity of around 10 ppt, because it is a mostly enclosed body of water with lots of river input. Conversely, the Red Sea is very salty (around 40 ppt), due to the lack of precipitation and the hot environment which leads to high levels of evaporation.

One of the saltiest large bodies of water on Earth is the Dead Sea, between Israel and Jordan. Salinity in the

Dead Sea is around 330 ppt, which is almost ten times saltier than the ocean. This extremely high salinity is a result of the hot, arid conditions in the Middle East that lead to high rates of evaporation. In addition, in the 1950s the flow from the Jordan River was diverted away from the Dead Sea, so there is no longer significant fresh water input. With no input and high evaporation, the water level in the Dead Sea is receding at a rate of about 1 m per year. The high salinity makes the water very dense, which creates buoyant forces that allow people to easily float at the surface. But the high salinity also means that the water is too salty for most living organisms, so only microbes are able to call it home; hence the name the Dead Sea. But as salty as the Dead Sea may be, it is not the saltiest body of water on Earth. That distinction currently belongs to Gaet'ale Pond in Ethiopia, with a salinity of 433 ppt!

Latitudinal Variations

While local conditions are important for determining salinity patterns in any single location, there are some global patterns that bear further investigation. Temperature is highest at the equator, and lowest near the poles, so we would expect higher rates of evaporation, and therefore higher salinity, in equatorial regions (Figure 5.3.2). This is generally the case, but in the figure below salinity right along the equator seems to be a little lower than at slightly higher latitudes. This is because equatorial regions also get a high volume of rain on a regular basis, which dilutes the surface water along the equator. So the higher salinities are found at subtropical, warm latitudes with high evaporation and less precipitation. At the poles there is little evaporation, which, coupled with ice and snow melting, produces a relatively low surface salinity. The image below shows high salinity in the Mediterranean Sea; this is located in a warm region with high evaporation, and the sea is largely isolated from mixing with the rest of the North Atlantic water, leading to high salinity. Lower salinities, such as those around southeast Asia, are the result of precipitation and high volumes of river input.

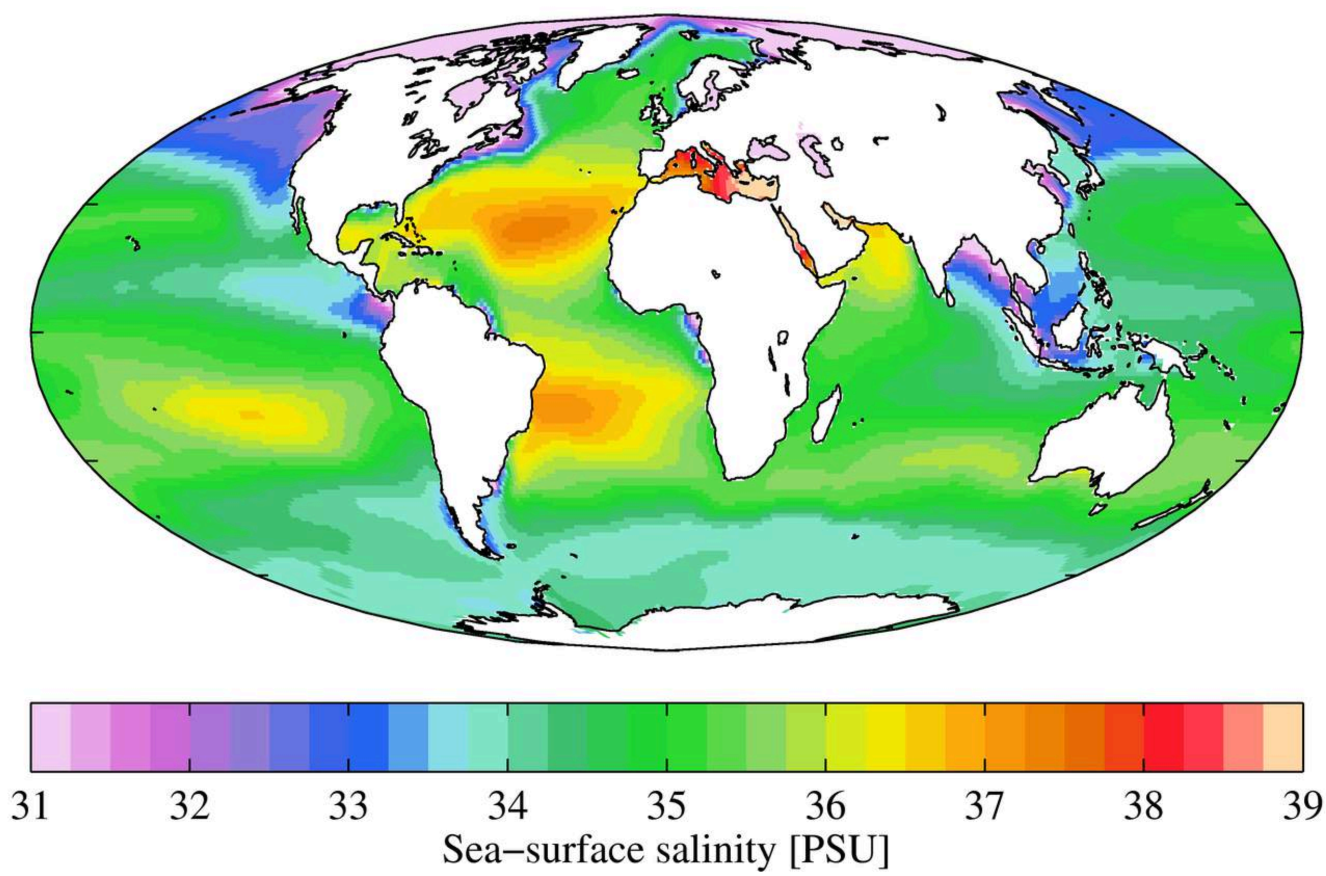


***Figure 5.3.2** Annual mean global sea surface salinity (Data from World Ocean Atlas 2009. By Plumbago – Own work, [CC BY-SA 3.0], via Wikimedia Commons).*

Figure 5.3.3 shows the mean global differences between evaporation and precipitation (evaporation – precipitation). Green colors represent areas where precipitation exceeds evaporation, while brown regions are where evaporation is greater than precipitations. Note the correlation between precipitation, evaporation, and surface salinity as seen in Figure 5.3.2.

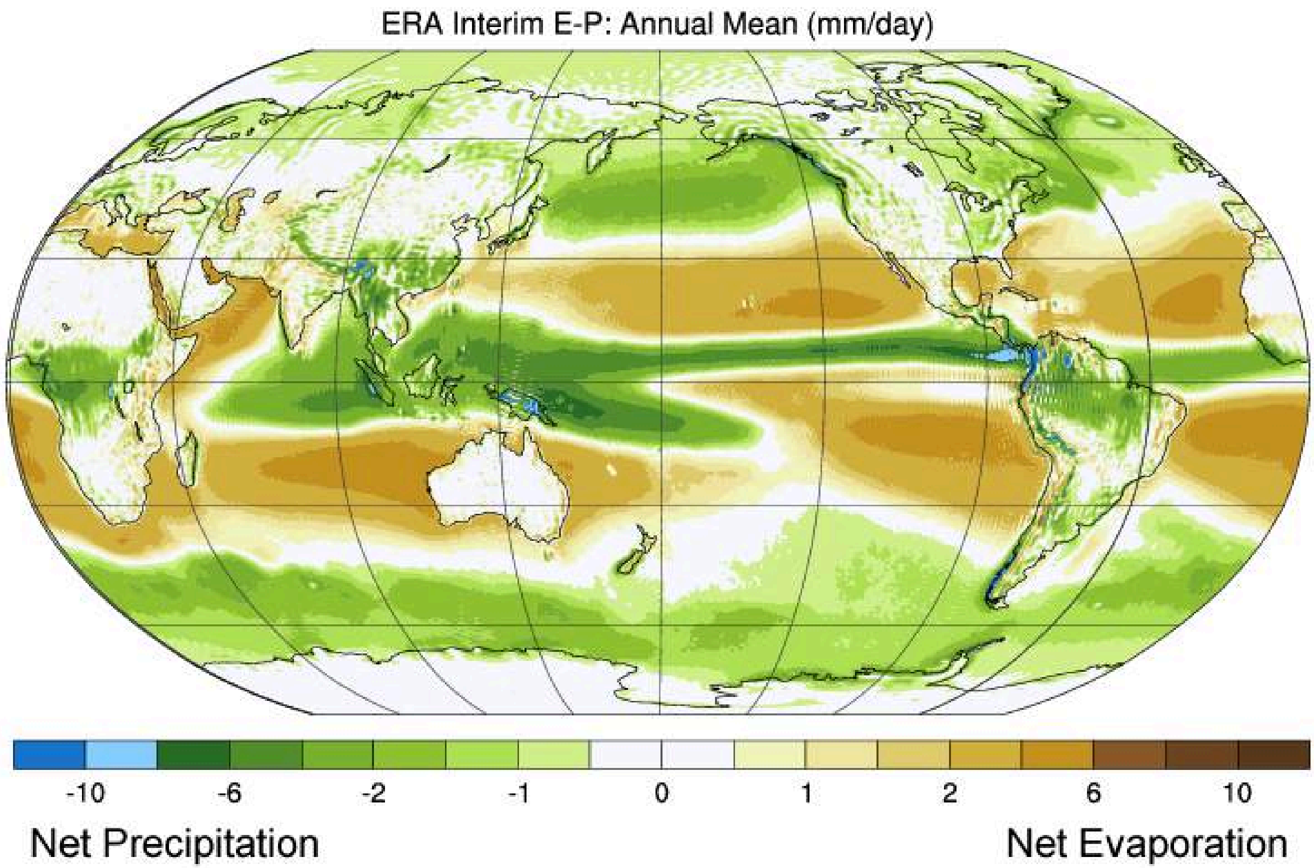


Figure 5.3.3 *Annual mean global differences in evaporation and precipitation (J.Fasullo. National Center for Atmospheric Research Staff (Eds). "The Climate Data Guide: ERA-Interim: derived components." Retrieved from https://climatedataguide.ucar.edu/climate-data/era-interim-derived-components).*

Vertical Variation

In addition to geographical variation in salinity, there are also changes in salinity with depth. As we have seen, most differences in salinity are due to variations in evaporation, precipitation, runoff, and ice cover. All of these process occur at the ocean surface, not at depth, so the most pronounced differences in salinity should be found in surface waters. Salinity in deeper water remains relatively uniform, as it is unaffected by these surface processes. Some representative salinity profiles are shown in Figure 5.3.4. At the surface, the top 200 m or so show relatively uniform salinity in what is called the **mixed layer**. Winds, waves, and surface currents stir up the surface water, causing a great deal of mixing in this layer and fairly uniform salinity conditions. Below the mixed layer is an area of rapid salinity change over a small change in depth. This zone of rapid change is called the **halocline,** and it represents a transition between the mixed layer and the deep ocean. Below the halocline, salinity may show little variation down to the seafloor, as this region is far removed from the surface processes that impact salinity. In the figure below, note the low surface salinity at high latitudes, and higher surface salinity at low latitudes as discussed above. Yet despite the surface differences, salinity at depth in both locations may be very similar.

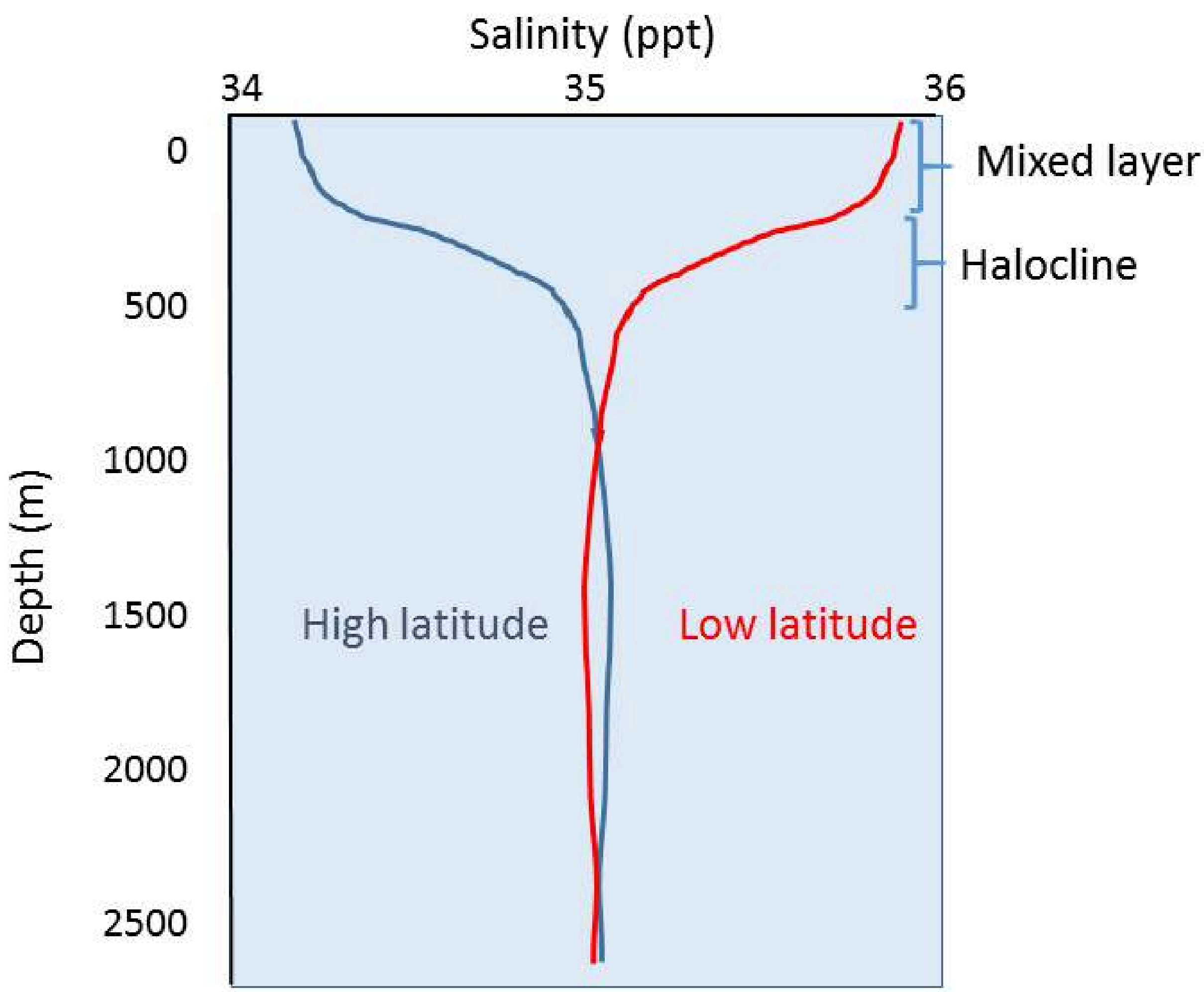


Figure 5.3.4 *Salinity profiles from two hypothetical sites in the open ocean, one from high latitude and one from low latitude (PW).*

5.4 Dissolved Gases: Oxygen

Ions are not the only materials that are dissolved in seawater. The oceans also contain dissolved gases that are very important to living organisms, particularly oxygen (O_2), carbon dioxide (CO_2), and nitrogen (N_2). Oxygen is required for respiration in marine plants, algae, and phytoplankton (the primary producers) and animals. Carbon dioxide is utilized by the primary producers to power photosynthesis, a byproduct of which is oxygen. Nitrogen gas dissolved in the ocean is fixed by bacteria and converted into the forms required for primary production, such as nitrate and nitrite.

All of these gases are found in the atmosphere, and can enter the ocean by dissolving into the water at the ocean's surface. But the amount of each gas in air is very different from the amount found in the ocean (Table 5.4.1).

Table 5.4.1 Percentage of total gas in each compartment

	Air	Total Ocean	Surface Ocean
N_2	78%	11%	48%
O_2	21%	6%	36%
CO_2	0.04%	83%	15%

The amount of each gas that can dissolve in the ocean depends on the solubility and saturation of the gas in water. **Solubility** refers to the amount of a dissolved gas that the water can hold under a particular set of conditions, which are usually defined as 0^{o} C and 1 atmosphere of pressure. The solubility of a gas increases with increasing pressure, decreased temperature, and decreased salinity. **Saturation** refers to the amount of gas currently dissolved in the water, relative to the maximum possible content. If the water is undersaturated, more gas can dissolve. If the water is saturated or supersaturated, gas may be released. Most atmospheric gases are saturated in the ocean, but O_2 and CO_2 are not saturated because they are rapidly used by living organisms.

Oxygen

Typical oceanic dissolved oxygen profiles are shown in Figure 5.4.1. The shape of the profile is determined by the various processes that add or remove oxygen from the water at different depths.

Oxygen content is highest at the surface for two main reasons; this is where oxygen dissolves into the ocean from the atmosphere, and the surface water is where oxygen is produced by phytoplankton through photosynthesis. Respiration is also occurring in the surface waters, but the rate of photosynthetic oxygen production is greater than the rate of removal through respiration. It should be noted that even though dissolved oxygen is highest at the surface, there is still far less oxygen in the water than is found in the air. Well-oxygenated surface water may only contain around 8 mg O_2/l, while the air contains 210 mg O_2/l.

As depth increases, dissolved oxygen declines, reaching a minimum between a few hundred meters and 1000 m deep, the aptly-named **oxygen minimum layer**. At these depths and below, the water is too far removed from the surface for any atmospheric exchange, and there is not enough light to support photosynthesis, so there is little if any oxygen added to the water. At the same time, oxygen is removed from the water through the respiration of deep water organisms, and the decomposition of organic material by bacteria as it sinks to depth.

Below the oxygen minimum layer there is often an increase in dissolved oxygen at the greatest depths (Figures 5.4.1, 5.4.2). This bottom water is usually colder than the surface water and is under enormous pressure; as stated above, lower temperatures and higher pressure increase the solubility of dissolved gases. But there is another reason that bottom water contains more oxygen than mid-water depths that has to do with the way water circulates throughout the deep ocean (see section 9.8). In polar regions, the cold surface water absorbs lots of oxygen. This cold, oxygen-rich water sinks to the bottom due to its high density, taking the oxygen with it. The oxygen-rich bottom water will then spend the next thousand years or so moving over the seafloor throughout the major ocean basins. This deep water circulation is the source of oxygen for bottom-dwelling (benthic) organisms. The oxygen-rich bottom water forms in the polar regions of the Atlantic, and slowly makes its way to the Pacific, with oxygen being removed for respiration along the way. This is why dissolved oxygen levels in Pacific deep water are generally lower than in the Atlantic (Figure 5.4.1).

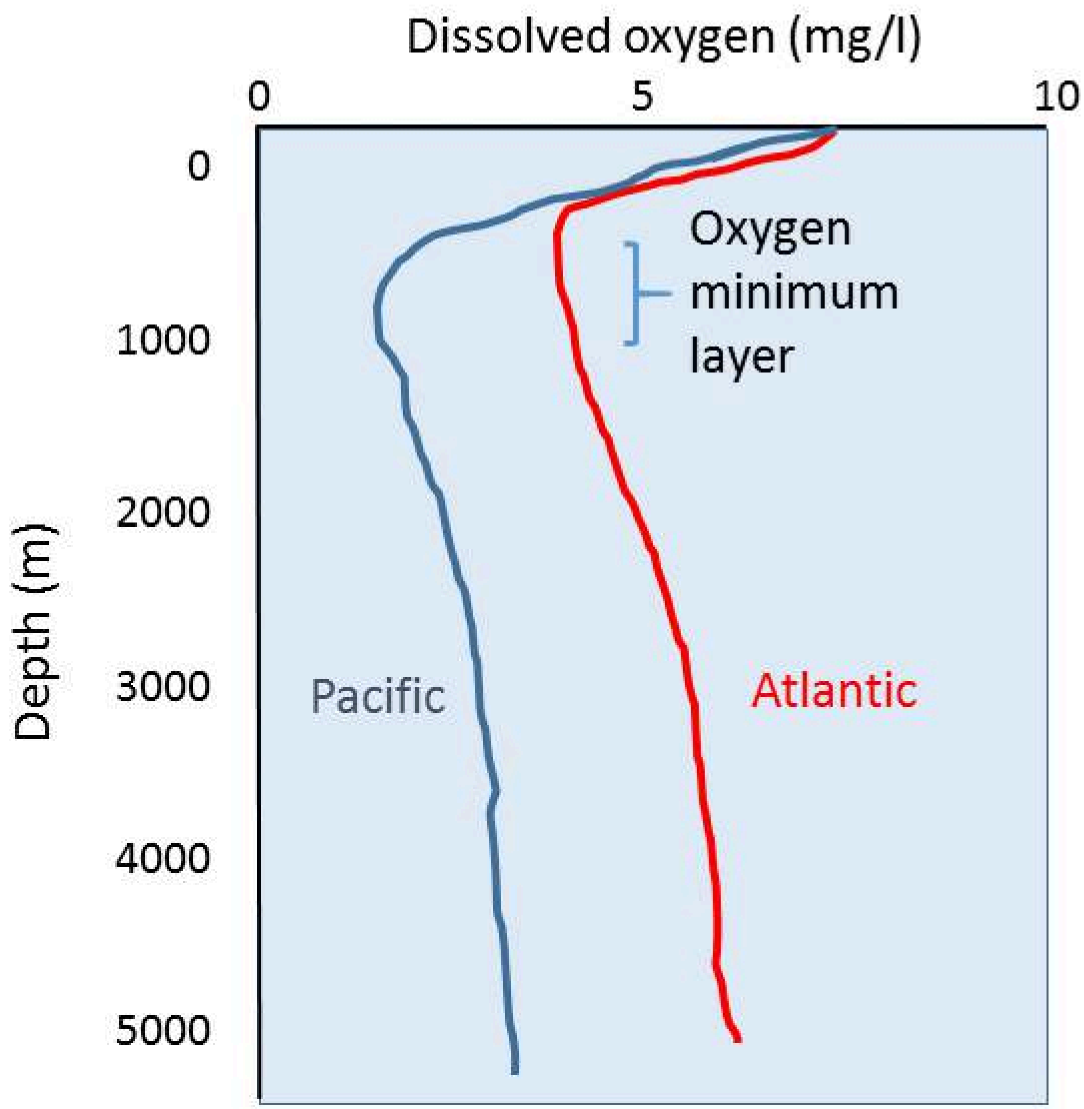


Figure 5.4.1 *Representative dissolved oxygen profiles for the Pacific and Atlantic oceans (PW).*

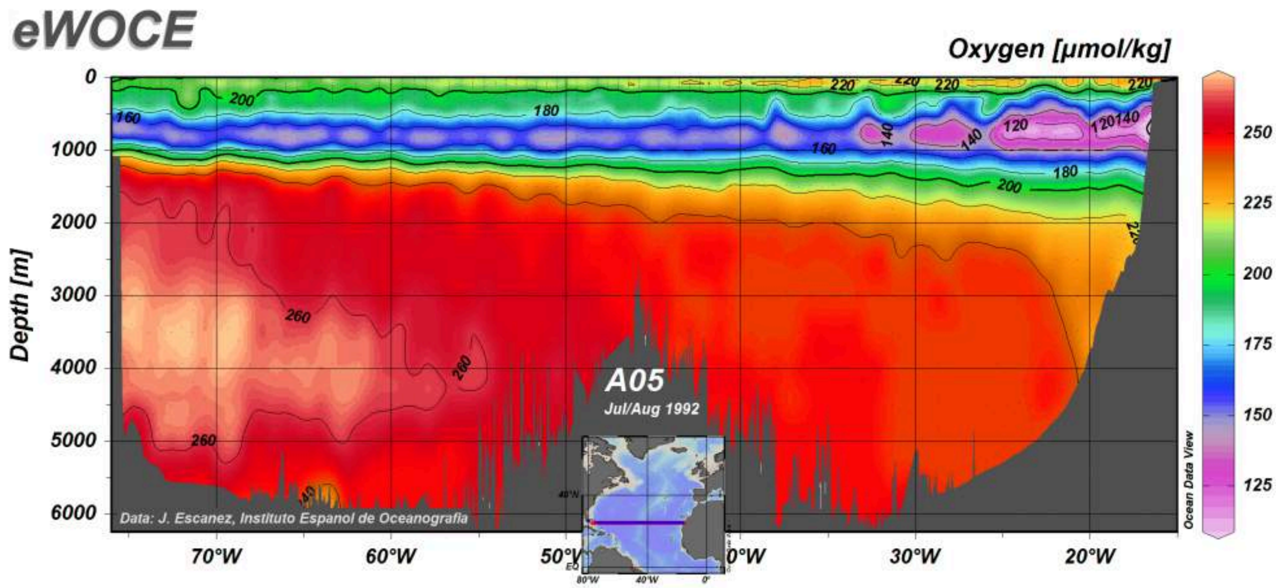


__Figure 5.4.2__ Dissolved oxygen profile from a transect across the Atlantic Ocean from Florida to the coast of Africa (inset). The oxygen minimum layer is visible between 500-1000 m (eWOCE, http://www.ewoce.org/gallery/eWOCE_Tables.html#Atlantic).

Areas where dissolved oxygen levels are too low to support most life are referred to as **hypoxic** zones (they are experiencing **hypoxia**, or low oxygen). Hypoxia is usually defined as oxygen levels below 2 mg/L. **Anoxic** zones (anoxia = without oxygen) show more severe forms of hypoxia, with oxygen below 0.5 mg/L. Some parts of the oceans may experience seasonal or temporary periods of hypoxia, while in other areas these conditions may last much longer. These hypoxic conditions often lead to mass die-offs of marine organisms who struggle to survive without sufficient oxygen.

Additional links for more information:

- NOAA site on oceanic hypoxic zones: http://oceanservice.noaa.gov/hazards/hypoxia/

5.5 Dissolved Gases: Carbon Dioxide, pH, and Ocean Acidification

Oxygen and carbon dioxide are involved in the same biological processes in the ocean, but in opposite ways; photosynthesis consumes CO_2 and produces O_2, while respiration and decomposition consume O_2 and produce CO_2. Therefore it should not be surprising that oceanic CO_2 profiles are essentially the opposite of dissolved oxygen profiles (Figure 5.5.1). At the surface, photosynthesis consumes CO_2 so CO_2 levels remain relatively low. In addition, organisms that utilize carbonate in their shells are common near the surface, further reducing the amount of dissolved CO_2.

In deeper water, CO_2 concentration increases as respiration exceeds photosynthesis, and decomposition of organic matter adds additional CO_2 to the water. As with oxygen, there is often more CO_2 at depth because cold bottom water holds more dissolved gases, and high pressures increase solubility. Deep water in the Pacific contains more CO_2 than the Atlantic as the Pacific water is older and has accumulated more CO_2 from the respiration of benthic organisms.

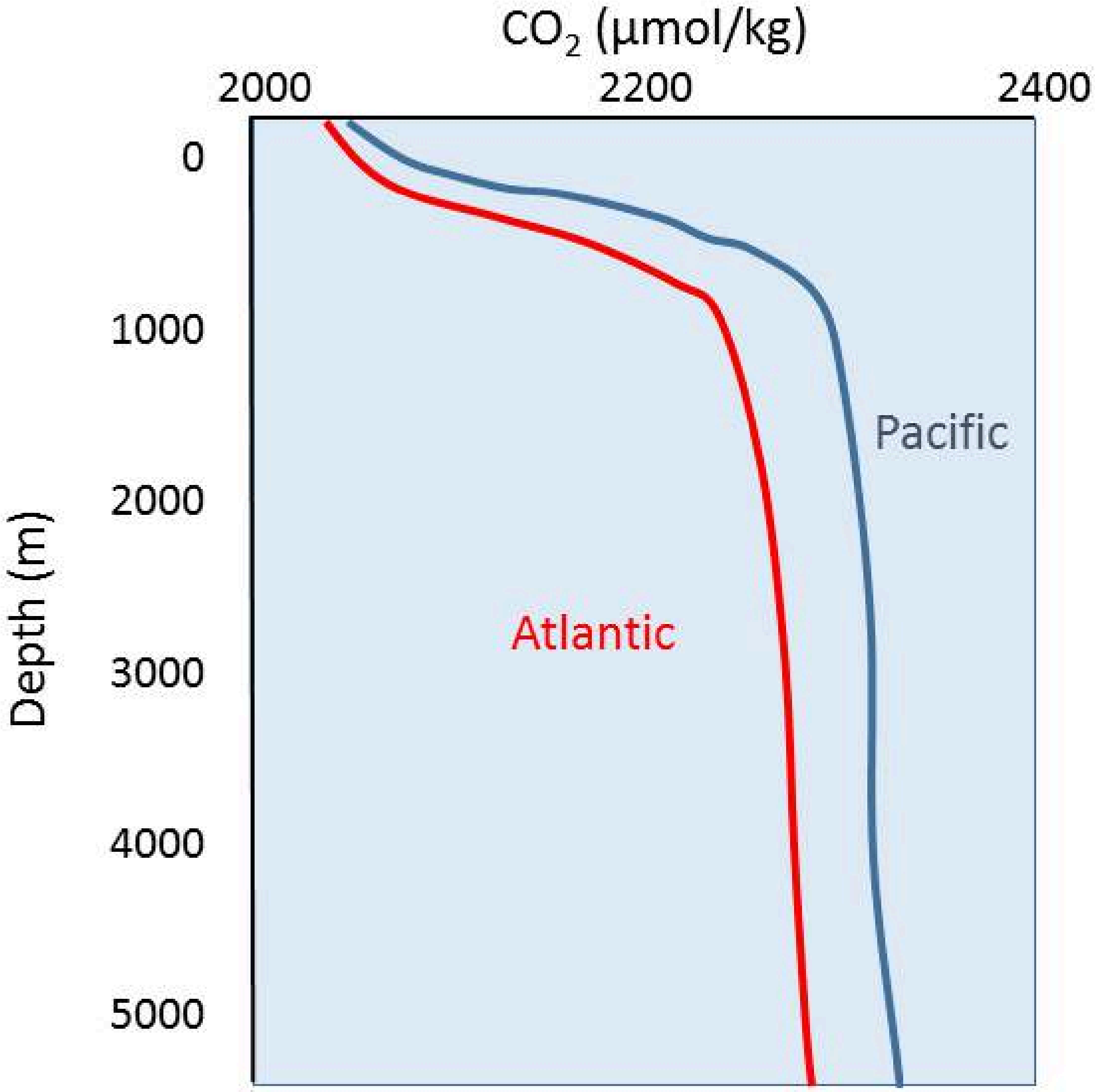


Figure 5.5.1 *Representative carbon dioxide profiles for the Pacific and Atlantic oceans (PW).*

But the behavior of carbon dioxide in the ocean is more complex than the figure above would suggest. When CO_2 gas dissolves in the ocean, it interacts with the water to produce a number of different compounds according to the reaction below:

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \leftrightarrow 2H^+ + CO_3^{2-}$$

CO_2 reacts with water to produce carbonic acid (H_2CO_3), which then dissociates into bicarbonate (HCO_3^-) and hydrogen ions (H^+). The bicarbonate ions can further dissociate into carbonate (CO_3^{2-}) and additional hydrogen ions (Figure 5.5.2).

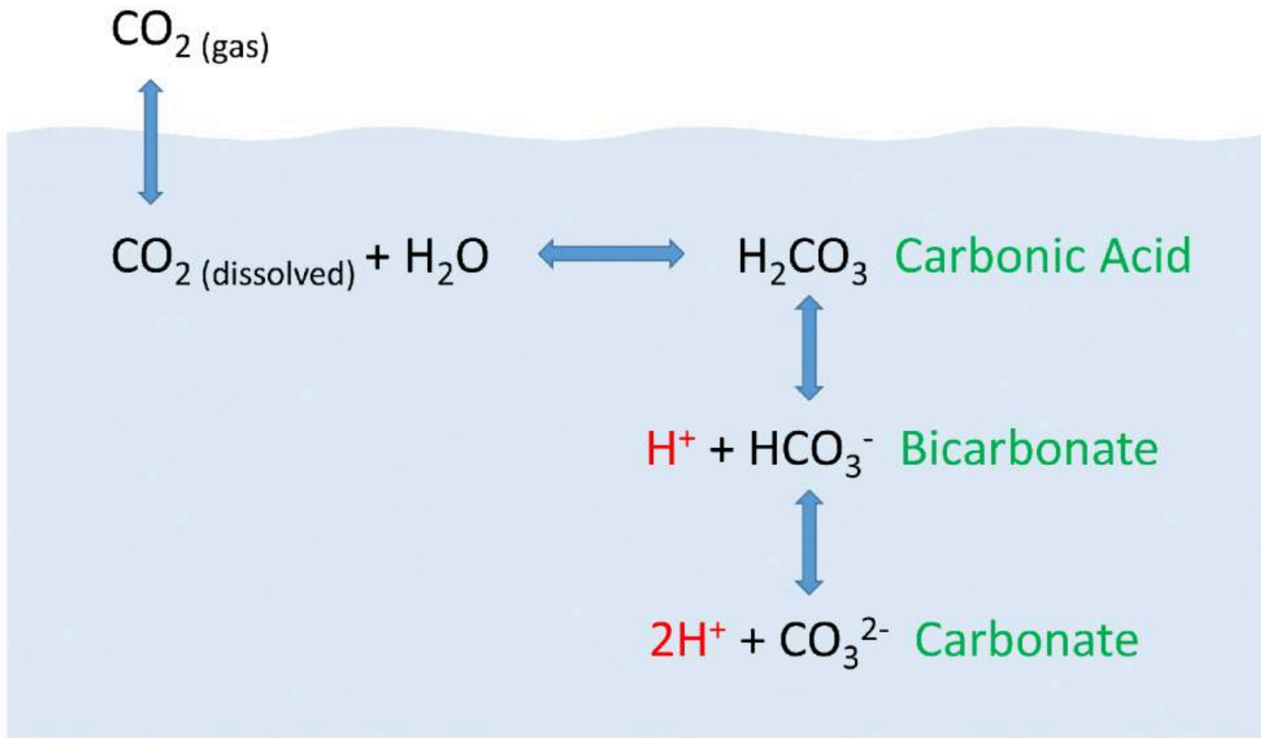


Figure 5.5.2 *The fate of dissolved carbon dioxide in the oceans. Most of the carbon ends up in the form of bicarbonate (PW).*

Most of the CO_2 dissolving or produced in the ocean is quickly converted to bicarbonate. Bicarbonate accounts for about 92% of the CO_2 dissolved in the ocean, and carbonate represents around 7%, so only about 1% remains as CO_2, and little gets absorbed back into the air. The rapid conversion of CO_2 into other forms prevents it from reaching equilibrium with the atmosphere, and in this way, water can hold 50-60 times as much CO_2 and its derivatives as the air.

CO_2 and pH

The equation above also illustrates carbon dioxide's role as a buffer, regulating the pH of the ocean. Recall that pH reflects the acidity or basicity of a solution. The pH scale runs from 0-14, with 0 indicating a very strong acid, and 14 representing highly basic conditions. A solution with a pH of 7 is considered neutral, as is the case for pure water. The pH value is calculated as the negative logarithm of the hydrogen ion concentration according to the equation:

$$pH = -\log_{10}[H^+]$$

Therefore, a high concentration of H^+ ions leads to a low pH and acidic condition, while a low H^+ concentration indicates a high pH and basic conditions. It should also be noted that pH is described on a logarithmic scale, so every one point change on the pH scale actually represents an order of magnitude (10 x) change in solution strength. So a pH of 6 is 10 times more acidic than a pH of 7, and a pH of 5 is 100 times (10 x 10) more acidic than a pH of 7.

Carbon dioxide and the other carbon compounds listed above play an important role in buffering the pH of the ocean. Currently, the average pH for the global ocean is about 8.1, meaning seawater is slightly basic. Because most of the inorganic carbon dissolved in the ocean exists in the form of bicarbonate, bicarbonate can respond to disturbances in pH by releasing or incorporating hydrogen ions into the various carbon compounds. If pH rises (low $[H^+]$), bicarbonate may dissociate into carbonate, and release more H^+ ions, thus lowering pH. Conversely, if pH gets too low (high $[H^+]$), bicarbonate and carbonate may incorporate some of those H^+ ions and produce bicarbonate, carbonic acid, or CO_2 to remove H^+ ions and raise the pH. By shuttling H^+ ions back and forth between the various compounds in this equation, the pH of the ocean is regulated and conditions remain favorable for life.

CO_2 and Ocean Acidification

In recent years there has been rising concern about the phenomenon of **ocean acidification**. As described in the processes above, the addition of CO_2 to seawater lowers the pH of the water. As anthropogenic sources of atmospheric CO_2 have increased since the Industrial Revolution, the oceans have been absorbing an increasing amount of CO_2, and researchers have documented a decline in ocean pH from about 8.2 to 8.1 in the last century. This may not appear to be much of a change, but remember that since pH is on a logarithmic scale, this decline represents a 30% increase in acidity. It should be noted that even at a pH of 8.1 the ocean is not actually acidic; the term “acidification” refers to the fact that the pH is becoming lower, i.e. the water is moving towards more acidic conditions.

Figure 5.5.3 presents data from observation stations in and around the Hawaiian Islands. As atmospheric levels of CO_2 have increased, the CO_2 content of the ocean water has also increased, leading to a reduction in seawater pH. Some models suggest that at the current rate of CO_2 addition to the atmosphere, by 2100 ocean pH may be further reduced to around 7.8, which would represent more than a 120% increase in ocean acidity since the Industrial Revolution.

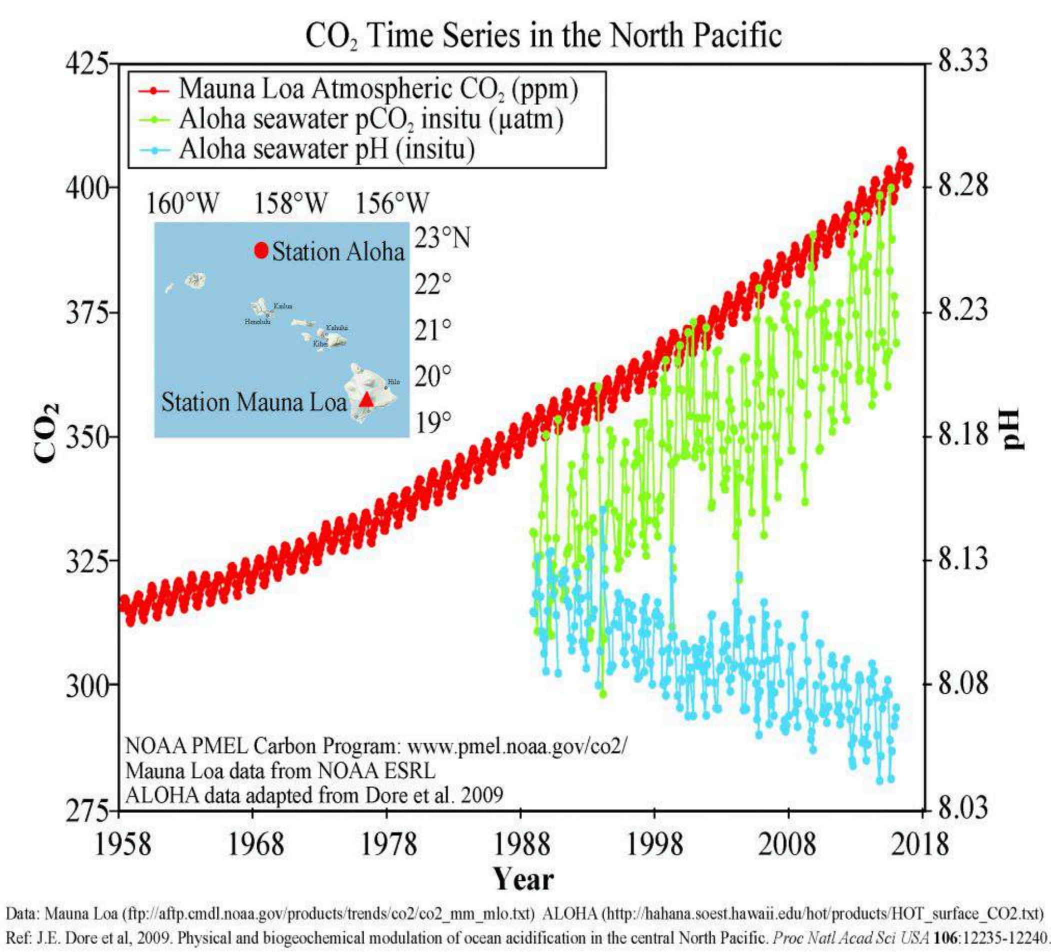


Figure 5.5.3 *Changes in atmospheric CO_2 (red), seawater CO_2 (green) and pH (blue) in the Hawaiian Islands (NOAA PMEL).*

Why is this important? Declining pH can impact many biological systems. Of particular concern are organisms that secrete calcium carbonate shells or skeletons, such as corals, shellfish, and may planktonic organisms. At lower pH levels, calcium carbonate dissolves, eroding the shells and skeletons of these organisms (Figure 5.5.4).

Figure 5.5.5 *The results of an experiment placing the calcium carbonate shells of pterapods in seawater with a pH of 7.8, the projected ocean pH for the year 2100 under current rates of acidification. The top row shows the shells before the experiment, and the bottom row shows the dissolution of the shells after 45 days of exposure (NOAA).*

Not only does a declining pH lead to increased rates of dissolution of calcium carbonate, it also diminishes the amount of free carbonate ions in the water. The relative proportions of the different carbon compounds in seawater is dependent on pH (Figure 5.5.6). As pH declines, the amount of carbonate declines, so there is less available for organisms to incorporate into their shells and skeletons. So ocean acidification both dissolves existing shells and makes it harder for shell formation to occur.

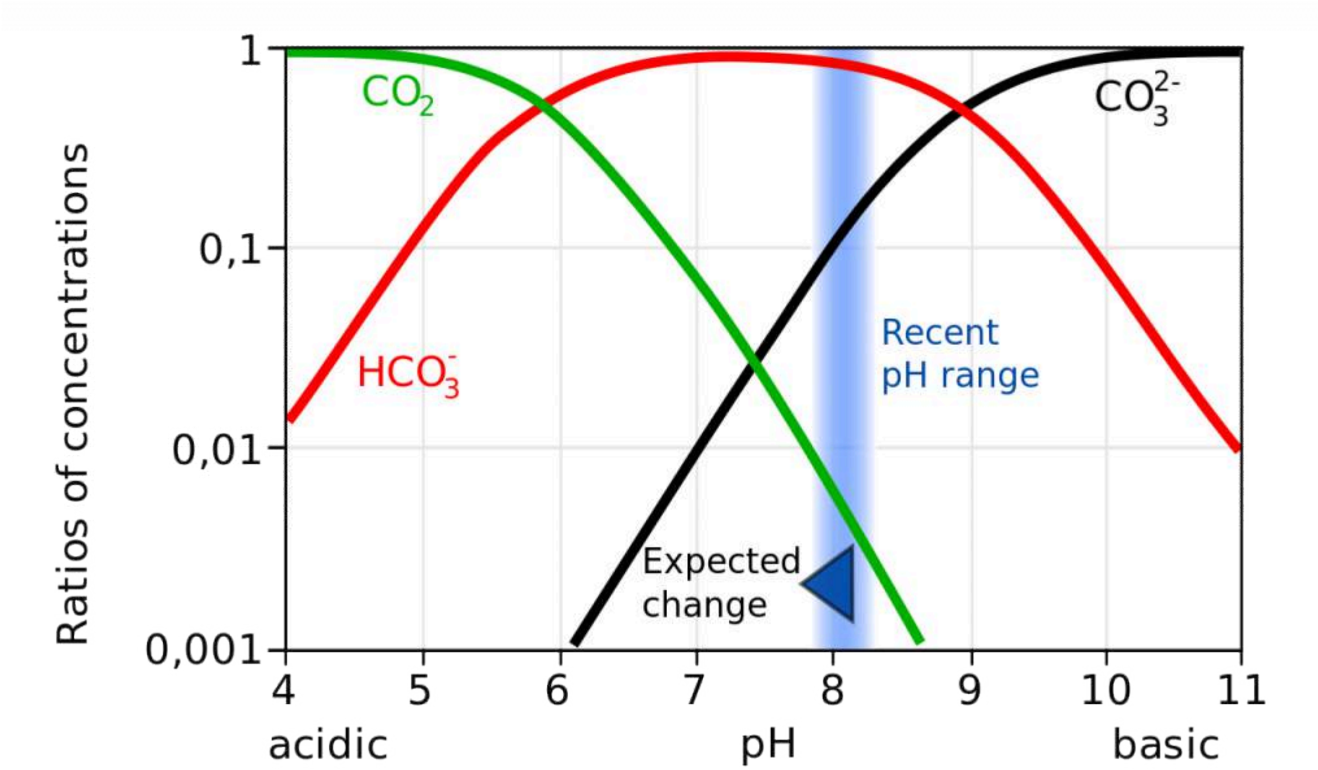


Figure 5.5.6 *Proportions of carbon compounds in the ocean at various pH levels. As the ocean pH declines, the proportion of carbonate ions also declines, reducing rates of shell formation (NOAA).*

Additional links for more information:

- NOAA Ocean Acidification Program website http://oceanacidification.noaa.gov/

5.6 Nitrogen and Nutrients

Nitrogen is the most abundant gas in the atmosphere, and like the other atmospheric gases it dissolves into the surface layers of the ocean. But most marine organisms cannot directly utilize dissolved nitrogen in the form in which it exists in air (N_2), so it must first be converted into other nitrogenous products by marine bacteria (Figure 5.6.1). Some bacteria (cyanobacteria) take the dissolved N_2 and convert it into ammonium (NH_4^+) through nitrogen fixation. Some of this ammonium can be used directly by phytoplankton, but the majority of it is converted by bacteria into nitrite (NO_2^{2-}) or nitrate (NO_3^-) through the process of nitrification. Nitrate is the main nitrogenous compound utilized by primary producers in the ocean; it is a major **nutrient** required for photosynthesis. Note that in this context, a nutrient refers to a chemical needed to support photosynthesis and primary production. It does not refer to the nutritional needs of consumer organisms. The nitrogen taken in by phytoplankton gets passed on to consumer organisms, and then gets returned to the ocean through decomposition of wastes and organic matter as these organisms die and sink into deeper water. Finally, the ammonium, nitrate and nitrite can undergo denitrification by yet another group of bacteria and get converted back into N_2, which can reenter the cycle or be exchanged with the atmosphere (Figure 5.6.1).

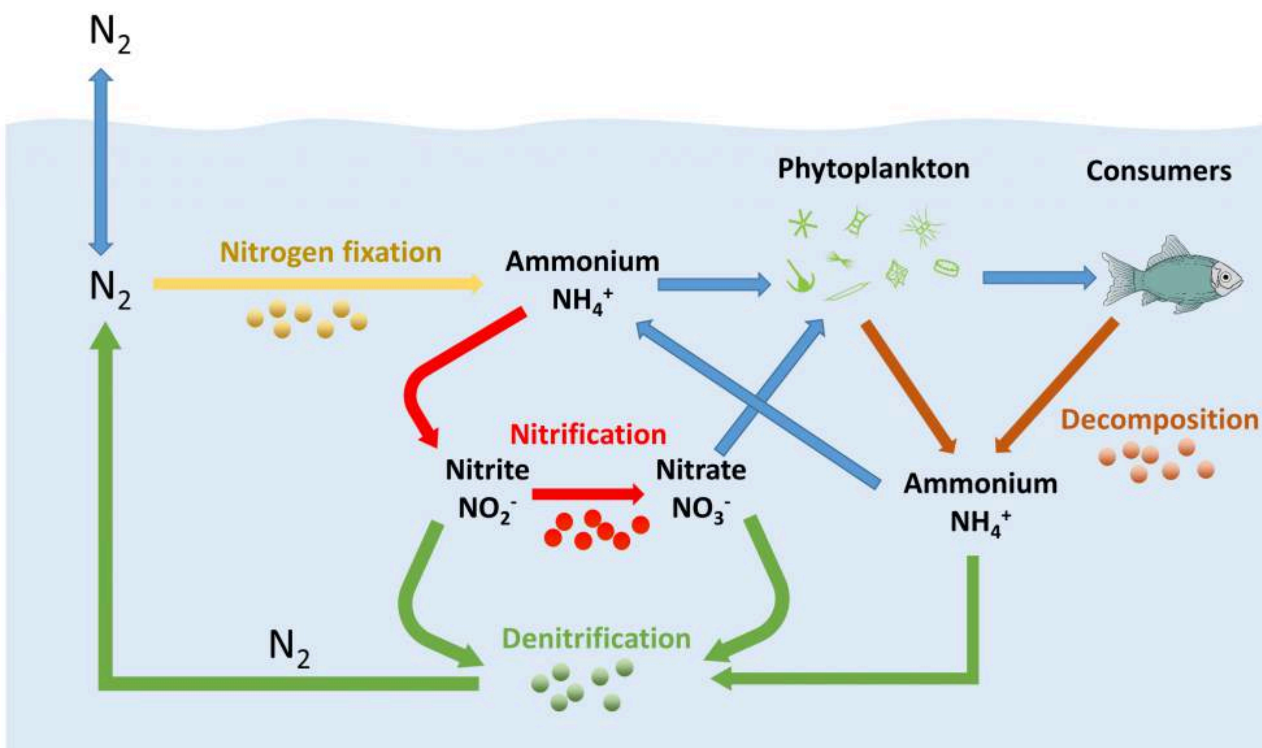


Figure 5.6.1 *Simplified nitrogen cycle in the ocean. Colored dots represent the marine bacteria responsible for nitrogen cycling (PW).*

Since nitrate is one of the most important nutrients, for now we will focus only on nitrate as we discuss general nutrient patterns in the ocean. Other important nutrients, such as phosphate and silica, show similar patterns to nitrate, and will be discussed in the section on primary production (chapter 7).

A representative nutrient (nitrate) profile is shown below (Figure 5.6.2). Since nutrients are rapidly used in biological processes, they are non-conservative, and their concentrations vary regionally and seasonally. Nutrient concentrations are low at the surface, because that is where the primary producers are located; the nutrients are rapidly consumed and they do not have the chance to accumulate. Nutrient levels increase at depth, as they are no longer being consumed by producers, and they are being regenerated through the decomposition of organic material by bacteria.

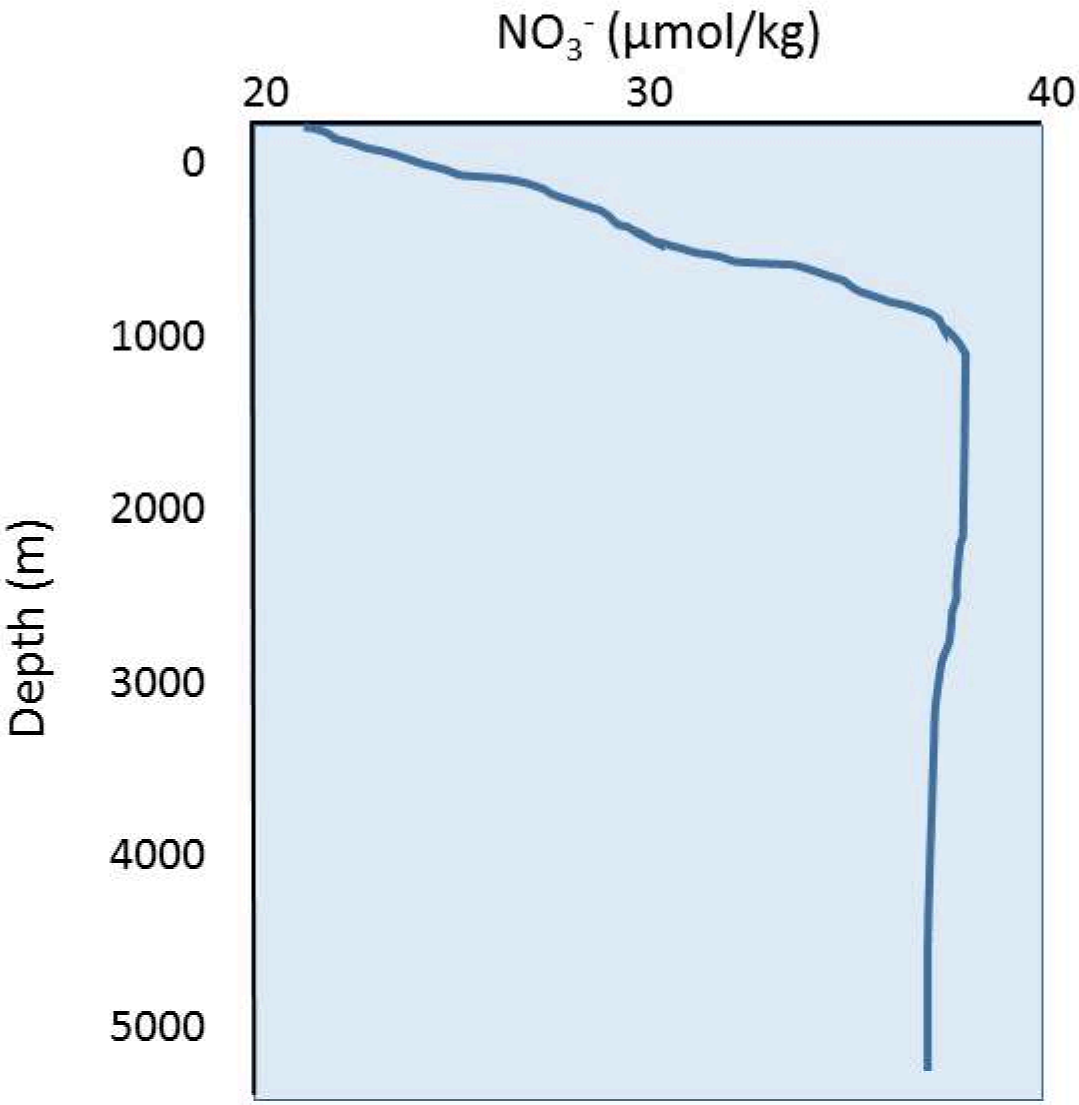


Figure 5.6.2 *Representative nutrient (nitrate) profile for the open ocean (PW).*

Comparisons of nutrient and dissolved oxygen profiles between the Pacific and Atlantic Oceans reveal some interesting differences (Figure 5.6.3). In general, the Atlantic has more dissolved oxygen but lower nutrient concentrations than the Pacific. Water masses form in the North Atlantic that are very cold and dense, so the water sinks to the bottom. This water will then spend the next thousand years or more moving along the seafloor from the Atlantic, to the Indian, and finally into the Pacific Ocean (see section 9.8). This water is initially oxygen-rich surface water, and as it sinks it brings oxygen to the deep seafloor. As the bottom water moves across the ocean basins, oxygen is removed through respiration and decomposition, and by the time it arrives in the Pacific it has been depleted of much of its oxygen. At the same time, decomposition of sinking organic matter adds nutrients to the deep water as it moves through the oceans, so nutrients accumulate and the Pacific water becomes nutrient-rich. Comparison of the ratios of oxygen to nutrients in the deep water can therefore provide an indication of the age of the water, i.e. how much time has passed since it initially sank from the surface in the North Atlantic. Water with a high oxygen and low nutrient content is relatively young, while older water will have less oxygen but higher nutrient concentrations.

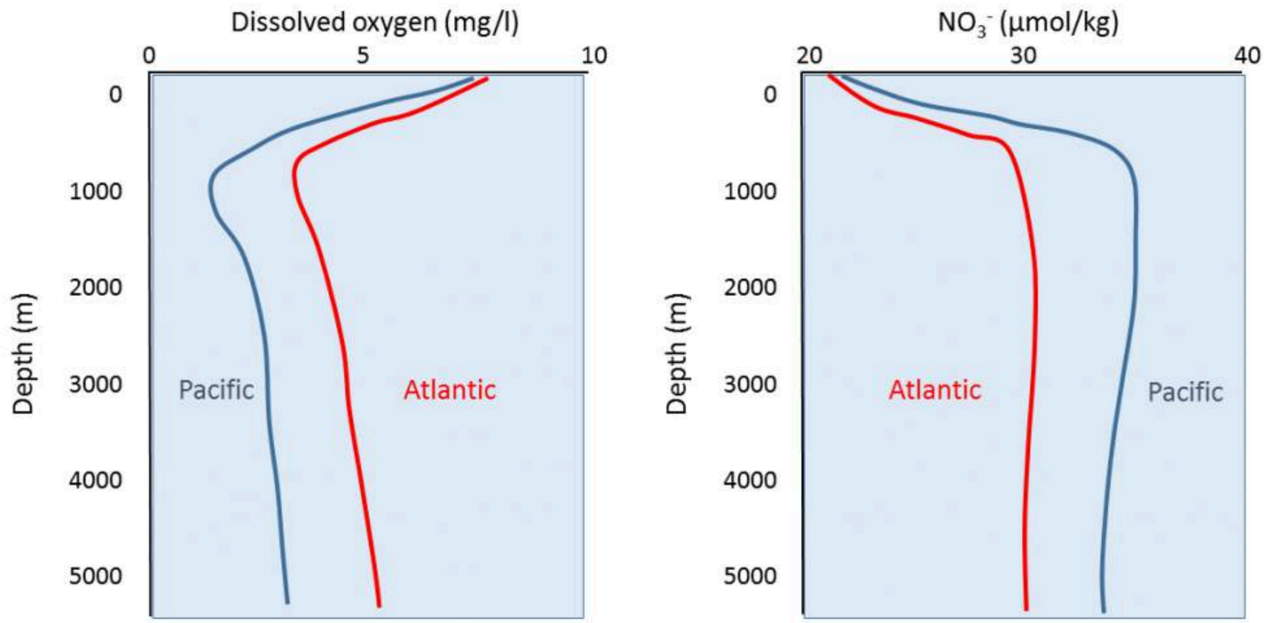


Figure 5.6.3 *Dissolved oxygen (left) and nutrient (right) profiles for the Pacific and Atlantic Oceans. As water circulates from the Atlantic to the Pacific, oxygen is consumed while nutrients accumulate (PW).*

5.7 Classifying Elements in Seawater

Now that we are familiar with the abundance and profiles of some common dissolved substances in the ocean, we can try to classify these materials based on their distribution. Dissolved materials are usually classified into one of several groups:

Conservative Elements are those whose concentration is relatively constant across the ocean, both vertically and horizontally. This category would include the major ions in seawater, such as sodium or chloride, which have very long residence times and whose concentration only changes through the addition or removal of fresh water (see section 5.3).

Nutrient-like Elements have a distribution similar to the profile we described for nitrate in the previous section. Concentrations of these substances are low at the surface, as they are rapidly used up by biological processes. Below the photic zone, concentrations of these materials will increase, as decomposition by bacteria cycles them back into the water column.

Scavenged Elements are those that react with other particles and are adsorbed to the particle surface. When the particles sink, those elements are removed to the sediment. Scavenged profiles generally show higher abundance at the surface, where the materials will enter the ocean, but declining levels with depth as they are removed by the sinking particles. This profile is common for metals, such as aluminum or lead.

Stable Gases dissolve into the ocean from the atmosphere. Because these substances are not very reactive, the ocean water becomes saturated with these gases. Since solubility of a gas increases in colder water, these gases are often found in greater concentrations in deep, cold water, and in lower concentrations in warmer surface water.

Additional links for more information

- Interactive periodic table of the elements, where you can click on any element to see its vertical profile in the ocean: https://www.mbari.org/science/upper-ocean-systems/chemical-sensor-group/periodic-table-of-elements-in-the-ocean/

CHAPTER 6: PHYSICAL OCEANOGRAPHY

Chapter 6: Physical Oceanography

Learning Objectives

After reading this chapter you should:

- understand the relationship between depth and pressure
- understand the terms thermocline, pycnocline, and mixed layer
- be able to describe temperature profiles for different regions of the ocean
- understand how density is related to temperature, pressure, and salinity
- understand how density changes with depth and latitude
- understand how sound travels through water, and what physical factors impact sound transmission
- understand the concept of the SOFAR channel
- be familiar with the electromagnetic spectrum
- understand the limitations of light penetration through water

Picture yourself swimming in the ocean. Are you imaging frolicking in warm, crystal-clear water? Well scrap that idea, as that idyllic setting represents the conditions in only a tiny portion of the global ocean. The conditions in the vast majority of the ocean are nothing like those at the surface; most of the ocean is very cold, dark, and subjected to crushing pressures. However, even in those hostile conditions, there is a great diversity of organisms managing to make a living at depth. This chapter examines the physical environment of the ocean, including processes that are important for marine life, such as temperature, and light and sound transmission. The field of physical oceanography typically also includes the fluid properties of ocean water, such as waves, tides, and circulation. Those topics will be addressed in subsequent chapters.

6.1 Pressure

When we talk about pressure in the ocean, we are referring to hydrostatic pressure, which is a result of the weight of the water column pressing down on an object due to gravity. The deeper you go, the more water that is above you, and the greater the weight (and thus pressure) of that water. At the surface we experience one atmosphere of pressure (1 atm = 101.3 kPa) due to the weight of the atmosphere above us. As you descend into the ocean, pressure increases linearly with depth; there is an increase in pressure of 1 atm for every 10 m increase in depth. So at 1000 m depth the pressure would be 101 atm (100 atm of pressure due to the 1000 m depth, plus the 1 atm that is present at the surface). This leads to very high pressure in the deeper parts of the ocean; if you consider that the average depth of the ocean is about 3800 m, the pressure at that depth would be 381 times greater than the pressure at the surface. The consequences of this great pressure can be seen in the video below, showing a hole being cut into a pipe at a depth of 6 miles (about 9.6 km). The pressure outside the pipe is over 960 atm, which is far higher than the pressure inside the pipe, and this pressure differential has serious implications for a passing crab...

A YouTube element has been excluded from this version of the text. You can view it online here: https://rwu.pressbooks.pub/webboceanography/?p=175

There are several important consequences of high pressure at depth. First, due to Boyle's Law, which states that the volume of a gas is inversely related to pressure, high pressure will act to compress air spaces, such as the lungs of a diving animal (or person), or the space inside a submarine. Submarines and submersibles must therefore have very strong hulls to resist this compression at extreme depths. Second, Henry's Law provides that at higher pressures a fluid will contain more dissolved gas. As seen in section 5.4, this means that deeper, high pressure water may contain more dissolved gases than surface water.

This also has implications for human divers. According to Henry's Law, if you increase pressure, you increase the amount of gas that can dissolve in a fluid (such as blood). Conversely, when you reduce pressure, the fluid holds less dissolved gas, and the excess gas will leave the solution, often in the form of bubbles. This is what happens when you open a bottle of a carbonated beverage. The contents in the bottle are sealed under pressure, and as you open the bottle, you release the pressure, and the fluid can no longer hold all of the CO_2 that was dissolved in it, so the CO_2 escapes, forming bubbles. Decompression sickness, or "the bends" occurs in SCUBA divers if they ascend too quickly after breathing compressed air. A slow ascent allows this excess gas to be removed from the blood and exhaled, but if the diver ascends too quickly, these gases will come out of solution and form bubbles in the blood that congregate near the joints, causing intense pain and perhaps death.

Additional links for more information

- How can deep-diving marine mammals avoid "the bends"?: http://www.oneworldoneocean.com/blog/entry/ocean-stemulation-how-marine-mammals-avoid-the-bends

6.2 Temperature

Generally ocean temperatures range from about -2^{o} to 30^{o} C (28-86^{o} F). The warmest water tends to be surface water in low latitude regions, while the surface water at the poles is obviously much colder (Figure 6.2.1). Note that at equivalent latitudes, water on the eastern side of the ocean basins is colder than the water on the western side. This has to do with the pattern of surface currents, as described in section 9.1. Even though surface water can be quite warm, most of the water in the oceans is deeper, colder water, so that the average temperature of the entire ocean is about 4^{o} C, which is roughly the temperature inside your refrigerator.

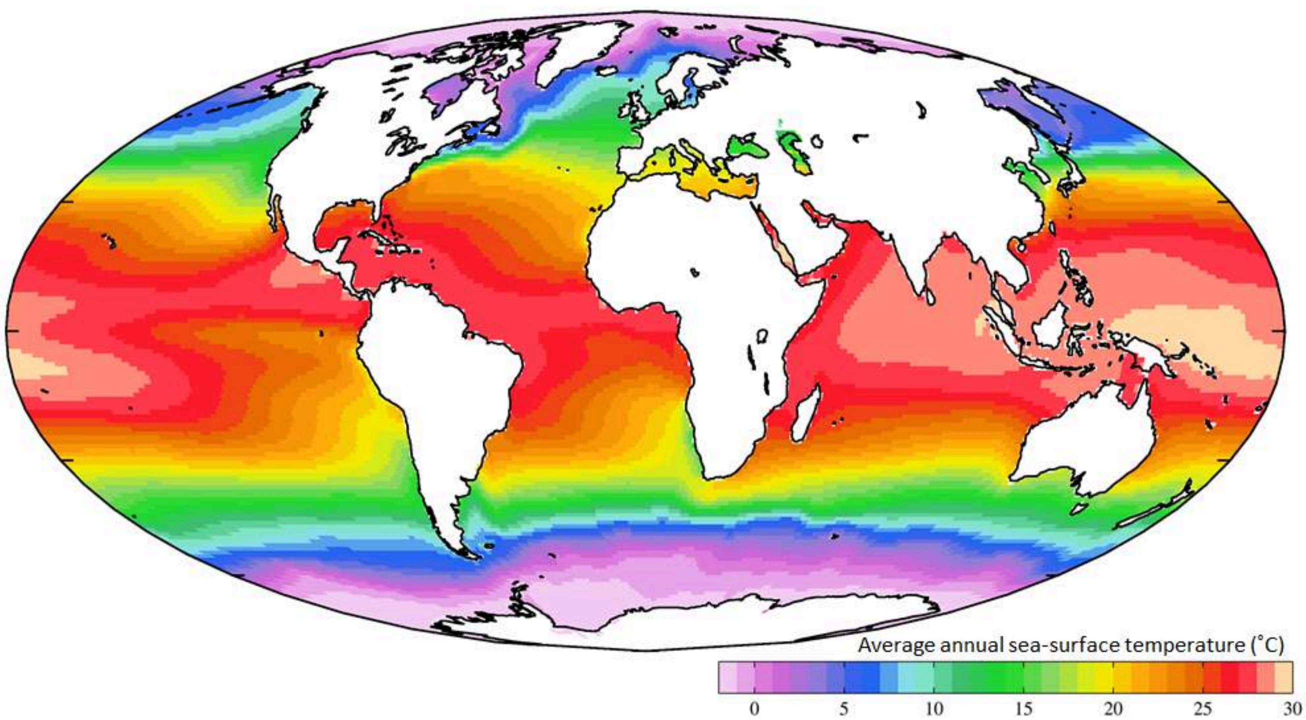


Figure 6.2.1 *Global average annual sea surface temperature (Stephen Earle, "Physical Geology").*

A typical temperature profile for open ocean, mid-latitude water is shown in Figure 6.2.2. Water is warmest at the surface, as it is warmed by the sun, and the sun's rays can only penetrate depths less than 1000 m (section 6.5). Since the surface water is warmer it is also less dense than the deep water (section 6.3), so it remains at the surface where it can be warmed even more. Temperature is fairly constant in the upper 100-200 m in what is called the **mixed layer**. The mixed layer results from surface winds, waves, and currents that mix the upper water and distribute the heat throughout this layer. Below the mixed layer there is a rapid decline in temperature over a fairly narrow increase in depth. This is called the **thermocline**. Below the thermocline the deep ocean temperature is fairly constant at about 2^{o} C, continuing down to the bottom. There is little temperature change in the deep ocean, as it is far removed from significant heat sources, making it one of the most thermally stable regions on earth. Temperature may fluctuate by less than half a degree per year in the deep ocean (Figure 6.2.3).

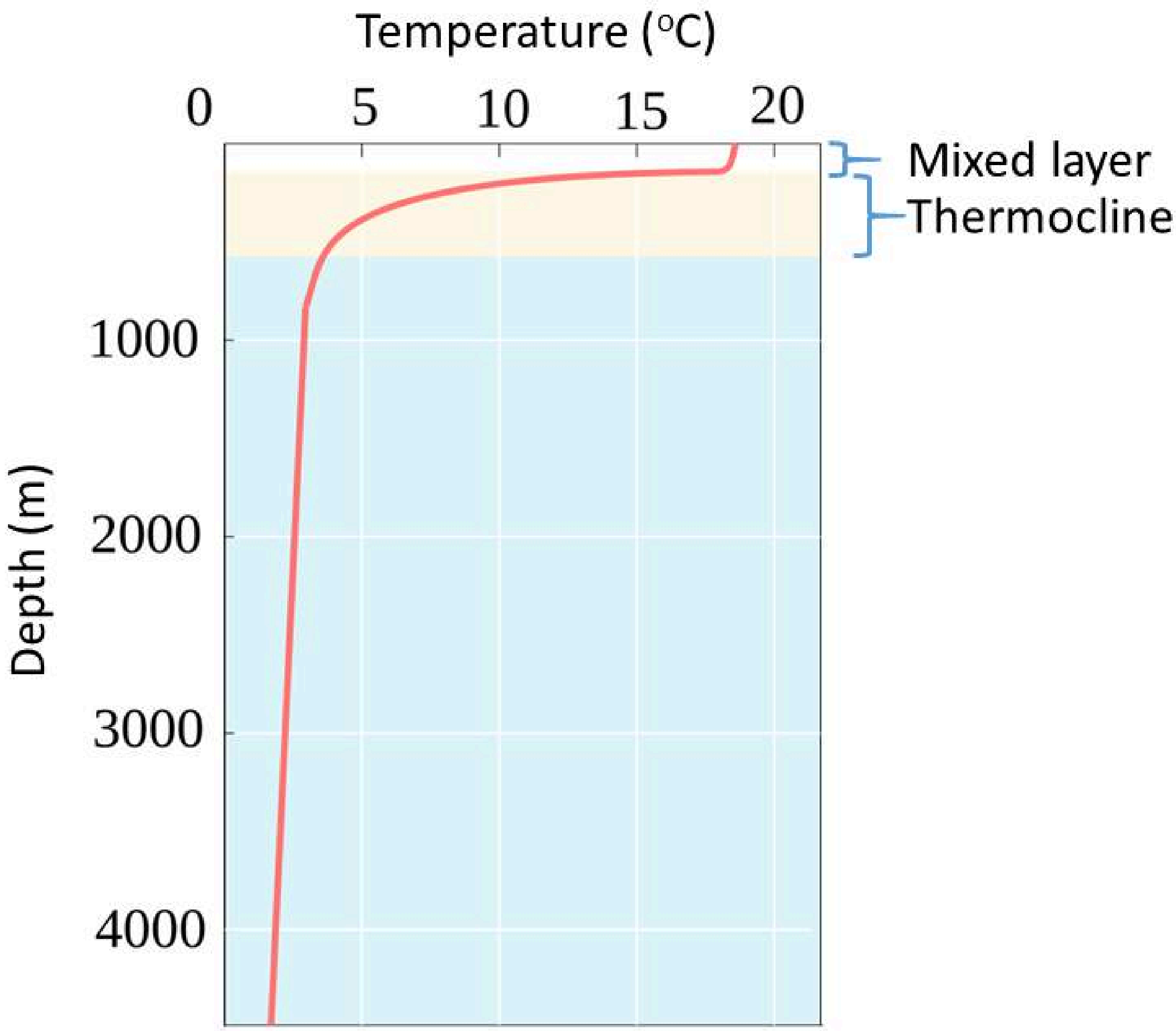


Figure 6.2.2 *Typical open ocean temperature profile for a mid-latitude region, showing the mixed layer, steep thermocline, and relatively stable temperature at depth (Public domain via Wikimedia Commons).*

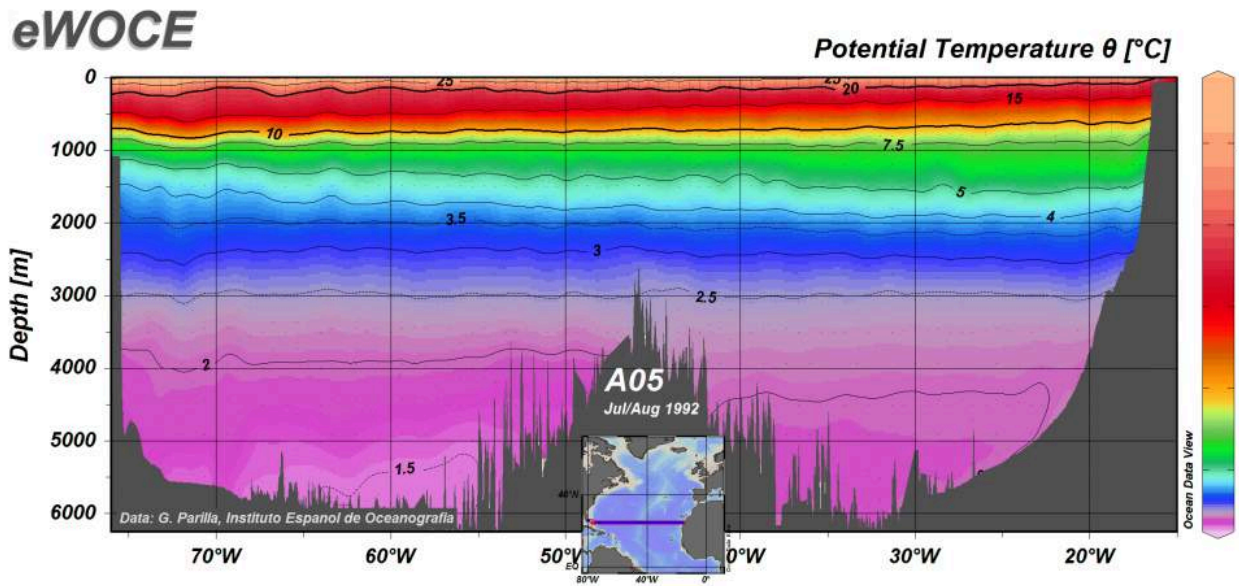


Figure 6.2.3 *Temperature profile across the Atlantic Ocean from the coast of Florida to the coast of Africa (inset). There is rapid temperature change near the surface in the thermocline zone, but the deep water temperature is fairly stable (eWOCE, http://www.ewoce.org/gallery/eWOCE_Tables.html#Atlantic).*

Temperature profiles vary at different latitudes, as the surface water is warmer near the equator and colder at the poles. In low latitude tropical regions the sea surface is much warmer, leading to a highly pronounced thermocline (Figure 6.2.4). Additionally, there is not much seasonal change in surface temperature in tropical regions, so there is little seasonal change in the profiles. In high latitude (polar) regions, there is little difference between the surface temperature and the deep water temperature, and temperature is fairly constant (and cold) at all depths. Polar waters therefore lack a strong thermocline, and as with tropical water, there is little seasonal change in temperatures. Mid-latitude temperate regions show greater seasonal fluctuations in surface temperature than the poles or the tropics; an 8-15^{o} C difference from summer to winter in temperate zones, compared to only ~2^{o} C in polar and tropical areas. In temperate regions, the surface water is much warmer in the summer and the thermocline is more pronounced compared to the winter months. But in the winter the thermocline is *deeper* at mid-latitudes than it is in the summer. This is because winter storms churn up the surface water more than occurs in the summer, creating a deeper mixed layer and thus a deeper thermocline (Figure 6.2.5).

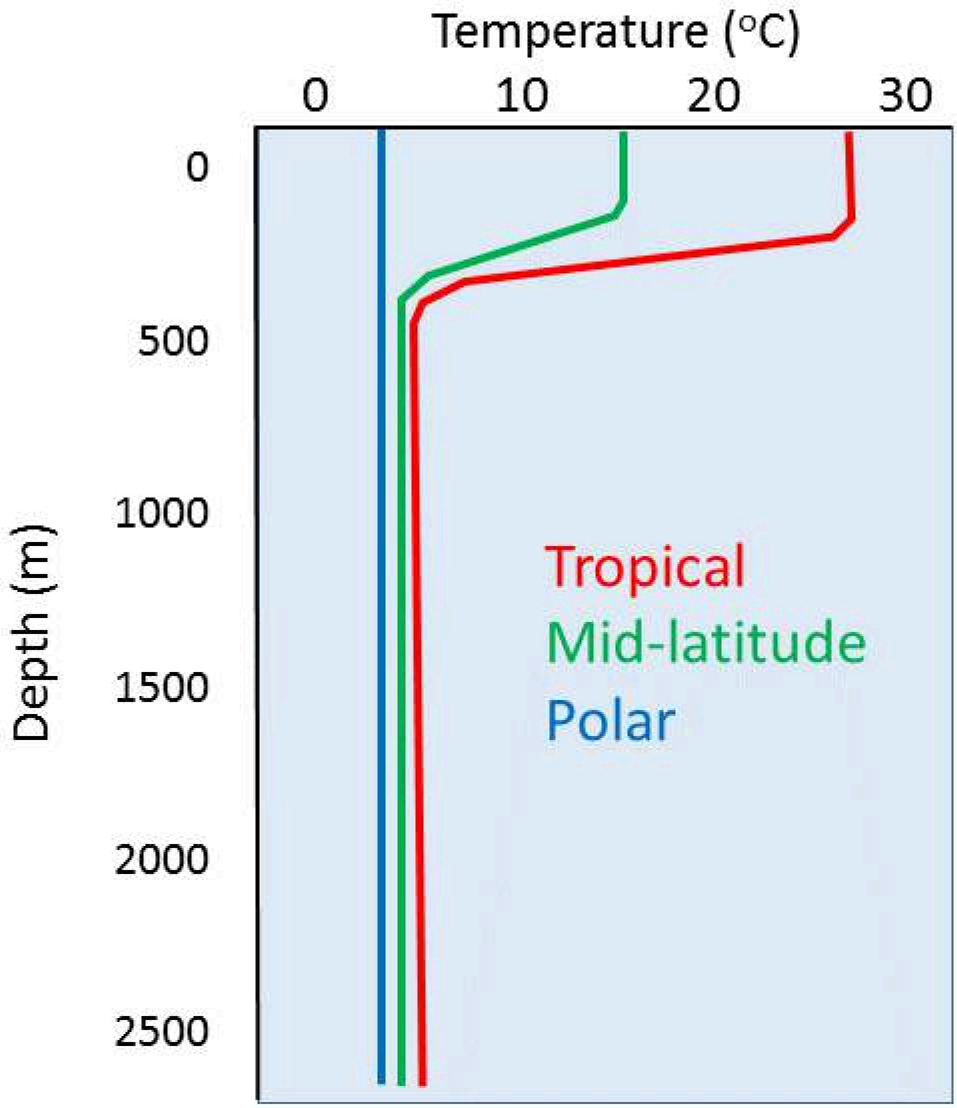


Figure 6.2.4 *Representative temperature profiles for tropical, mid-latitude, and polar regions (PW).*

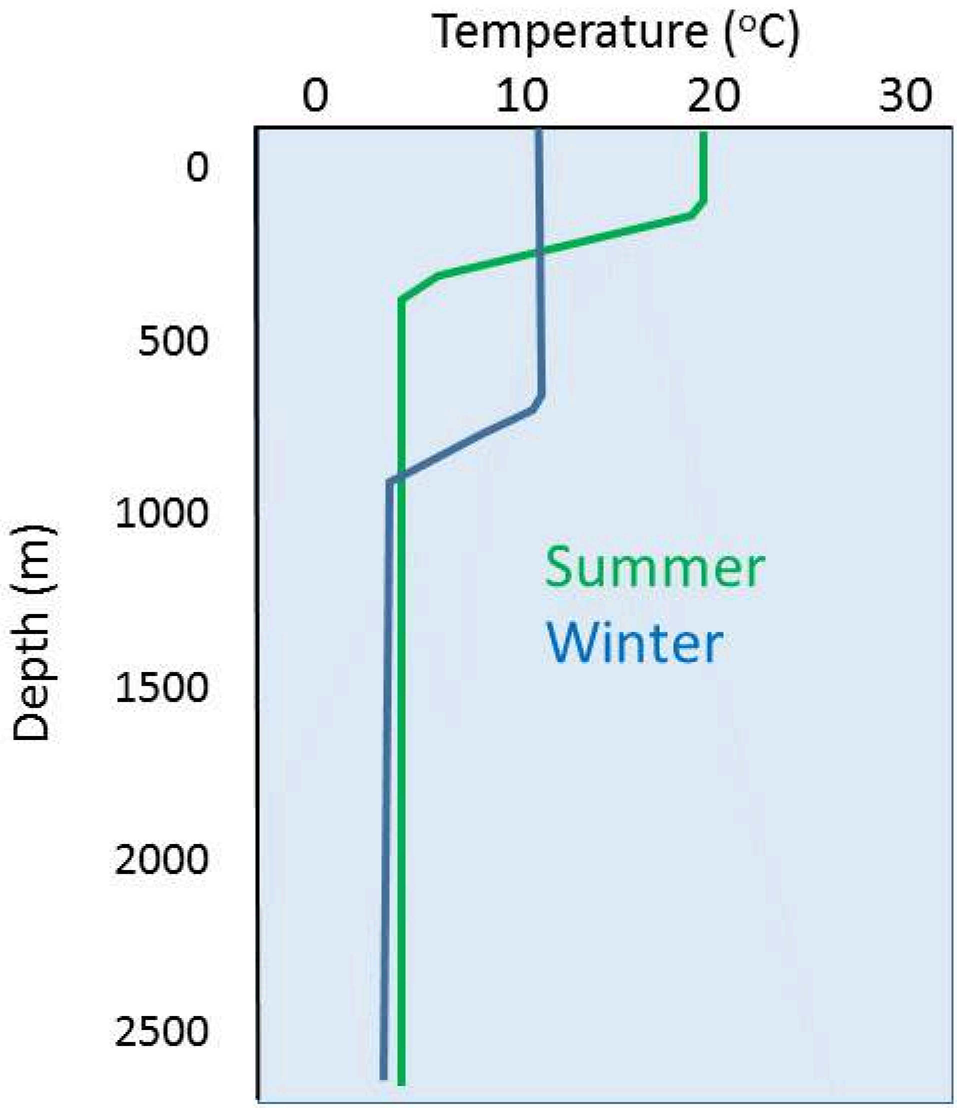


Figure 6.2.5 *In temperate regions, the mixed layer is deeper and the thermocline less pronounced in the winter compared to the summer (PW).*

Due to the high heat capacity of water, daily fluctuations in ocean temperature are fairly insignificant.

6.3 Density

Density refers to the amount of mass per unit volume, such as grams per cubic centimeter (g/cm^3). The density of fresh water is 1 g/cm^3 at 4^o C (see section 5.1), but the addition of salts and other dissolved substances increases surface seawater density to between 1.02 and 1.03 g/cm^3. The density of seawater can be increased by reducing its temperature, increasing its salinity, or increasing the pressure. Pressure has the least impact on density as water is fairly incompressible, so pressure effects are not very significant except at extreme depths. However, if not for the slight compression of water due to pressure, sea level would be approximately 50 m higher than it is today! That leaves temperature and salinity as the primary factors determining density, and of these, temperature has the greatest impact (Figure 6.3.1).

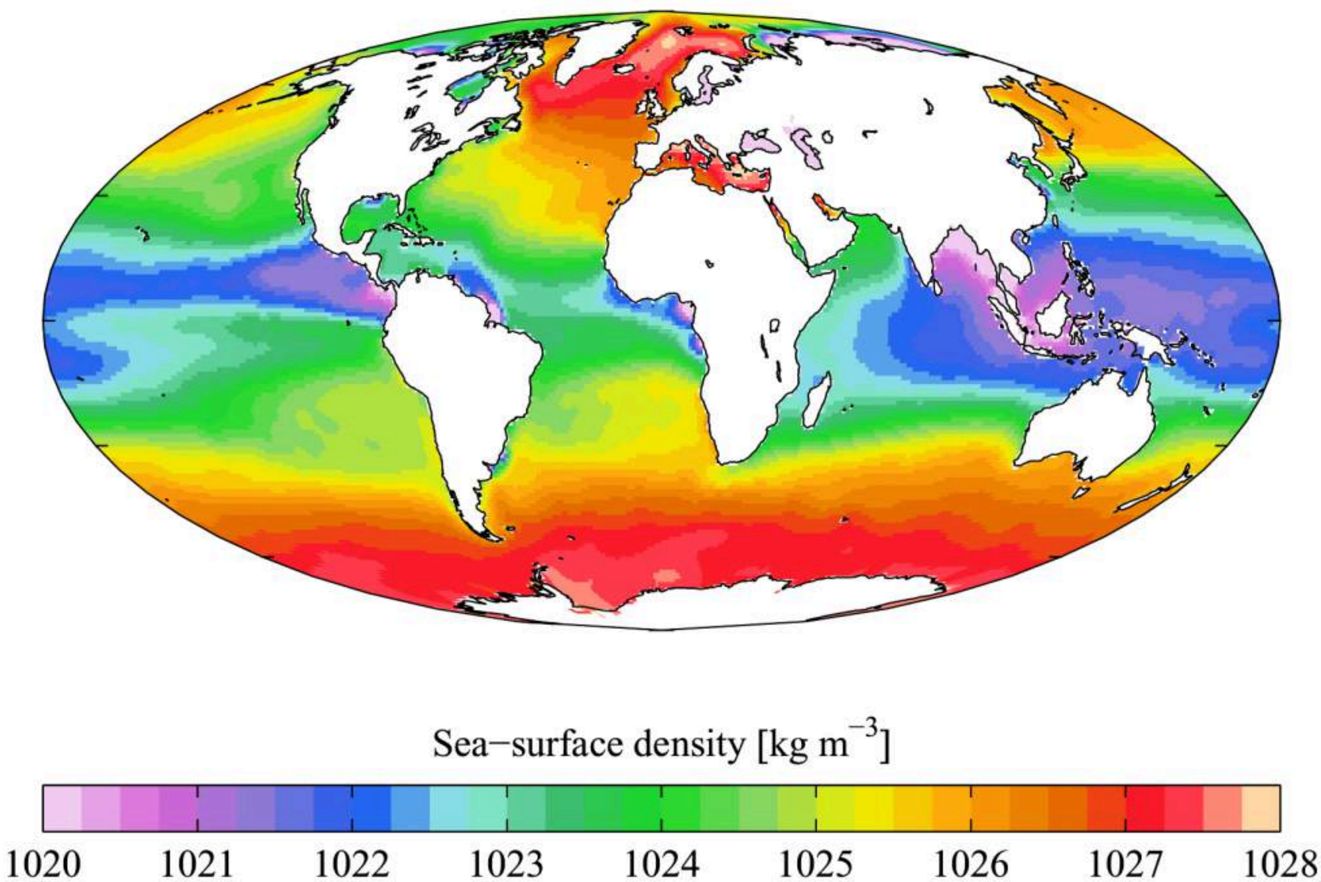


Figure 6.3.1 *Global sea surface density. Colder polar regions display higher densities than warmer tropical zones (By Plumbago (Own work) [CC BY-SA 3.0, or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons).*

Since temperature has the greatest effect on density, density profiles are usually mirror images of temperature profiles (Figure 6.3.2). Density is lowest at the surface, where the water is the warmest. As depth increases, there is a region of rapidly increasing density with increasing depth, which is called the **pycnocline**. The pycnocline coincides with the thermocline, as it is the sudden decrease in temperature that leads to the increase in density. Below the pycnocline, density may be fairly constant (as is temperature), or it may continue to increase slightly towards the bottom.

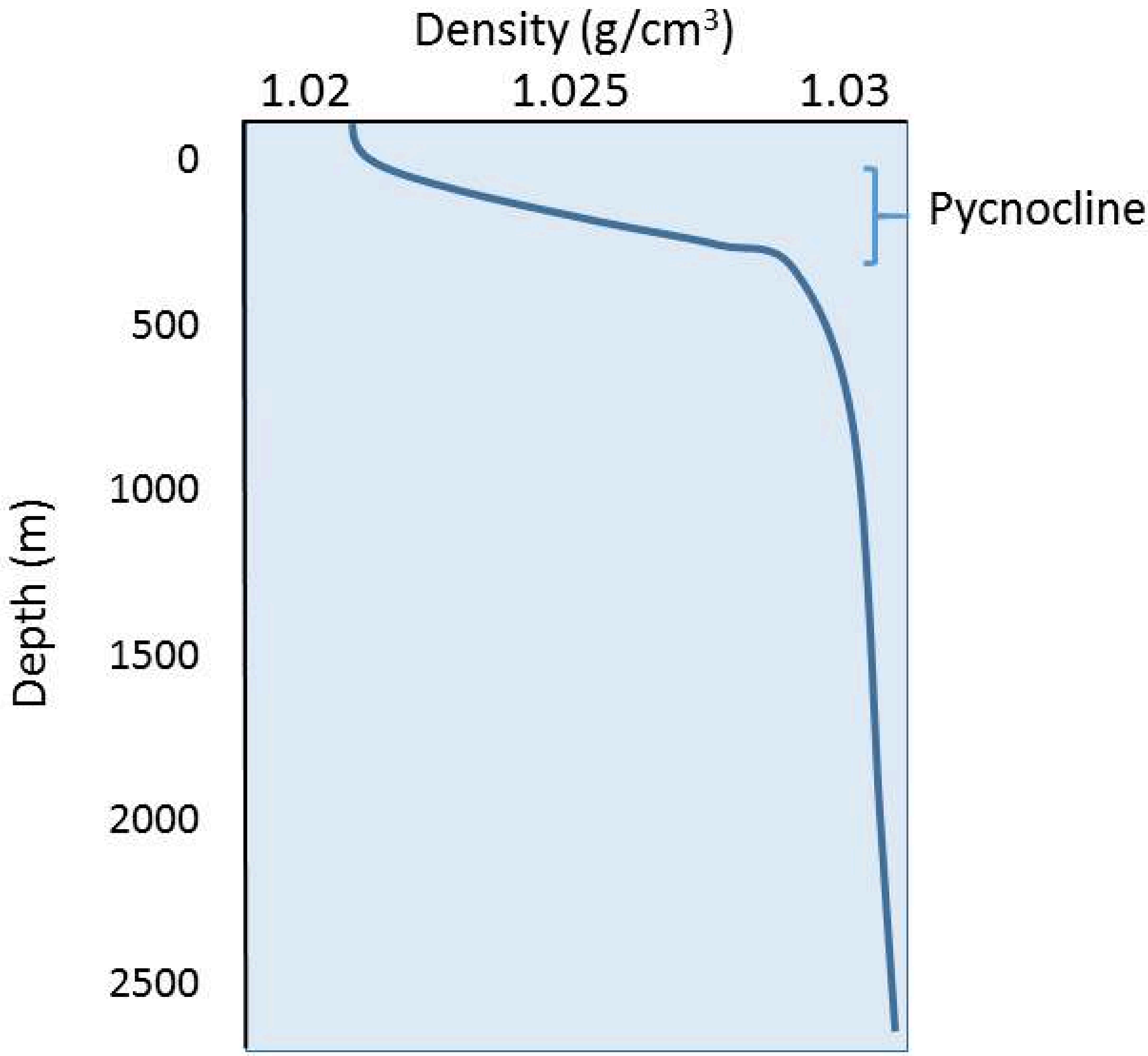


Figure 6.3.2 *Representative density profile for the open ocean at mid-latitudes. The warm surface water causes a decrease in surface density (PW).*

The profile above represents a stable state, or a high degree of stratification, where the warm, low density layer sits atop the colder, denser layer. If denser water happened to form at the surface, the water masses would be unstable, and the denser water would sink to the bottom, to be replaced by less dense water at the surface. This vertical movement of water masses based on density (as determined by temperature and salinity) is referred to as thermohaline circulation, which is the topic of section 9.8. By creating a stratified water column, the thermocline and pycnocline together create a barrier that prevents mixing between the warmer, less dense surface water and the colder, denser bottom water. In this way, nutrient-rich deep water may be prevented from coming to the surface to support primary production.

As with temperature, there are also latitudinal differences in density. In the tropics the surface water is warm and low density, and there is a pronounced thermocline separating it from the colder, denser deep water. As stated above, this stratification prevents nutrient-rich water from reaching the surface and as a result tropical regions often have low productivity. In the high latitudes the water is uniformly cold at all depths, so there is little density stratification. The lack of a pycnocline (or a thermocline) allows cold, nutrient-rich deep water to more easily mix with the surface water, leading to higher primary production in polar regions.

6.4 Sound

Sound is a form of energy transmitted through pressure waves; longitudinal or compressional waves similar to the seismic P-waves we discussed in section 3.3. With ocean sounds, the energy is transmitted via water molecules vibrating back and forth parallel to the direction of the sound wave, and passing on the energy to adjacent molecules. Therefore, sound travels faster and more efficiently when the molecules are closer together and are better able to transfer their energy to neighboring particles. In other words, sound travels faster through denser materials. Since water is much denser than air, the speed of sound in water (about 1500 m/s) is approximately five times faster than the speed in air (around 330 m/s). This helps explain why we sometimes have difficulty localizing the source of a sound that we hear underwater. We localize sound sources when our brains detect the tiny differences in the time of arrival of sounds reaching our ears. A sound coming from our left will reach our left ear a fraction of a second before reaching our right ear. Our brains can process that small difference in time of arrival to recognize the direction from which the sound came. In water, the sound is so much faster that the difference in arrival time between our ears becomes too small for us to interpret, and we lose the ability to localize the source.

However, as with sound in air, the speed of sound in the ocean is not constant; it is influenced by a number of variables including temperature, salinity, and pressure, and an increase in any of these factors will lead to an increase in the speed of sound. We have seen that these variables change with depth and location; so to will the speed of sound differ in different regions of the ocean.

To examine the way the speed of sound changes as a function of depth, we need to consider the vertical profiles for temperature and pressure. At the surface, the pressure is low, but the temperature is at its highest point in the water column. The temperature effects dominate at the surface, so the speed of sound is fast in surface waters. As depth increases, the temperature and the speed of sound decline. Near the bottom, the extreme pressure dominates, and even though temperatures are low, the speed of sound increases with depth. At moderate depths (between a few hundred and one thousand meters) there is a zone where both temperature and pressure are relatively low, so the speed of sound is at a minimum. This zone of minimum speed is called the **SOFAR channel** (**So**und **F**ixing **A**nd **R**anging) or the Deep Sound Channel (Figure 6.4.1).

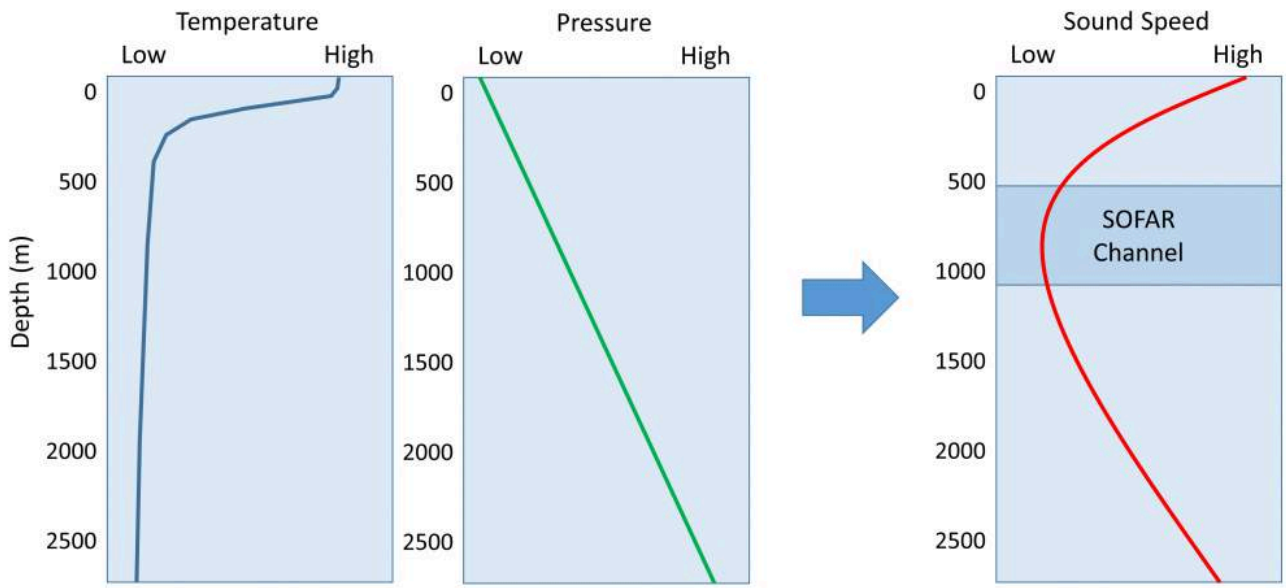


Figure 6.4.1 *Profiles of temperature, pressure, and sound speed with depth. Sound speed is high at the surface due to the high temperatures, and is high at depth because of the high pressure. At moderate depths lies the SOFAR channel, the region of slowest sound speed (PW).*

The SOFAR channel is important because sounds produced in that region can be propagated over very long distances with little attenuation (loss of energy). Sound waves produced in the channel radiate out in all directions. Waves that travel into shallower or deeper water outside of the sound channel are entering a region of faster sound transmission. As we saw with seismic waves, when these sound waves encounter a region of differing transmission speed, the waves tend to be refracted or bent back towards the region of lower speed. As a result, sound waves moving from the SOFAR channel into shallower water will be refracted back towards the channel. As the sound waves go deeper below the channel, they will be refracted upwards, back into the channel and the region of slower speed. In this way, much of the sound does not dissipate out into the water in all directions, but instead is trapped within the channel, and can travel very long distances with little loss of energy (Figure 6.4.2).

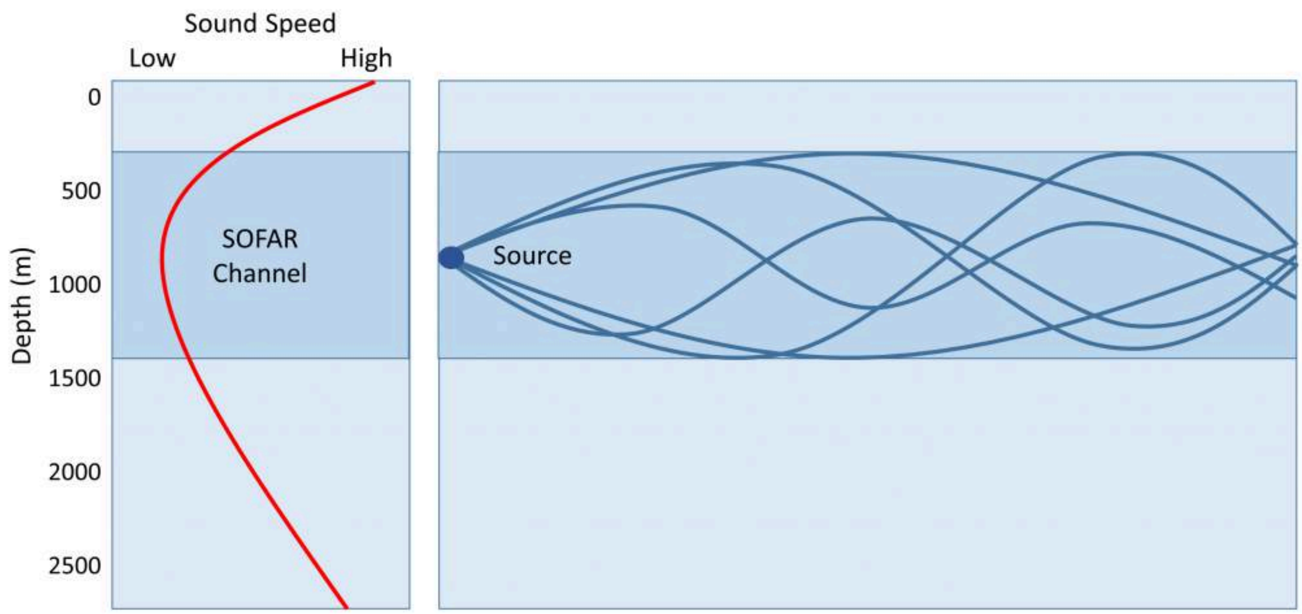


Figure 6.4.2 *Sound propagation in the SOFAR channel. Sound waves emanating from a source will be refracted towards the region of slower speed, "trapping" the sound in the SOFAR channel (PW).*

There are several practical applications of the SOFAR channel. Baleen whales are thought to use the SOFAR channel to communicate with each other over long distances of hundreds to thousands of kilometers. Their vocalizations are very loud and are low frequency calls, which travel farther than high frequency sounds in the oceans. The military has been able to track submarines using the SOFAR channel, and during World War II it was used to locate downed pilots or missing ships and planes. A stranded pilot could drop a small device into the water, and once it sank into the SOFAR channel it would explode, creating a sound that could be heard at multiple listening stations. Using the time of arrival of the sound at the various receivers, the location of the source could be determined through triangulation. In the 1990s it was suggested that the SOFAR channel could be used to monitor global ocean temperatures. A project known as ATOC (Acoustic Thermometry of Ocean Climate) was proposed where loud, low frequency sounds produced near Hawaii and California would travel through the SOFAR channel to receiving stations around the Pacific. By monitoring the time it took for the sounds to reach the receivers, scientists could monitor changes in ocean temperatures on a global scale, as sounds would move faster through a warming ocean.

Since sound travels better through water than air, the energy required to transmit a given sound wave is higher in air than in water. The energy, or intensity (loudness) of a sound is measured on the decibel (dB) scale. It turns out that it takes about 61 times more energy to transmit a sound through air than through water. Because of this energy difference, there is a 61 dB difference between sounds transmitted through air and water, such that a sound intensity of 120 dB in water would be equivalent to an intensity of about 60 dB in air. This should be kept in mind when trying to compare sounds in the ocean with sounds in the air. A sound of 130 dB in air is about equivalent to standing 100 m from a jet engine at takeoff. A sound of 130 dB in water is equivalent to about 70 dB in air, which is the intensity of the sound of a vacuum cleaner. It should also be pointed out that on the dB scale, an increase of 10 dB means the sound is 10 times louder. In other words, 20 dB is 10 x louder than 10 dB, while 30 dB is 100 x louder than 10 dB.

Additional links for more information:

- Discovery of Sound in the Sea website: http://www.dosits.org/

6.5 Light

Radiant energy from the sun is important for several major oceanic processes:

- Climate, winds, and major ocean currents are ultimately dependent on solar radiation reaching the Earth and heating different areas to different degrees.
- Sunlight warms the surface water where much oceanic life lives.
- Solar radiation provides light for photosynthesis, which supports the entire ocean ecosystem.

The energy reaching Earth from the sun is a form of electromagnetic radiation, which is represented by the electromagnetic spectrum (Figure 6.5.1). Electromagnetic waves vary in their frequency and wavelength. High frequency waves have very short wavelengths, and are very high energy forms of radiation, such as gamma rays and x-rays. These rays can easily penetrate the bodies of living organisms and interfere with individual atoms and molecules. At the other end of the spectrum are low energy, long wavelength waves such as radio waves, which do not pose a hazard to living organisms.

Most of the solar energy reaching the Earth is in the range of visible light, with wavelengths between about 400-700 nm. Each color of visible light has a unique wavelength, and together they make up white light. The shortest wavelengths are on the violet and ultraviolet end of the spectrum, while the longest wavelengths are at the red and infrared end. In between, the colors of the visible spectrum comprise the familiar "ROYGBIV"; red, orange, yellow, green, blue, indigo, and violet.

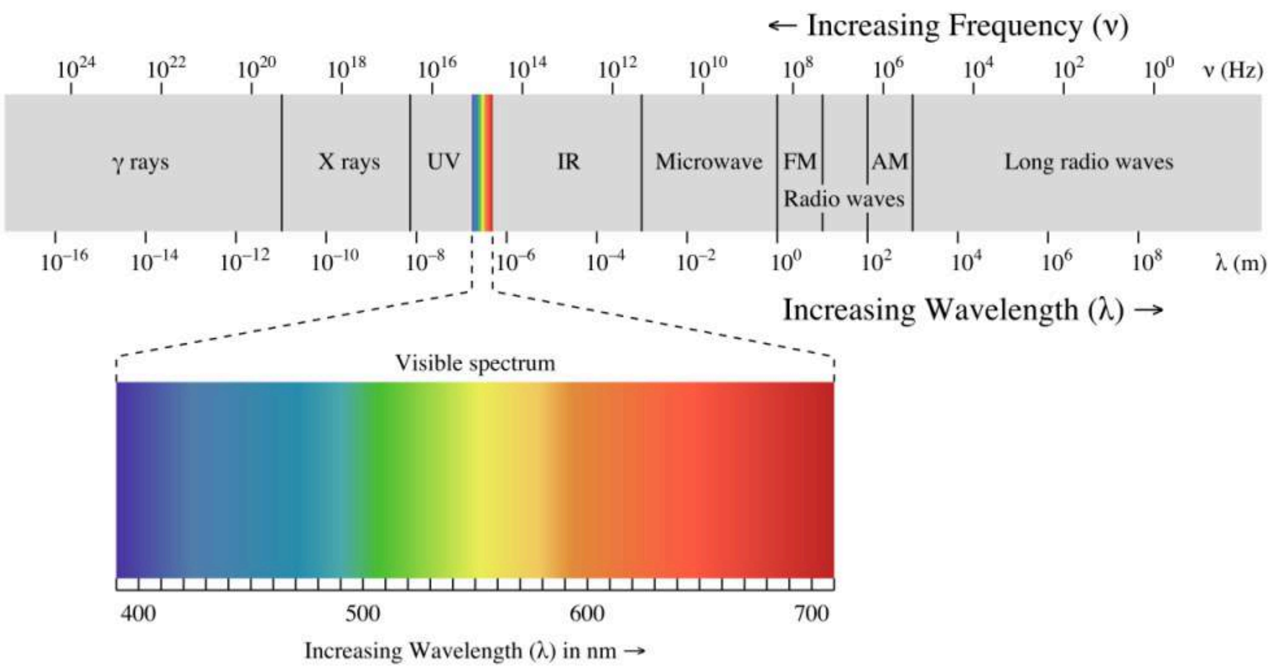


Figure 6.5.1 *The electromagnetic spectrum. Frequency is expressed in Hertz (Hz), or waves per second, while wavelengths are expressed in meters (Phillip Roman, CC BY-SA 3.0, via Wikimedia Commons).*

Water is very effective at absorbing incoming light, so the amount of light penetrating the ocean declines rapidly (is attenuated) with depth (Figure 6.5.2). At 1 m depth, only 45% of the solar energy that falls on the ocean surface remains. At 10 m depth only 16% of the light is still present, and only 1% of the original light is left at 100 m. No light penetrates beyond 1000 m.

In addition to overall attenuation, the oceans absorb the different wavelengths of light at different rates (Figure 6.5.2). The wavelengths at the extreme ends of the visible spectrum are attenuated faster than those wavelengths in the middle. Longer wavelengths are absorbed first; red is absorbed in the upper 10 m, orange by about 40 m, and yellow disappears before 100 m. Shorter wavelengths penetrate further, with blue and green light reaching the deepest depths.

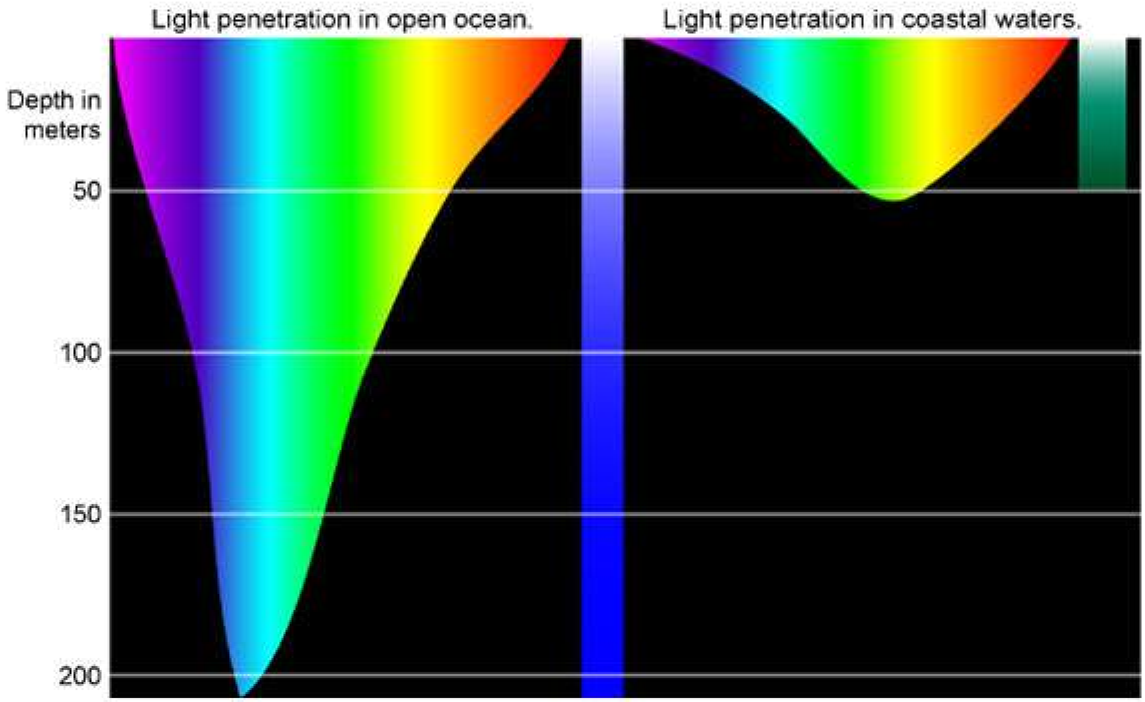


Figure 6.5.2 *Light penetration in open ocean and coastal water, showing the different depths to which each color will penetrate (By NOAA – National Oceanic and Atmospheric Administration [Public domain], via Wikimedia Commons).*

This explains why everything appears blue under water. The colors we perceive depends on the wavelengths of light that are received by our eyes. If an object appears red to us, that is because the object reflects red light but absorbs all of the other colors. So the only color reaching our eyes is red. Under water, blue is the only color of light still available at depth, so that is the only color that can be reflected back to our eyes, and everything has a blue tinge under water. A red object at depth will not appear red to us because there is no red light available to reflect off of the object. Objects in water will only appear as their real colors near the surface where all wavelengths of light are still available, or if the other wavelengths of light are provided artificially, such as by illuminating the object with a dive light.

Water in the open ocean appears clear and blue because it contains much less particulate matter, such as phytoplankton or other suspended particles, and the clearer the water, the deeper the light penetration. Blue light penetrates deeply and is scattered by the water molecules, while all other colors are absorbed; thus the water appears blue. On the other hand, coastal water often appears greenish (Figure 6.5.2). Coastal water contains much more suspended silt and algae and microscopic organisms than the open ocean. Many of these organisms, such as phytoplankton, absorb light in the blue and red range through their photosynthetic pigments, leaving green as the dominant wavelength of reflected light. Therefore the higher the phytoplankton concentration in water, the greener it appears. Small silt particles may also absorb blue light, further shifting the color of water away from blue when there are high concentrations of suspended particles.

The ocean can be divided into depth layers depending on the amount of light penetration, as discussed in section 1.3 (Figure 6.5.3). The upper 200 m is referred to as the **photic** or **euphotic zone**. This represents the region where enough light can penetrate to support photosynthesis, and it corresponds to the epipelagic zone. From 200-1000 m lies the **dysphotic zone**, or the **twilight zone** (corresponding with the mesopelagic zone). There is still some light at these depths, but not enough to support photosynthesis. Below 1000 m is the **aphotic** (or midnight) zone, where no light penetrates. This region includes the majority of the ocean volume, which exists in complete darkness.

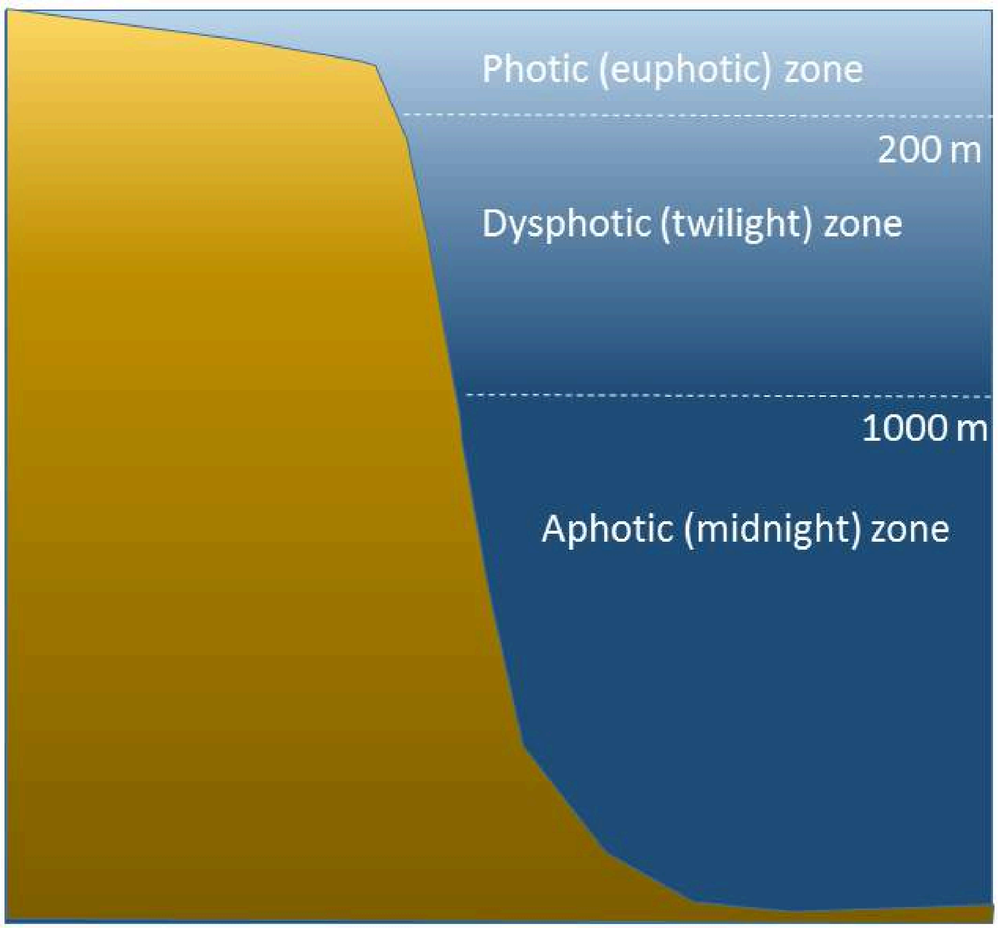


Figure 6.5.3 *The zones of the water column as defined by the amount of light penetration (PW).*

CHAPTER 7: PRIMARY PRODUCTION

Chapter 7: Primary Production

Learning Objectives

After reading this chapter you should:

- be able to define primary production, gross and net production, new and regenerated production
- know some of the major organisms involved in oceanic primary productivity
- know what a red tide is
- know how certain phytoplankton can cause toxic blooms
- know the major requirements for primary production to occur
- understand how light availability and thus primary production changes with depth
- understand the concept of compensation depth in relation to primary production
- know the major nutrients needed for primary production
- know the general depth profiles for nutrients in the ocean, and the factors responsible for those curves
- be able to describe the seasonal patterns in productivity in polar, tropical, and temperate regions, and the factors responsible for those patterns

While this book does not delve too deeply into marine biology topics, one biological process that is crucial to an understanding of oceanography is primary production. Oceanic primary production obviously forms the base of marine food webs, but it also produces about half of the oxygen that we breathe in terrestrial systems. Furthermore, you cannot fully understand the profiles of dissolved gases (section 5.4) or the distribution of biogenous sediments (section 12.6) without considering the roles played by primary production.

This chapter begins by defining primary production, before introducing the organisms responsible for that productivity. We will then examine the requirements for marine productivity, which leads to a discussion of how primary production changes seasonally and along latitudinal gradients.

7.1 Primary Production

Primary production is the creation of new organic matter from inorganic substrates, and it is this organic matter that serves as the base of the food web for most marine consumers. Primary production generally refers to the process of **photosynthesis,** or the utilization of light energy to produce chemical fuels that is undertaken by plants and algae according to the reaction:

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

Here, powered by light energy, carbon dioxide and water combine to produce glucose and oxygen. However, primary production is also carried out by bacteria in the absence of light through **chemosynthesis**. Instead of light providing the energy for the reaction, the energy comes from the oxidation of inorganic materials, such as hydrogen sulfide (see section 4.11 on hydrothermal vents). Here we will concentrate on photosynthesis because it plays a much larger role in total oceanic productivity than chemosynthesis.

The organisms responsible for oceanic primary production include a wide diversity of marine plants and algae. While many people may be more familiar with the larger seagrasses and macroalgae (seaweeds), by far the greatest amount of photosynthesis in the ocean comes from microscopic algae, the **phytoplankton**. The term "plankton" refers to organisms that drift with the currents, and the phytoplankton are the free-floating algae that undergo photosynthesis (contrast this with the **zooplankton**, who are the drifting animals). In section 7.2 we will take a closer look at the organisms responsible for oceanic primary production.

The total amount of organic material created by the producers is called the **gross primary productivity**, or **total production**. However, the primary producers consume a portion of this organic matter themselves through respiration, so the total amount that is left to support the consumers in the ocean is called **net production** (gross productivity – respiration). Gross production can be divided into two components, new production and regenerated production. **New production** is supported by nutrients brought in from outside of the local ecosystem through processes such as upwelling or ocean currents. **Regenerated production** results from the recycling of nutrients within an ecosystem.

Overall, marine productivity is similar to terrestrial production. Marine net production is about 35-50 billion metric tons per year, while terrestrial production reaches 50-70 billion tons per year. However, the biomass responsible for that production in the ocean is about 1-2 billion metric tons, compared to 600-1000 billion metric tons of biomass in terrestrial systems. So the oceans are producing almost as much organic material as terrestrial producers, but are doing it from only a fraction of the amount of producer biomass. One reason for this discrepancy is that the phytoplankton are constantly being consumed, while much of the terrestrial biomass is much longer-lived than the plankton.

7.2 The Producers

Although the phytoplankton are microscopic in size compared to marine plants and macroalgae like seaweeds and seagrasses, they account for by far the greatest amount of photosynthesis in the oceans; about 95% of all marine primary productivity. Most of the production by phytoplankton comes from three groups, the diatoms, dinoflagellates, and coccolithophores.

Diatoms are single-celled algae consisting of cellular material inside a shell, or test, made of silica, a component of glass. Diatoms are relatively large, reaching up to about 1 mm in diameter, and come in a range of shapes, from circular disks to elongated or triangular forms (Figure 7.2.1). In some species of diatoms individual cells link together into multicellular chains. Diatoms are very efficient producers, with up to 55% of the absorbed sunlight energy incorporated into carbohydrate formation; this is one of the highest photosynthetic efficiencies known. Diatoms are most abundant in coastal and cold, nutrient-rich waters. Where diatoms are bountiful, the underlying sediments are rich in their silica shells, creating siliceous sediment and diatomaceous earth (see section 12.3).

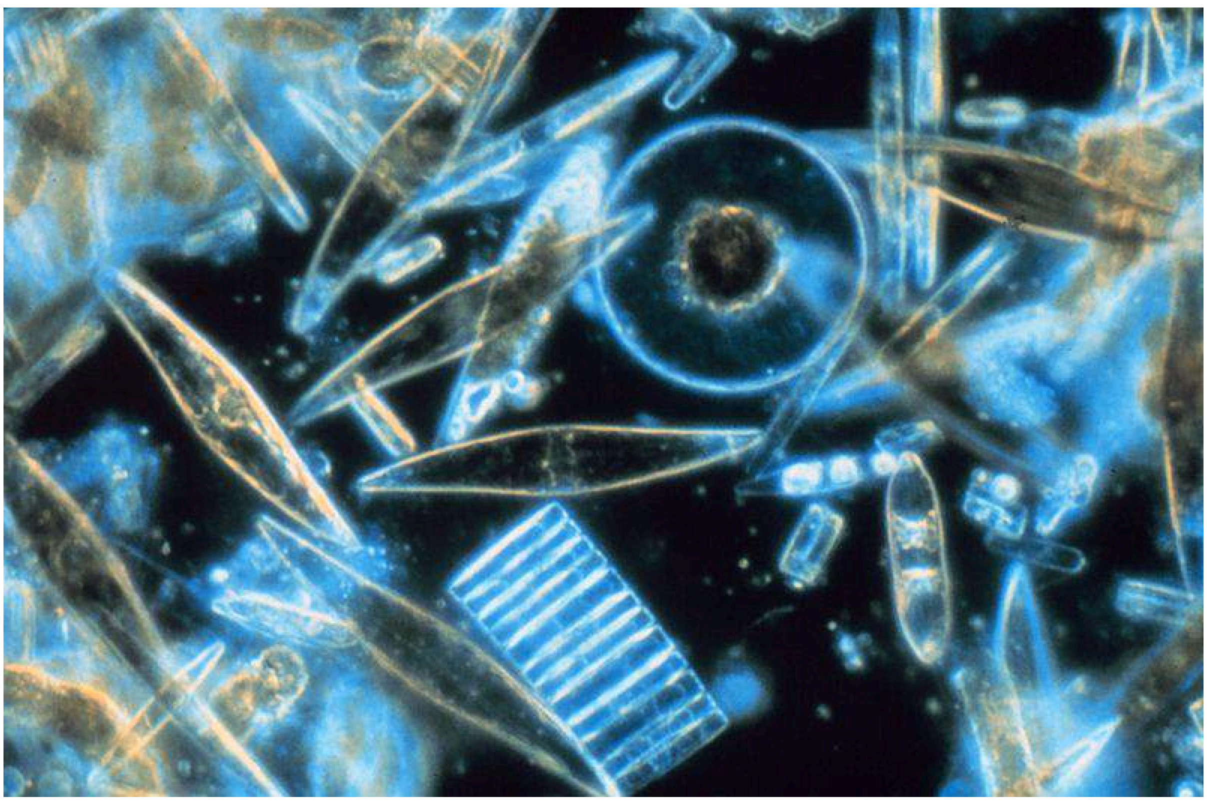


Figure 7.2.1 *A variety of diatom species from the Southern Ocean, seen under a light microscope (Prof. Gordon T. Taylor, Stony Brook University, [Public domain], via Wikimedia Commons).*

Dinoflagellates are another form of single-shelled photosynthetic algae that are generally smaller than diatoms, most with sizes in the 0.015 – 0.04 mm range. Most dinoflagellates have a characteristic pair of flagella (hence the name); small whip-like "tails" that they use for locomotion. Usually there is one flagellum that trails from the body to provide forward movement, and another that encircles the cell to make it spin as it moves. Unlike diatoms, dinoflagellates do not have a mineralized shell. Instead, many are covered by cellulose, which easily decomposes in seawater, so their shells do not really contribute to sediment formation (Figure 7.2.2). While most dinoflagellates undergo photosynthesis, some species will also ingest prey.

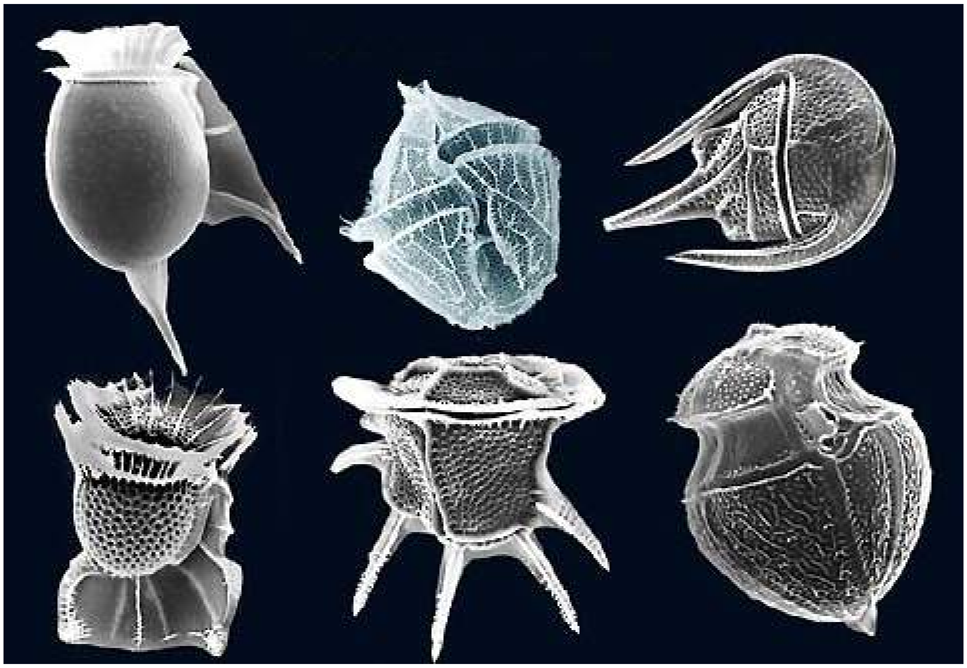


Figure 7.2.2 *Electron microscope image of several dinoflagellate species. In some of the species, a groove can be seen running around the circumference of the organism; this groove is usually home to one of the flagella (fickleandfreckled, https://www.flickr.com/photos/fickleandfreckled/6939384773, CC-BY 2.0).*

A third, much smaller type of phytoplankton includes the **coccolithophores**, which range from about 5-100 micrometers wide. As with diatoms and dinoflagellates, these are single-celled photosynthetic algae that contribute significantly to oceanic primary production, but their cellular material is encased in a very different kind of shell. The test (shell) is made up of a number of interlocking circular plates composed of calcium carbonate that link together to form a sphere (Figure 7.2.3). Coccolithophores are most abundant in warm, open ocean waters, and their sinking tests can lead to calcium carbonate sediments in some parts of the ocean (see section 12.3).

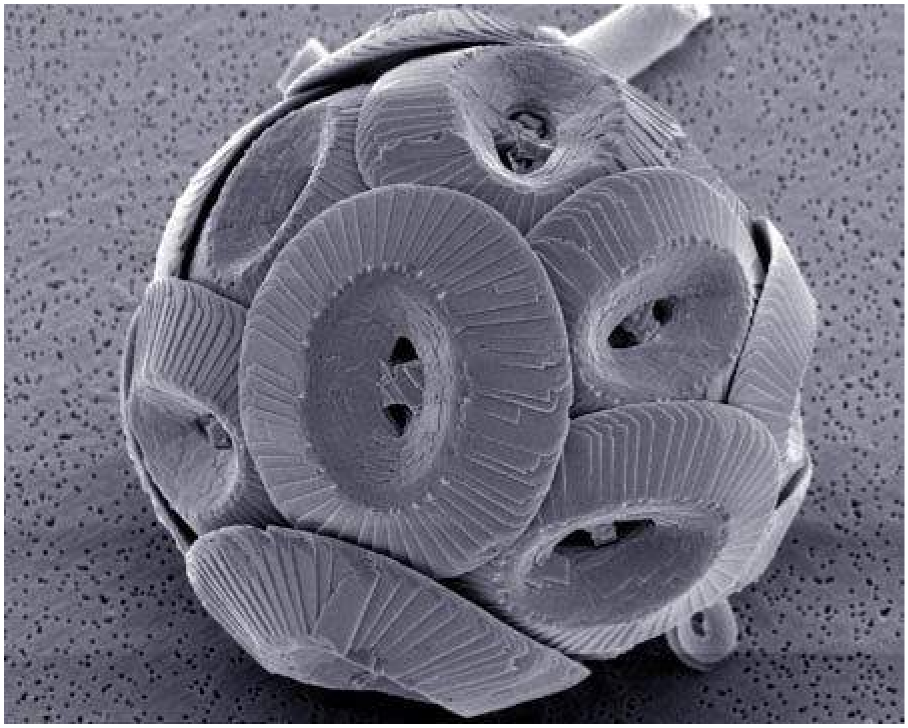


Figure 7.2.3 *Electron microscope image of the North Atlantic coccolithophore Coccolithus pelagicus (Richard Lampitt, Jeremy Young, The Natural History Museum, London (http://planktonnet.awi.de/) [CC BY 2.5], via Wikimedia Commons).*

Recent evidence suggests that another group of organisms, the **bacteria**, or **picoplankton**, may be very important primary producers. Although they are very small, on the order of 0.2-2 micrometers long, they can be found in very high concentrations, and may be responsible for up to 70% of all productivity in some parts of the ocean.

Harmful algal blooms

Primary production provides plentiful food resources for ocean consumers, so a high abundance of phytoplankton is a good thing, right? As in many other cases, too much of a good thing can sometimes be dangerous, and an overabundance of dinoflagellates or diatoms can often create serious concerns. These events are referred to as **harmful algal blooms**, or HABs. HABs can occur for a number of reasons, although a common one is an overabundance of nutrients, which is often due to excessive terrestrial runoff of fertilizers or other nitrogen- and phosphate-containing materials. These conditions lead to an explosion in algal populations that can change the color of the water if the cells are in high enough concentrations. Figure 7.2.4 shows a massive bloom that contained so many dinoflagellate cells that it turned the water reddish-brown, a so-called "**red tide**." (It has been suggested that Biblical references to seas being "turned to blood" may have actually been describing red tide events).

Figure 7.2.4 A red tide caused by dinoflagellates near the Scripps Institution of Oceanography Pier, La Jolla California (Alejandro Díaz, Public Domain via Wikimedia Commons).

These massive algal blooms can have some serious consequences. For one, when all of this algae eventually dies off and sinks, their decomposition uses up the dissolved oxygen in the water, leaving anoxic or hypoxic conditions that can lead to the mass die-off of fish and invertebrates. Dinoflagellates and diatoms are also capable of producing toxins themselves. These phytoplankton get eaten by fish, shellfish and other organisms, and in high abundances the toxins become concentrated in the tissues of the consumers. When humans or other higher-level consumers then eat these organisms, the toxins are concentrated enough to cause sickness or even death. For example, some dinoflagellates produce a toxin that causes **paralytic shellfish poisoning**, which can occur in humans as soon as 30 minutes after eating infected shellfish. This toxin attacks the nervous system, producing symptoms of dizziness, nausea, slurred speech, loss of feeling, and uncoordinated movements, and can ultimately be fatal. Diatoms produce a toxin called domoic acid that causes **amnesic shellfish poisoning**, leading to memory loss, seizures, and potentially death. Domoic acid poisoning also affects marine animals; it is thought to have been responsible for an event in Capitola, California in 1961 where flocks of seabirds acted crazily, even attacking humans. This event inspired the Alfred Hitchcock movie "The Birds."

Additional links for more information:

- For more on harmful algal blooms, visit NOAA's HAB site: http://www.noaa.gov/what-is-harmful-algal-bloom

7.3 Factors Influencing Production

For terrestrial plants, many factors affect productivity, including light, temperature, nutrients, soil, and water. For phytoplankton, soil is obviously not needed, and water availability is not an issue. Temperature is generally more stable in the ocean than on land, so for phytoplankton, productivity comes down to the availability of light and nutrients.

Light

Since light is vital for photosynthesis, phytoplankton and other primary producers are limited to the uppermost layers of the ocean where light is abundant enough to sustain the reaction. As depth increases, light intensity decreases until there reaches a depth where photosynthesis can no longer occur (Figure 7.3.1). The region through which sufficient light for photosynthesis can penetrate is called the photic or euphotic zone, which extends down to about 200 m (section 6.5).

In addition to undergoing photosynthesis, phytoplankton also respire, consuming some of the organic compounds they produce. Rates of respiration are not light dependent, and respiration occurs at all depths and light levels. Therefore, as depth increases the rate of photosynthesis declines as light is diminished, until a point is reached where the rate of photosynthesis equals the respiration rate (Figure 7.3.1). This depth is the **compensation depth**, and it marks the lower level of the photic zone, and represents the depth where net primary production ends. Below this depth, there is net respiration.

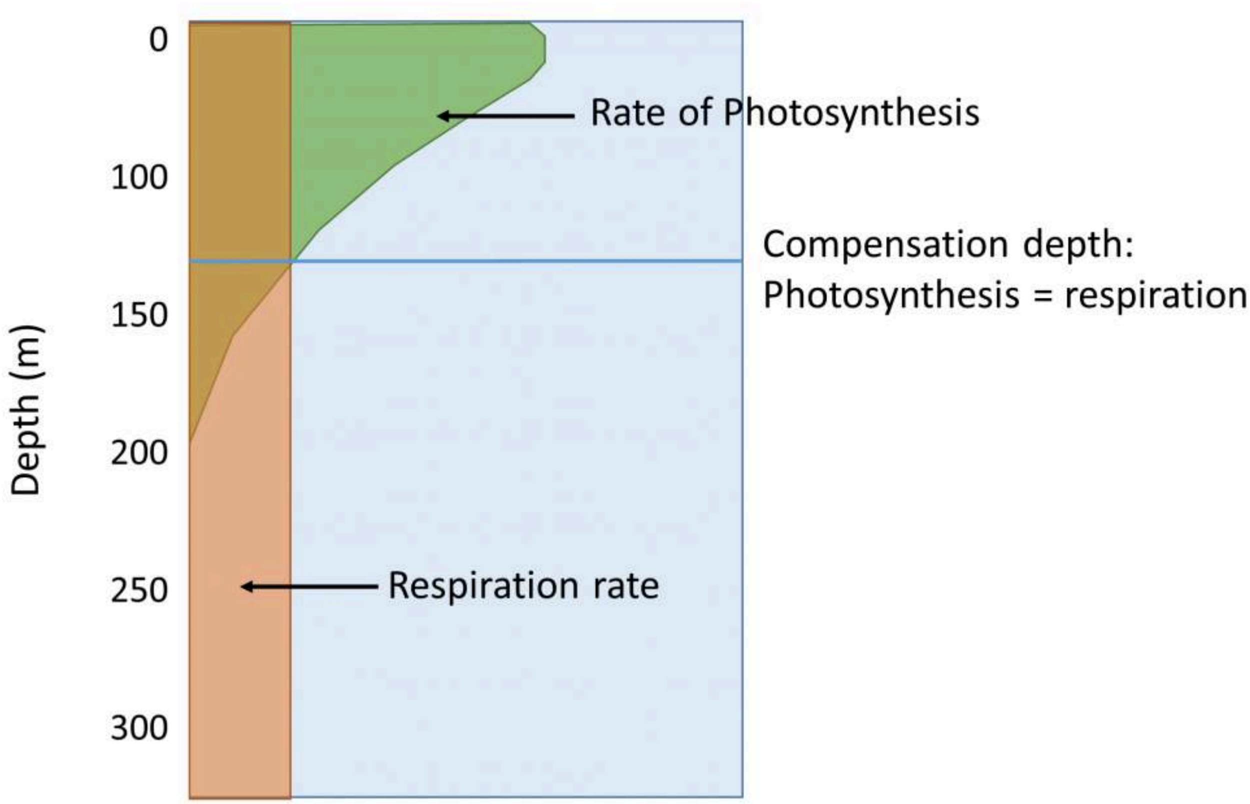


Figure 7.3.1 *As depth increases the rate of photosynthesis declines as light becomes limited. The rate of respiration remains consistent at all depths. The depth where photosynthesis equals respiration is the compensation depth (PW).*

Nutrients

Nutrients are required by all of the marine primary producers. The major nutrients required by phytoplankton are nitrogen and phosphorus, in the forms of nitrate NO_3^-, nitrite NO_2^{2-}, ammonium NH_4^+, and phosphate PO_4^{3-}. Many phytoplankton, particularly the diatoms, also require silica, SiO_2, for shell formation. All of these nutrients occur in very small amounts in seawater, so they are often the limiting factors for phytoplankton growth in most situations, particularly the nitrogen compounds. For example, agricultural soil contains 0.5% nitrogen in the upper meter of soil, while surface ocean water contains about 0.00005% nitrogen, 1/10,000 the amount in soil.

As we saw in section 5.6, nutrients are not distributed evenly throughout the water column (Figure 7.3.2). Near the surface nutrients are quickly utilized by phytoplankton as they become available, so surface waters are nutrient-poor. But as the phytoplankton are consumed or die they are recycled into particles of organic matter, such as fecal pellets or carcasses, that sink into deeper water. Once in deep water, decomposition of these materials releases the nutrients back into the water column. Because there are no producers to utilize them at depth, nutrients are more abundant in deeper water.

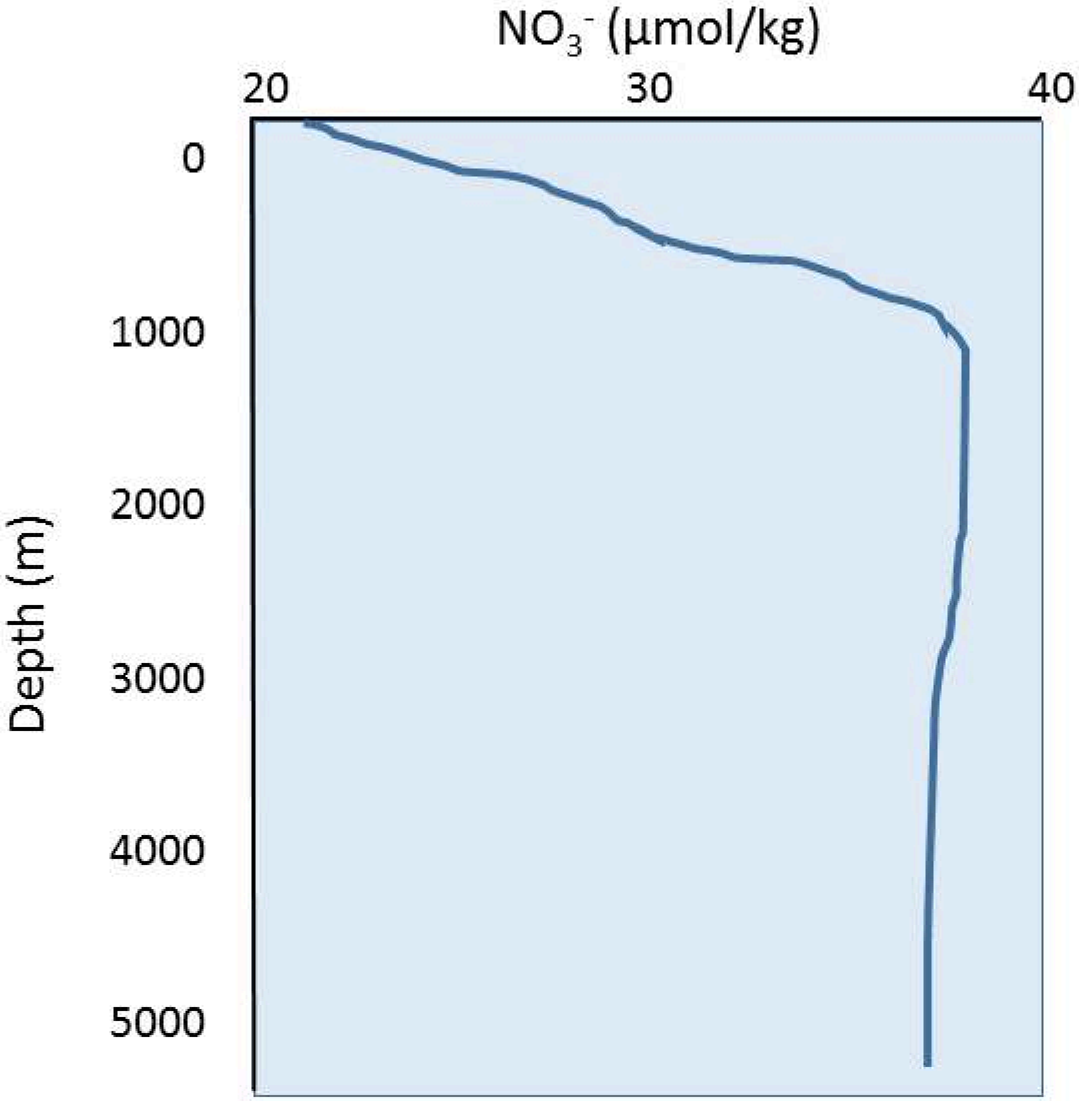


Figure 7.3.2 *Representative nutrient profile in the open ocean. While this profile shows nitrate abundance, the profile is similar for other nutrients such as phosphate and silica (PW).*

These deep water nutrients are out of reach of the phytoplankton at the surface. The thermocline and density stratification of the water column generally prevents the nutrient-rich deep water from mixing with the surface water. However, under certain conditions this nutrient-rich deep water may be brought to the surface through the process of upwelling (see section 9.5). Where upwelling occurs there is usually high productivity as the phytoplankton can take advantage of the input of nutrients.

7.4 Patterns of Primary Production

Primary productivity varies both geographically and seasonally. Geographically, phytoplankton abundance generally decreases as you move from coastal to oceanic waters (Figure 7.4.1). Coastal waters are more productive than the central ocean for two main reasons. First, runoff from land often contains a high abundance of nutrients which get deposited in coastal waters and stimulate production. Second, the shallower bottom along the continental shelf can trap nutrients and prevent them from sinking to greater depths. It is easier for these nutrients to be brought back to the surface when they remain trapped in the shallows. Conversely, the central ocean generally has very low primary production, as these areas are far removed from any terrestrial sources of nutrients, and the great depth prevents the deep nutrients from returning to the surface.

Global averages for ocean surface primary production are about 75-150 g C/m^2/yr. Some highly productive areas include the California coast (200-300 g C/m^2/year), the Southern Ocean (200-400 g C/m^2/year), and the coast of Peru (200-400 g C/m^2/year), all regions with significant upwelling. The central ocean, by contrast, produces less than 50 g C/m^2/year.

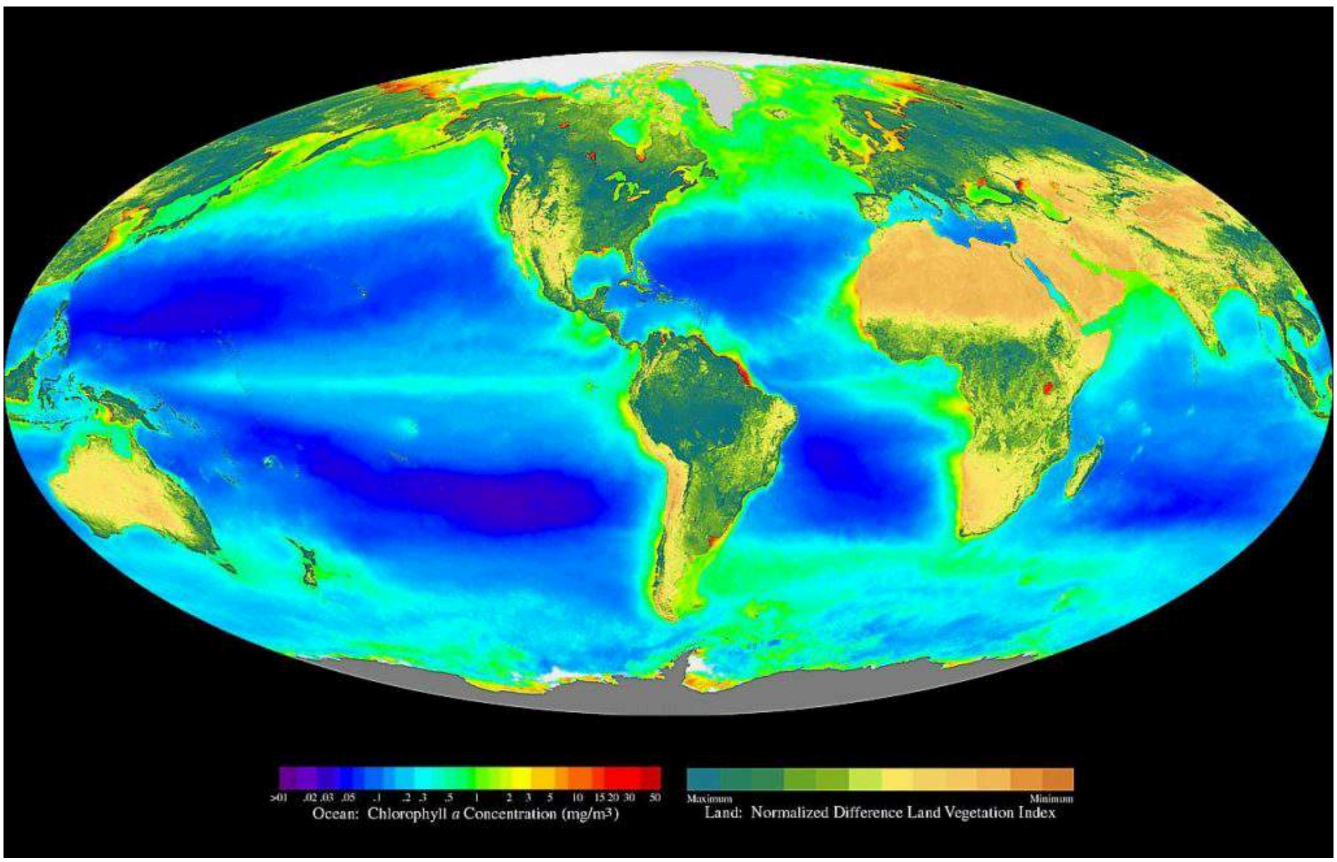


Figure 7.4.1 *Global surface ocean primary productivity, as measured by chlorophyll concentration (Provided by the SeaWiFS Project, Goddard Space Flight Center and ORBIMAGE [Public domain], via Wikimedia Commons).*

Regional and seasonal changes in primary production are due to a combination of the availability of light and the amount of nutrients provided by water mixing above the thermocline. In tropical regions sunlight is plentiful throughout the year, so light is not a limiting factor. The surface water is always warm and there is always a pronounced thermocline, leading to highly stratified water that prevents the nutrient-rich bottom water from reaching the surface (section 6.2). Thus productivity in tropical water is always nutrient-limited, and productivity is low throughout the year (Figure 7.4.2). Because tropical water is nutrient-poor with little phytoplankton production, the water is very clear, as is the case with water in the central ocean.

At the poles, the water is uniformly cold at all depths, so there is not much of a thermocline and little stratification, allowing mixing to occur year-round (section 6.2). This mixing distributes nutrients throughout the water column, so that for much of the year productivity will not be nutrient-limited. However, the polar regions may experience several months with little or no light during the winter, and the fluctuation in light levels leads to variation is seasonal productivity. In the winter months, mixing is occurring and nutrients are abundant, but there is no light, so there is no productivity. By late spring the sunlight returns, and combined with the abundance of nutrients, a spring/summer bloom of phytoplankton occurs (Figure 7.4.2). By late summer, the nutrients have been depleted and zooplankton have been grazing on the phytoplankton, so the bloom begins to decline. In the autumn, light levels decline and prevent further production throughout the winter. But during the winter the mixing is distributing nutrients throughout the water, ready for the sun to return and stimulate a bloom in the following summer.

In temperate regions there is much more seasonal variation in the depth and intensity of the thermocline (see section 6.2). The thermocline is shallower and stronger in the summer, and is deeper and weaker in the winter, so mixing of deep and surface water is more pronounced in the winter months. As with the poles, this winter mixing creates nutrient-rich water during the winter, but the lack of light limits winter productivity. When light levels increase in the spring, the combination of abundant light and nutrients creates a spring bloom of productivity. By late summer there is still plenty of light, but the nutrients have been depleted by the phytoplankton bloom, and the summer thermocline has prevented further mixing, so productivity declines. In the autumn, cooler temperatures weaken the thermocline, and increasing storms cause a deeper mixed layer to form, bringing nutrients back to the surface. At the same time, there is still sufficient light available that a smaller autumn bloom occurs (Figure 7.4.2). But this bloom is short-lived, as light declines throughout the autumn and into the winter. Again, there is little production during the winter due to light limitations, but the winter storms and deep thermocline recharge the water with nutrients for the next spring bloom.

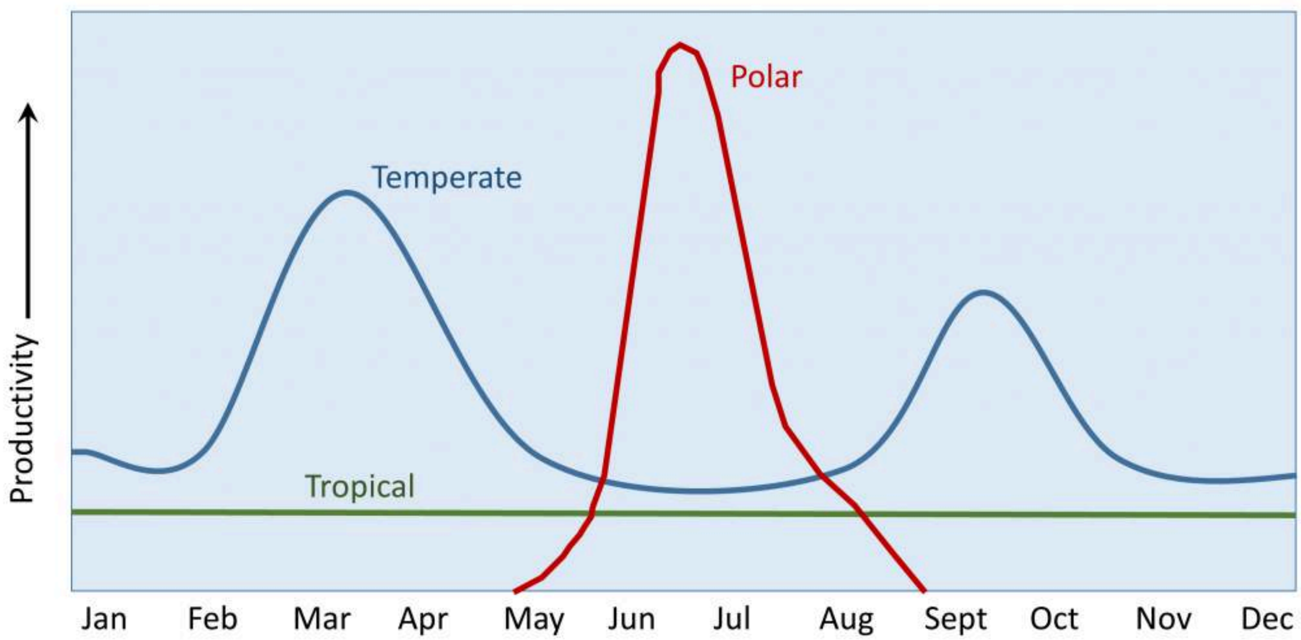


Figure 7.4.2 *Seasonal patterns of productivity in the Northern Hemisphere. Tropical regions are always nutrient-limited and show low productivity. Polar regions are light limited in the winter and only display production during the late spring and summer months when light is available. Northern temperate regions have a spring bloom, and a smaller autumn bloom (PW).*

CHAPTER 8: OCEANS AND CLIMATE

Chapter 8: Oceans and Climate

Learning Objectives

After reading this chapter you should:

- understand the concept of the heat budget of the earth (you do not need to remember all of the percentages, just general trends)
- know the difference and relative importance of radiation, conduction and phase change in exchanging heat with the atmosphere
- understand the mechanisms and causes of the greenhouse effect and global warming
- understand how evaporation and condensation transport heat in the atmosphere and oceans
- understand how the curvature of the Earth results in differential heating of the surface
- understand the role played by albedo in radiation reaching the Earth
- understand how and why atmospheric convection cells form
- understand the reasons for and the results of the Coriolis Effect
- be able to derive the major wind patterns on Earth
- understand the relationship between climatic zones and convection cells – e.g. why are rainforests along the equator?
- understand the effects of altitude on air masses and climate
- understand how seasonal and daily cycles effect things like land and sea breezes
- understand how hurricanes form
- understand the potential impacts of climate change

The ultimate source of energy driving the motion of the atmosphere and the ocean is radiant energy from the sun, which falls on different parts of the Earth in differing amounts. The oceans are the recipient of most of this solar energy, and they are therefore a major factor in regulating Earth's climate.

Remember that compared to land temperatures, ocean temperatures do not undergo large swings from day to night or seasonally. This is due to a number of factors;

- Water has a very high heat capacity, so it can absorb a large amount of heat without much of an increase in temperature. Water can also release large amounts of heat back to the atmosphere without its temperature declining as much as land temperatures would.
- On land, the solar energy only hits the surface, which can heat up dramatically, but the heat does not penetrate very far below the surface. In water, light penetrates for a few hundred meters, so the heat is distributed through a greater area, and water does not heat up as quickly as land.
- Mixing of water in the top few hundred meters also distributes heat. Mixing does not happen on land.

Because of water's ability to regulate heat exchange and climate, areas near the oceans usually have a much milder climate than regions in the center of the continents. Furthermore, areas in the Southern Hemisphere have a much more moderate climate than regions of similar latitude in the Northern Hemisphere, because a larger proportion of the Southern Hemisphere is covered by oceans.

In this chapter we will examine the ways that the oceans and atmosphere interact with solar radiation to influence wind and atmospheric circulation, local weather phenomena, and global climatic zones.

8.1 Earth's Heat Budget

The balance of incoming and outgoing heat on Earth is referred to as its **heat budget**. As with any budget, to maintain constant conditions the budget must be balanced so that the incoming heat equals the outgoing heat. The heat budget of Earth appears below (Figure 8.1.1).

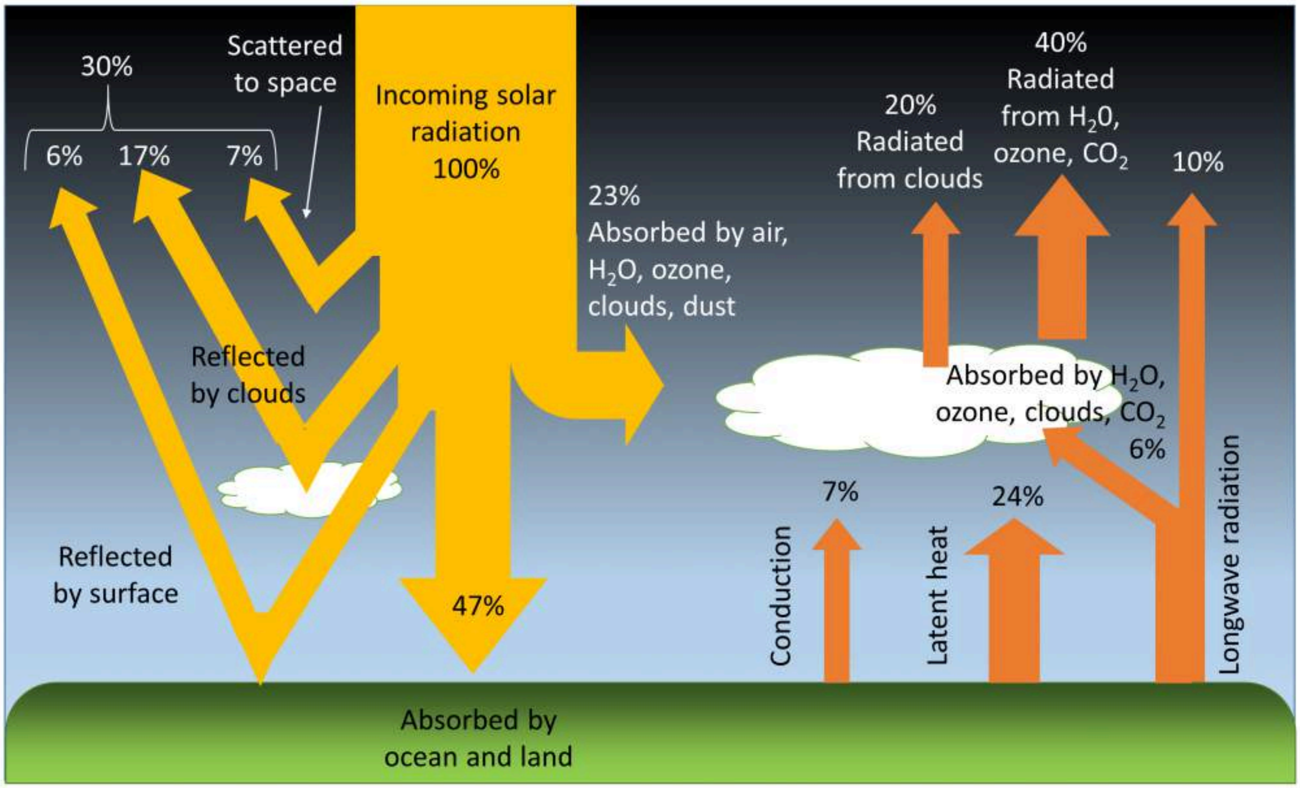


Figure 8.1.1 *Earth's heat budget. Of all of the solar radiation reaching Earth, 30% is reflected back to space and 70% is absorbed by the Earth (47%) and atmosphere (23%). The heat absorbed by the land and oceans is exchanged with the atmosphere through conduction, radiation, and latent heat (phase change). The heat absorbed by the atmosphere is eventually radiated back into space (PW).*

Of all of the solar energy reaching the Earth, about 30% is reflected back into space from the atmosphere, clouds, and surface of the Earth. Another 23% of the energy is absorbed by the water vapor, clouds, and dust in the atmosphere, where it is converted into heat. Just under half (47%) of the incoming solar radiation is absorbed by the land and ocean, and this energy heats up the Earth's surface. The energy absorbed by the Earth returns to the atmosphere through three processes; conduction, radiation, and latent heat (phase change)(Figure 8.1.1).

Conduction is the transfer of heat through direct contact between the surface and the atmosphere. Air is a relatively poor thermal conductor (which means it is a good insulator), so conduction represents only a small part of the energy transfer between the Earth and the atmosphere; equal to about 7% of the incoming solar energy.

All bodies with a temperature above absolute zero (-273^{o} C) radiate heat in the form of longwave, infrared **radiation** (see the electromagnetic spectrum in section 6.5). The warmed Earth is no exception, and about 16%

of the original solar energy is radiated from the Earth to the atmosphere (Figure 8.1.1). Some of this radiated energy will dissipate into space, but a significant amount of heat will be absorbed by the atmosphere. This is the basis for the **greenhouse effect** (Figure 8.12). In the greenhouse effect, shortwave solar radiation passes through the atmosphere and reaches the Earth's surface where it gets absorbed. When the radiation is re-emitted by the Earth, it is now in the form of long wavelength, infrared radiation, which does not easily pass through the atmosphere. Instead, this infrared radiation is absorbed by the atmosphere, particularly by the greenhouse gases such as CO_2, methane, and water vapor. As a result, the atmosphere heats up. Without the greenhouse effect, the average temperature on Earth would be about -18^{o} C, which is too cold for liquid water, and therefore life as we know it could not exist!

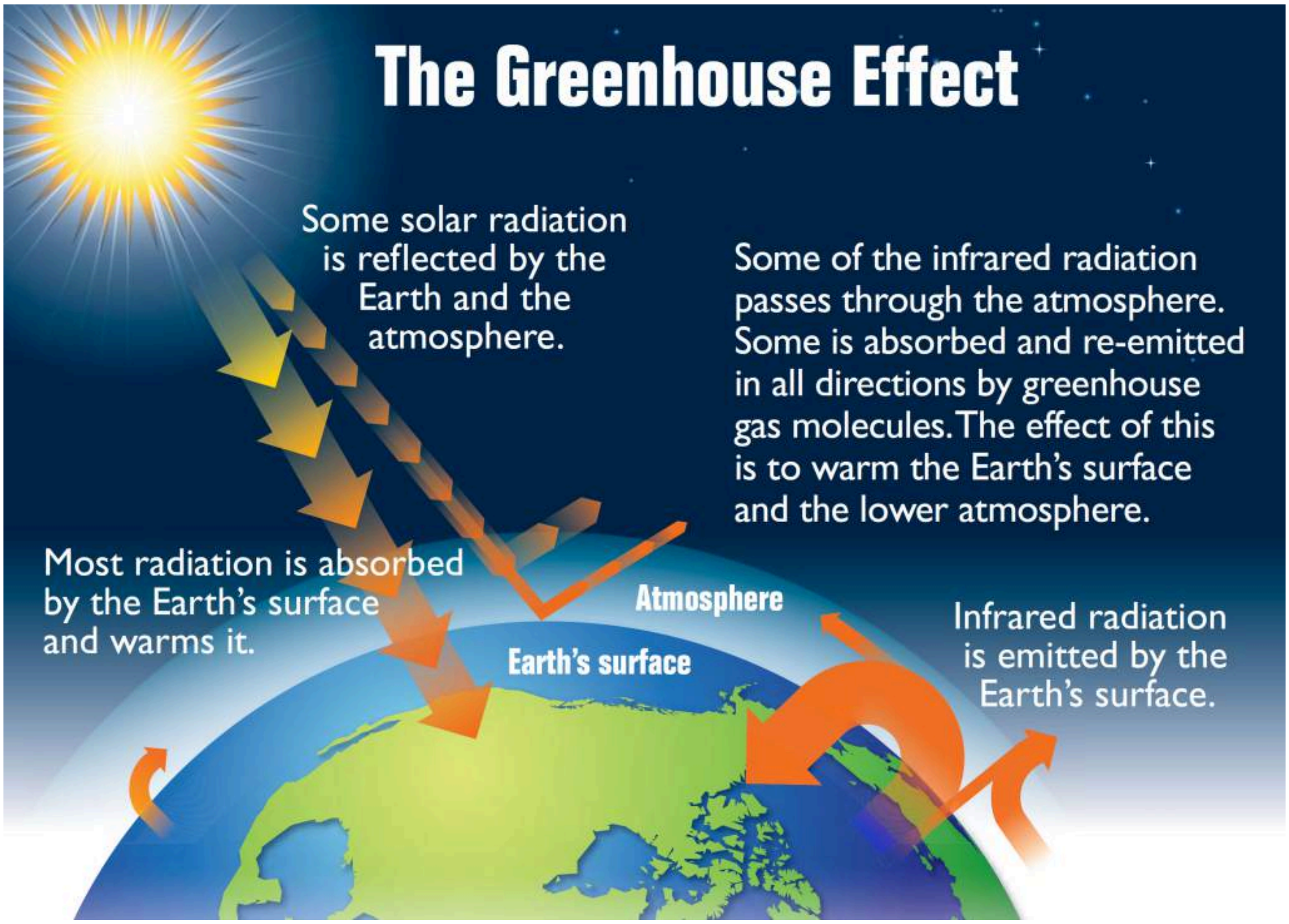


Figure 8.1.2 *An explanation of the greenhouse effect (By US EPA [Public domain], via Wikimedia Commons).*

There is a great deal of concern about the greenhouse effect across the globe; not because of the presence of the effect itself, but because the effect is intensifying, causing climate change or global warming. Since the Industrial Revolution the atmospheric concentrations of the major greenhouse gases, particularly CO_2 and methane, have increased dramatically due to industrialization, the burning of fossil fuels, and deforestation. At the same time, there has been rapid warming of the global climate; CO_2 concentrations have increased more than 25% and global temperature has risen by 0.5^{o} C over the past century. Unless production of these greenhouse gases is curbed, this rapid warming trend may continue, with potentially dire consequences. See section 8.5 for detailed information on the causes and effects of climate change.

The largest pathway for heat exchange between the land or oceans and the atmosphere is **latent heat** transferred through **phase changes**; heat released or absorbed when water moves between solid, liquid, and vapor forms (see section 5.1). Heat must be added to liquid water to make it evaporate, and when water vapor is formed, that heat is removed from the ocean and transferred to the atmosphere along with the water vapor. When water vapor condenses into rain, that heat is then returned to the oceans. The same process happens with the formation and melting of ice. Heat is absorbed by ice when it melts, and heat is released when ice forms, and these phase changes transfer heat between the oceans and the atmosphere.

To complete the heat budget, the heat that is absorbed by the atmosphere either directly from solar radiation or as a result of conduction, radiation and latent heat, is eventually radiated back into space (Figure 8.1.1).

Differential Heating of Earth's Surface

If the Earth was a flat surface facing the sun, every part of that surface would receive the same amount of incoming solar radiation. However, because the Earth is a sphere, sunlight is not equally distributed over the Earth's surface, so different regions of Earth will be heated to different degrees. This differential heating of Earth's surface occurs for a number of reasons. First, because of the curvature of Earth, sunlight only falls perpendicularly to the surface at the center of the sphere (equatorial regions). At any other point on Earth, the angle between the surface and the incoming solar radiation is less than 90^{o}. Because of this, the same amount of incoming solar radiation will be concentrated in a smaller area at the equator, but will be spread over a much larger area at the poles (Figure 8.1.3). Thus the tropics receive more intense sunlight and a greater amount of heating per unit of area than the polar regions.

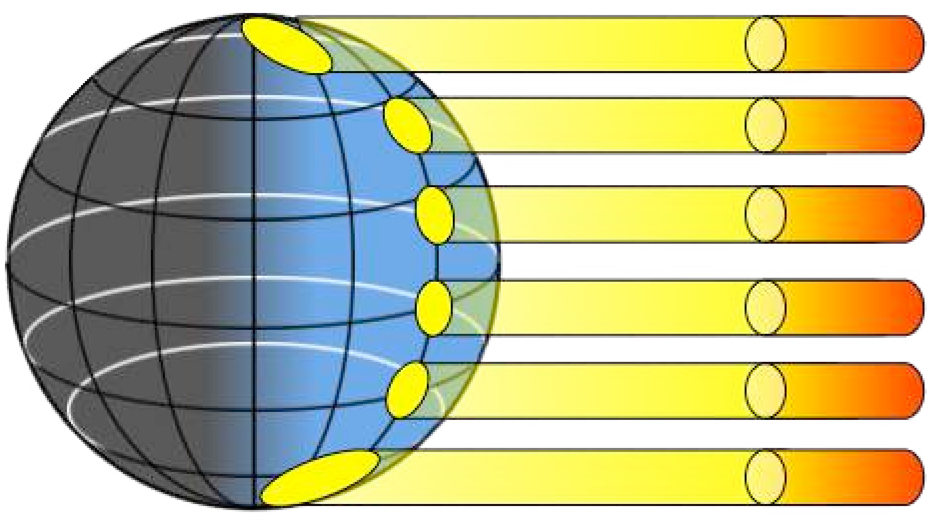


Figure 8.1.3 *Because of the curvature of the Earth, the same amount of sunlight will be spread out over a larger area at the poles compared to the equator. The equator therefore receives more intense sunlight, and a greater amount of heat per unit of area (By Thebiologyprimer (Own work) [CC0], via Wikimedia Commons).*

The angle at which sunlight strikes the Earth contributes to differential heating of the surface in an additional way. At the poles, because of the angle at which the solar energy strikes the surface, more of the light will glance off of the surface and the atmosphere and be reflected back into space. At the equator, the direct angle with which light reaches the surface results in more of the energy being absorbed rather than reflected. Finally, the poles reflect more solar energy than other parts of the Earth because the poles have a higher **albedo**. The albedo refers to reflectivity of a surface. Lighter surfaces are more reflective than darker surfaces (which absorb more energy), and therefore have a higher albedo. At the poles, the ice, snow and cloud cover create a much higher albedo, and the poles reflect more and absorb less solar energy than the lower latitudes. Through all of these mechanisms, the poles absorb much less solar radiation than equatorial regions, which is why the poles are cold and the tropics are very warm.

But there is an interesting twist to this global distribution of heat. The tropical regions actually receive more radiant heat than they emit, and the poles emit more heat than they receive (Figure 8.1.4). We should therefore expect that the tropics will be getting continually warmer, while the poles become increasingly cold. Yet this is not the case; so what is happening? Rather than the heat remaining isolated near the equator, about 20% of the heat from the tropics is transported to the poles before it is emitted. This large scale transport of energy moderates the climates at both extremes. The mechanisms for this heat transfer are ocean and atmospheric circulation, the topic of the next section.

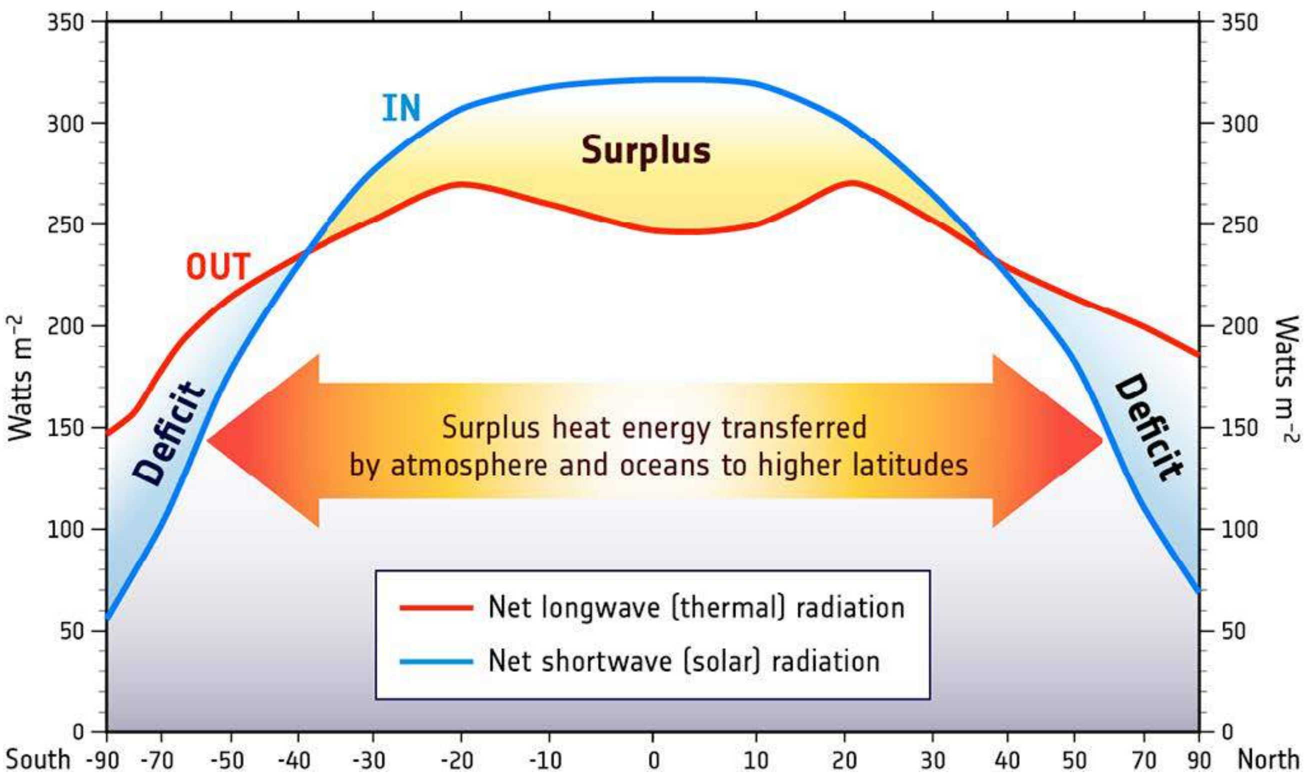


Figure 8.1.4 *The balance between heat gain and heat loss as a function of latitude. Excess heat received near the equator is transferred towards the poles (National Oceanography Centre (NOC). Creative Commons 3.0 unported license).*

The idea of differential heating of the Earth's surface is fundamental to understanding a wide range of oceanographic and atmospheric processes. This differential heating leads to atmospheric convection, which creates winds, which blow over the water and create waves and surface currents, and these currents influence nutrient distribution, which promotes primary production, which then supports the rest of the ocean ecosystem. So there's a lot riding on the simple fact that more light reaches the tropics than the poles!

8.2 Winds and the Coriolis Effect

Differential heating of the Earth's surface results in equatorial regions receiving more heat than the poles (section 8.1). As air is warmed at the equator it becomes less dense and rises, while at the poles the cold air is denser and sinks. If the Earth was non-rotating, the warm air rising at the equator would reach the upper atmosphere and begin moving horizontally towards the poles. As the air reached the poles it would cool and sink, and would move over the surface of Earth back towards the equator. This would result in one large atmospheric convection cell in each hemisphere (Figure 8.2.1), with air rising at the equator and sinking at the poles, and the movement of air over the Earth's surface creating the winds. On this non-rotating Earth, the prevailing winds would thus blow from the poles towards the equator in both hemispheres (Figure 8.2.1).

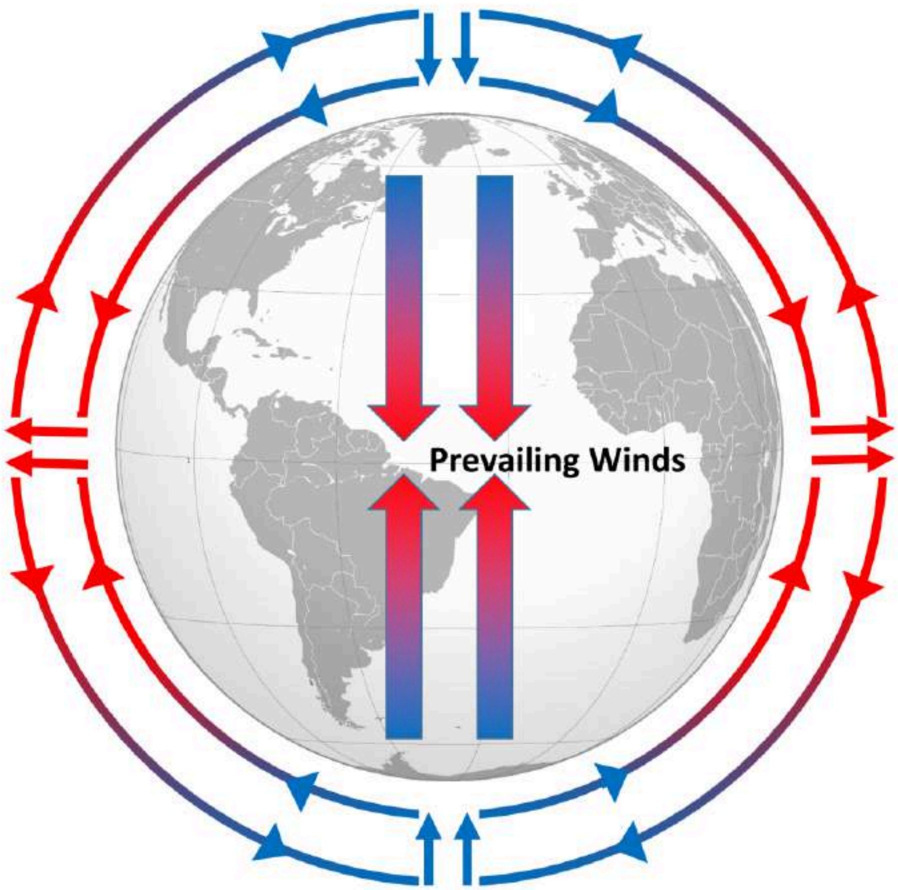


Figure 8.2.1 *Hypothetical atmospheric convection cells on a non-rotating Earth. Air rises at the equator and sinks at the poles, creating a single convection cell in each hemisphere. The prevailing winds moving over the Earth's surface blow from the poles towards the equator in both hemispheres (Modified by PW from globe image by Location_of_Cape_Verde_in_the_globe.svg: Eddo derivative work: Luan fala! [CC BY-SA 3.0], via Wikimedia Commons).*

The non-rotating situation in Figure 8.2.1 is of course only hypothetical, and in reality the Earth's rotation makes this atmospheric circulation a bit more complex. The paths of the winds on a rotating Earth are deflected by the **Coriolis Effect**. The Coriolis Effect is a result of the fact that different latitudes on Earth rotate at different speeds. This is because every point on Earth must make a complete rotation in 24 hours, but some points must travel farther, and therefore faster, to complete the rotation in the same amount of time. In 24 hours a point on the equator must complete a rotation distance equal to the circumference of the Earth, which is about 40,000 km. A point right on the poles covers no distance in that time; it just turns in a circle. So the speed of rotation at the equator is about 1600 km/hr, while at the poles the speed is 0 km/hr. Latitudes in between rotate at intermediate speeds; approximately 1400 km/hr at 30^{o} and 800 km/hr at 60^{o}. As objects move over the surface of the Earth they encounter regions of varying speed, which causes their path to be deflected by the Coriolis Effect.

To explain the Coriolis Effect, imagine a cannon positioned at the equator and facing north. Even though the cannon appears stationary to someone on Earth, it is in fact moving east at about 1600 km/hr due to Earth's rotation. When the cannon fires the projectile travels north towards its target; but it also continues to move to the east at 1600 km/hr, the speed it had while it was still in the cannon. As the shell moves over higher latitudes, its momentum carries it eastward faster than the speed at which the ground beneath it is rotating. For example, by 30^{o} latitude the shell is moving east at 1600 km/hr while the ground is moving east at only 1400 km/hr. Therefore, the shell gets "ahead" of its target, and will land to the east of its intended destination. From the point of view of the cannon, the path of the projectile appears to have been deflected to the right (red arrow, Figure 8.2.2). Similarly, a cannon located at 60^{o} and facing the equator will be moving east at 800 km/hr. When its shell is fired towards the equator, the shell will be moving east at 800 km/hr, but as it approaches the equator it will be moving over land that is traveling east *faster* than the projectile. So the projectile gets "behind" its target, and will land to the west of its destination. But from the point of view of the cannon facing the equator, the path of the shell still appears to have been deflected to the right (green arrow, Figure 8.2.2). Therefore, in the Northern Hemisphere, the apparent Coriolis deflection will *always* be to the **right**.

In the Southern Hemisphere the situation is reversed (Figure 8.2.2). Objects moving towards the equator from the south pole are moving from low speed to high speed, so are left behind and their path is deflected to the left. Movement from the equator towards the south pole also leads to deflection to the left. In the Southern Hemisphere, the Coriolis deflection is always to the **left** from the point of origin.

The magnitude of the Coriolis deflection is related to the difference in rotation speed between the start and end points. Between the poles and 60^{o} latitude, the difference in rotation speed is 800 km/hr. Between the equator and 30^{o} latitude, the difference is only 200 km/hr (Figure 8.2.2). Therefore the strength of the Coriolis Effect is stronger near the poles, and weaker at the equator.

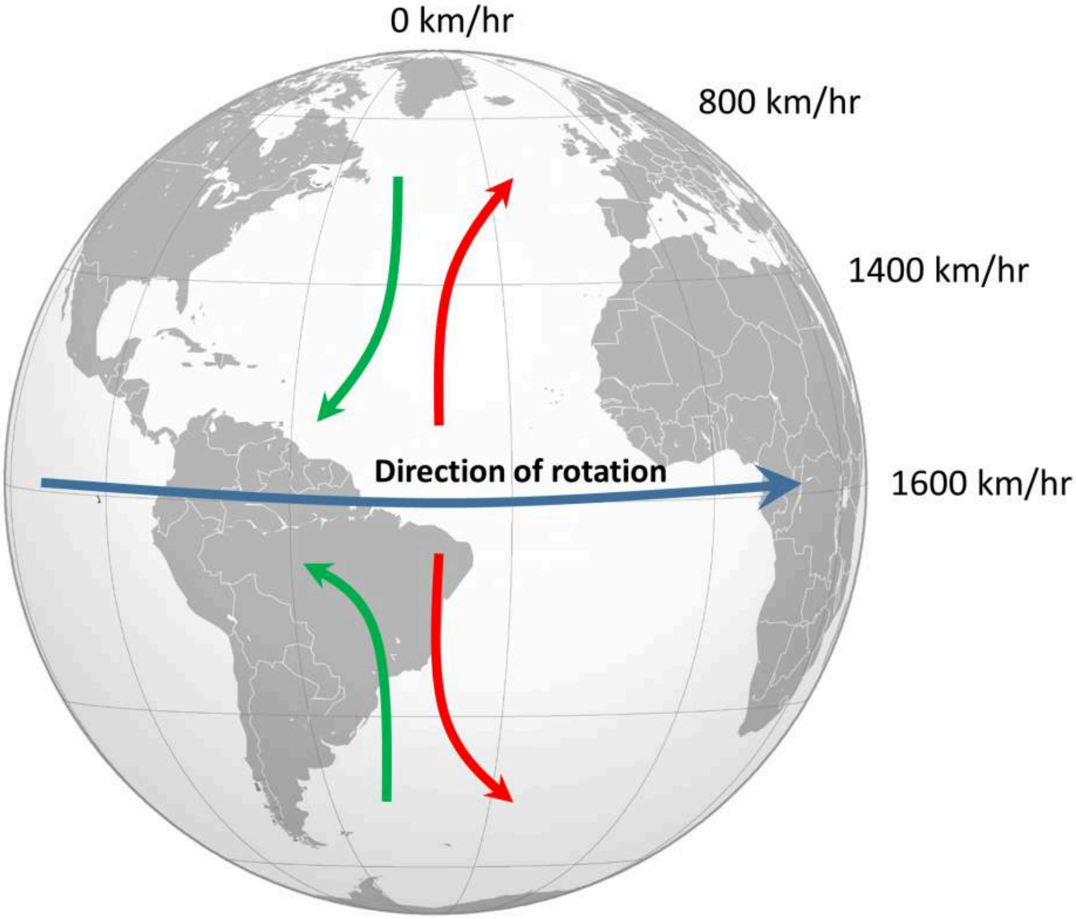


Figure 8.2.2 *The Coriolis Effect. Objects moving from the equator towards the poles (red arrows) move into a region of slower rotational speed and their paths are deflected "ahead" of their point of origin. Movement from high latitudes to low latitudes (green arrows) goes from a region of low speed to a region of higher rotation speed, and there is deflection "behind" their point of origin. In the Northern Hemisphere this deflection is always to the right from the point of origin, and in the Southern Hemisphere the deflection is always to the left (Modified by PW from globe image by Location_of_Cape_Verde_in_the_globe.svg: Eddo derivative work: Luan fala! [CC BY-SA 3.0], via Wikimedia Commons).*

Because of the rotation of the Earth and the Coriolis Effect, rather than a single atmospheric convection cell in each hemisphere, there are three major cells per hemisphere. Warm air rising at the equator cools as it moves through the upper atmosphere, and it descends at around 30^{o} latitude. The convection cells created by rising air at the equator and sinking air at 30^{o} are referred to as **Hadley Cells**, of which there is one in each hemisphere. The cold air that descends at the poles moves over the Earth's surface towards the equator, and by about 60^{o} latitude it begins to rise, creating a **Polar Cell** between 60^{o} and 90^{o}. Between 30^{o} and 60^{o} lie the **Ferrel Cells**, composed of sinking air at 30^{o} and rising air at 60^{o} (Figure 8.2.3). With three convection cells in each hemisphere that rotate in alternate directions, the surface winds no longer always blow from the poles towards the equator as in the non-rotating Earth in Figure 8.2.1. Instead, surface winds in both hemispheres blow towards the equator between 90^{o} and 60^{o} latitude, and between 0^{o} and 30^{o} latitude. Between 30^{o} and 60^{o} latitude, the surface winds blow towards the poles (Figure 8.2.3).

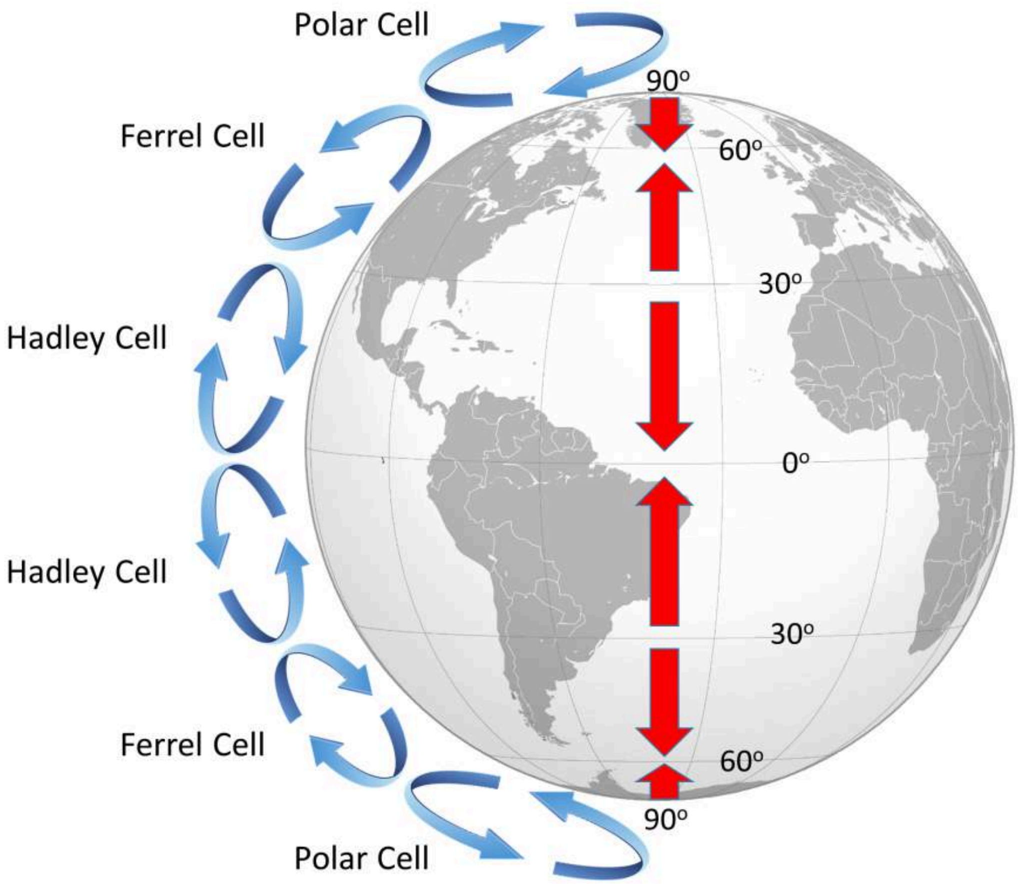


Figure 8.2.3 *On a rotating Earth, there are three atmospheric convection cells in each hemisphere, leading to alternating bands of surface winds (red arrows) (Modified by PW from globe image by Location_of_Cape_Verde_in_the_globe.svg: Eddo derivative work: Luan fala! [CC BY-SA 3.0], via Wikimedia Commons).*

The surface winds created by the atmospheric convection cells are also influenced by the Coriolis Effect as they change latitudes. The Coriolis Effect deflects the path of the winds to the right in the Northern Hemisphere and to the left in the Southern Hemisphere. Adding this deflection leads to the pattern of prevailing winds illustrated in Figure 8.2.4. Between the equator and 30^{o} latitude are the **trade winds**; the northeast trade winds in the Northern Hemisphere and the southeast trade winds in the Southern Hemisphere (note that winds are named based on the direction from which they originate, not where they are going). The **westerlies** are the dominant winds between 30^{o} and 60^{o} in both hemispheres, and the **polar easterlies** are found between 60^{o} and the poles.

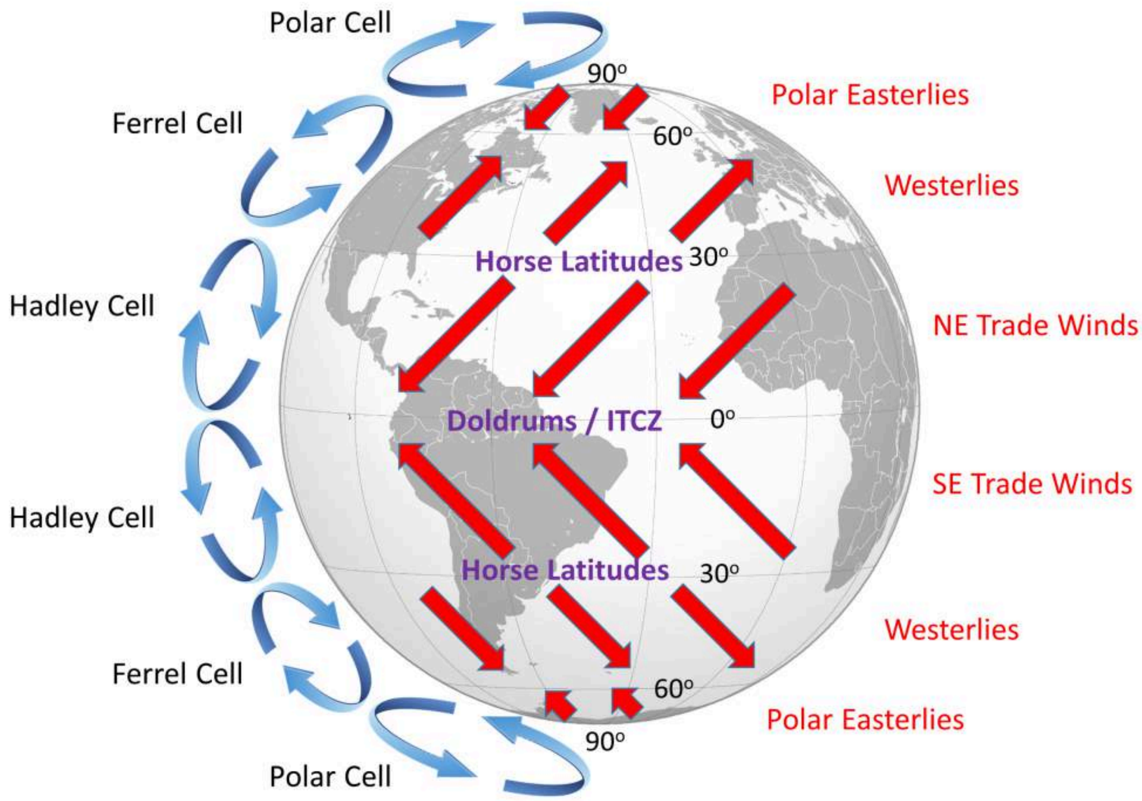


Figure 8.2.4 *The prevailing wind patterns of Earth (Modified by PW from globe image by Location_of_Cape_Verde_in_the_globe.svg: Eddo derivative work: Luan fala! [CC BY-SA 3.0], via Wikimedia Commons).*

In between these wind bands lie regions of high and low pressure. High pressure zones occur where air is descending, while low pressure zones indicate rising air. Along the equator the rising air creates a low pressure region called the **doldrums**, or the **Intertropical Convergence Zone** (ITCZ)(convergence zone because this is where the trade winds converge). At 30^{o} latitude there are high pressure zones of descending air known as the **horse latitudes**, or the subtropical highs. Finally, at 60^{o} lies another low pressure region called the **polar front**. It should be noted that these high and low pressure zones are not fixed in place; their latitude fluctuates depending on the season, and these fluctuations have important implications for regional climates.

Doldrums? Horse latitudes? Trade winds?

These may seem like some odd names for these atmospheric phenomena, but many of them can be traced back to maritime traditions and lore.

The **doldrums** refer to regions of low pressure around the equator. In these areas, air is rising rather than moving horizontally, so these regions commonly encounter very light winds. The lack of wind could leave sailing ships becalmed for days or weeks at a time, which was not good for the morale of the ship's crew.

Like the doldrums the **horse latitudes** are also areas with light winds, this time due to descending air, which could leave ships becalmed. One explanation for the term "horse latitudes" is that when these ships became stranded they ran the risk of running out of food or water. To conserve these resources, sailors would throw their dead or dying horses overboard, hence the "horse latitudes." Another explanation is that many sailors received part of their pay before a voyage, and often spent it before departing. This meant that they would spend the first part of the voyage working without pay and in debt, a period called the "dead horse" time, which might last for a few months. When they started earning their pay once again, they had a "dead horse" ceremony and threw a pretend horse overboard. The timing of this ceremony often coincided with reaching the horse latitudes, leading to the association of the ceremony with the location. A third explanation is that a ship was referred to as "horsed" when winds were weak and the ship instead had to rely on ocean currents to move them. This could be a common occurrence in the high pressure zones around 30^{o} latitude, so they were referred to as the horse latitudes.

The term **trade winds** may have originally derived from the terms for "track" or "path", but the term may have become more common during European exploration and commercialization of the New World. Mariners sailing from Europe to the New World could sail south until they reached the trade winds, which would then propel their ships across the Atlantic to the Caribbean. To return to Europe, ships could sail to the northeast until they entered the westerlies, which would then steer them back to Europe.

8.3 Winds and Climate

In the previous section we learned that rising air creates low pressure systems, and sinking air creates high pressure. In addition to their role in creating the surface winds, these high and low pressure systems also influence other climatic phenomena. Along the equator air is rising as it is warmed by solar radiation (section 8.2). Warm air contains more water vapor than cold air, which is why we experience humidity during the summer and not during the winter. The water content of air roughly doubles with every 10° C increase in temperature. So the air rising at the equator is warm and full of water vapor; as it rises into the upper atmosphere it cools, and the cool air can no longer hold as much water vapor, so the water condenses and forms rain. Therefore, low pressure systems are associated with precipitation, and we see wet habitats like tropical rainforests near the equator (Figure 8.3.1).

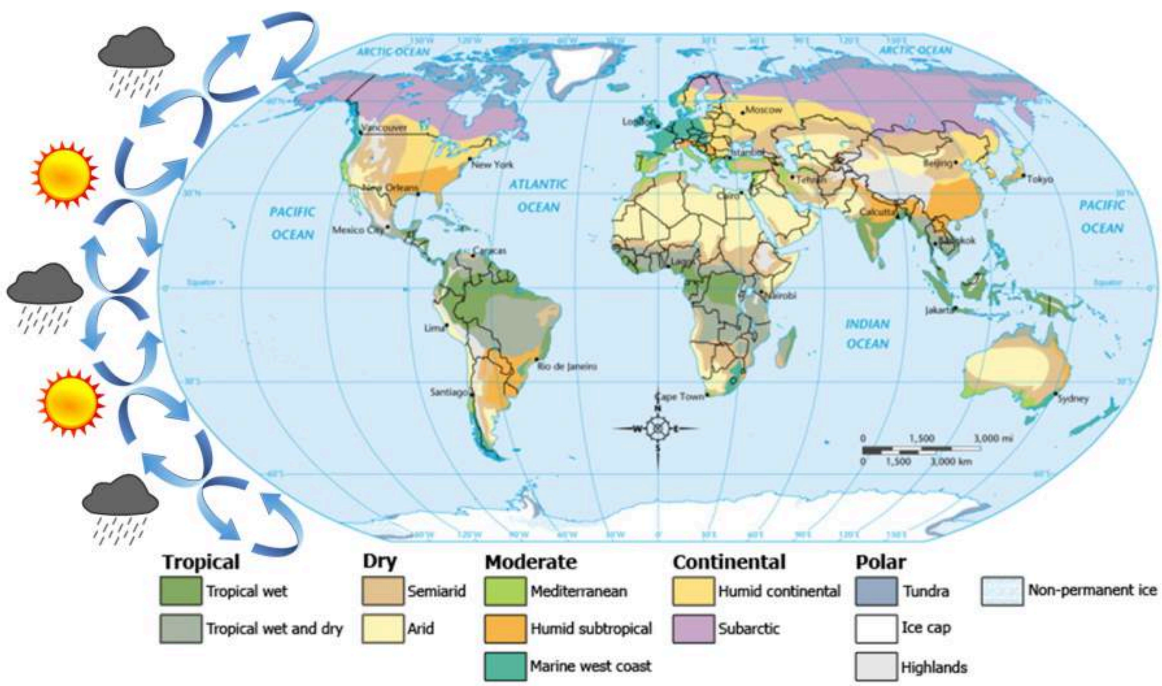


Figure 8.3.1 *Major global climatic regions in relation to atmospheric convection cells. Rising air and low pressure creates rain and wet environments at 0° and 60° latitudes, while high pressure, sinking air creates drier conditions at 30° and 90° latitudes. (Modified by PW; Map by Waitak at en.wikipedia Later version(s) were uploaded by Splette at en.wikipedia; Sun by Inductiveload (Own work Based on File:Nuvola_apps_kweather.svg); Raincloud by Calusarul (Own work); all [CC BY-SA 3.0], via Wikimedia Commons).*

After rising and producing rain near the equator, the air masses move towards 30° latitude and sink back towards Earth as part of the Hadley convection cells. This air has lost most of its moisture after producing the equatorial rains, so the sinking air is dry, resulting in arid climates near 30° latitude in both hemispheres. Many of the major desert regions on Earth are located near 30° latitude, including much of Australia, the Middle East, and the Sahara Desert of Africa (Figure 8.3.1). The air also becomes compressed and heats up as it sinks, absorbing any moisture from the clouds and creating clear skies. Thus high pressure systems are associated with dry weather and clear skies. This cycle of high and low pressure regions continues with the Ferrel and Polar convection cells, leading to rain and the boreal forests at 60° latitude in the Northern Hemisphere (there are no corresponding large land masses at these latitudes in the Southern Hemisphere). At the poles, descending, dry air produces little precipitation, leading to the polar desert climate.

The elevation of the land also plays a role in precipitation and climactic characteristics. As moist air moves over land and encounters mountains it rises, expands, and cools because of the declining pressure and temperature. The cool air holds less water vapor, so condensation occurs and rain falls on the windward side of the mountains. As the air passes over the mountains to the leeward side, it is now dry air, and as it sinks the pressure increases, it heats back up, any moisture revaporizes, and it creates dry, deserts regions behind the mountains (Figure 8.3.2). This phenomenon is referred to as a **rain shadow**, and can be found in areas such as the Tibetan Plateau and Gobi Desert behind the Himalayas, Death Valley behind the Sierra Nevada mountains, and the dry San Joaquin Valley in California.

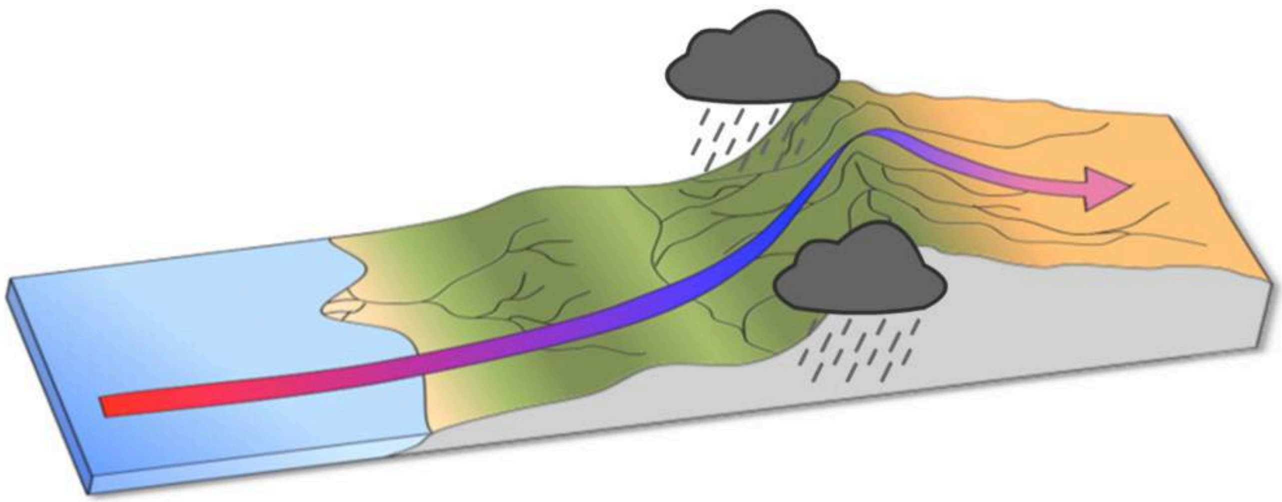


Figure 8.3.2 *A rain shadow. Air rising over mountains cools and condenses and forms rain, leaving dry descending air and arid conditions on the other side of the mountain (Modified by PW from Thebiologyprimer, Public domain via Wikimedia Commons).*

Rising and falling air are also responsible for more localized, short-term wind patterns in coastal areas. Due to the high heat capacity of water, land heats up and cools down about five times faster than water. During the day the sun heats up the land faster than it heats the water, setting up a convection cell of warmer rising air over the land and sinking cooler air over the water. This creates winds blowing from the water towards the land during the day and early evening; a **sea breeze** (Figure 8.3.3). The opposite occurs at night, when the land cools more quickly than the ocean. Now the ocean is warmer than the land, so air rises over the water and sinks over the land, creating a convection cell where winds blow from land towards the water. This is a **land breeze**, which blows at night and into the early morning (Figure 8.3.3).

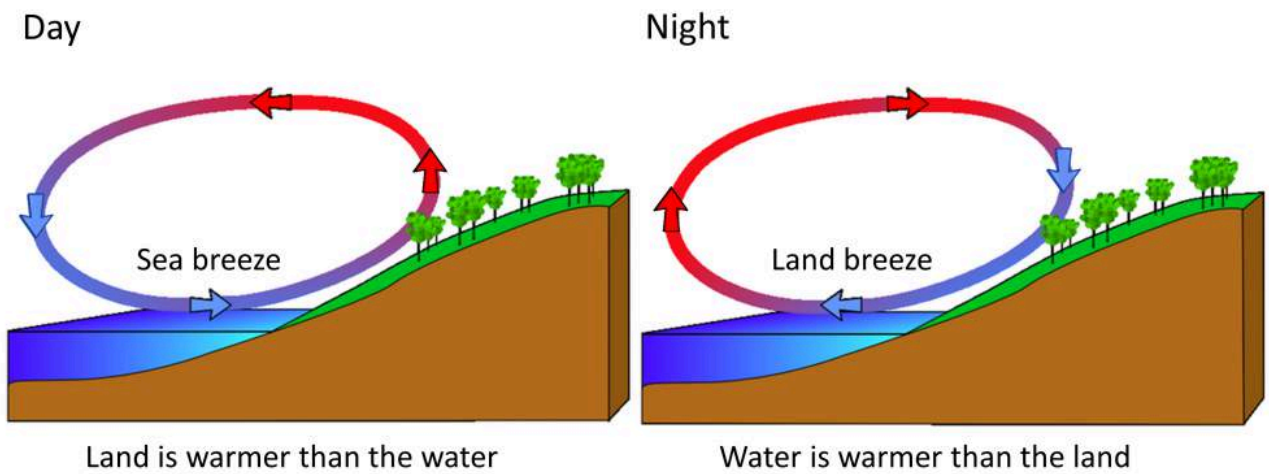


Figure 8.3.3 *Land heats and cools faster than the ocean, so during the day the land is warmer than the water leading to rising air over land and a sea breeze. At night, the ocean is warmer than the land, creating a land breeze (Modified by PW derivative work: Ingwik (Diagrama de formacion de la brisa-breeze.png) [CC-BY-SA-3.0 or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons).*

The same phenomenon leads to seasonal climatic changes in many areas. During the winter the lower pressure is over the warmer ocean, and the high pressure is over the colder land, so winds blow from land to sea. In summer the land is warmer than the ocean, causing low pressure over the land and winds to blow from the ocean towards the land. The winds blowing from the ocean contain a lot of water vapor, and as the moist air passes over land and rises, it cools and condenses causing seasonal rains, such as the summer **monsoons** of southeast Asia (Figure 8.3.4).

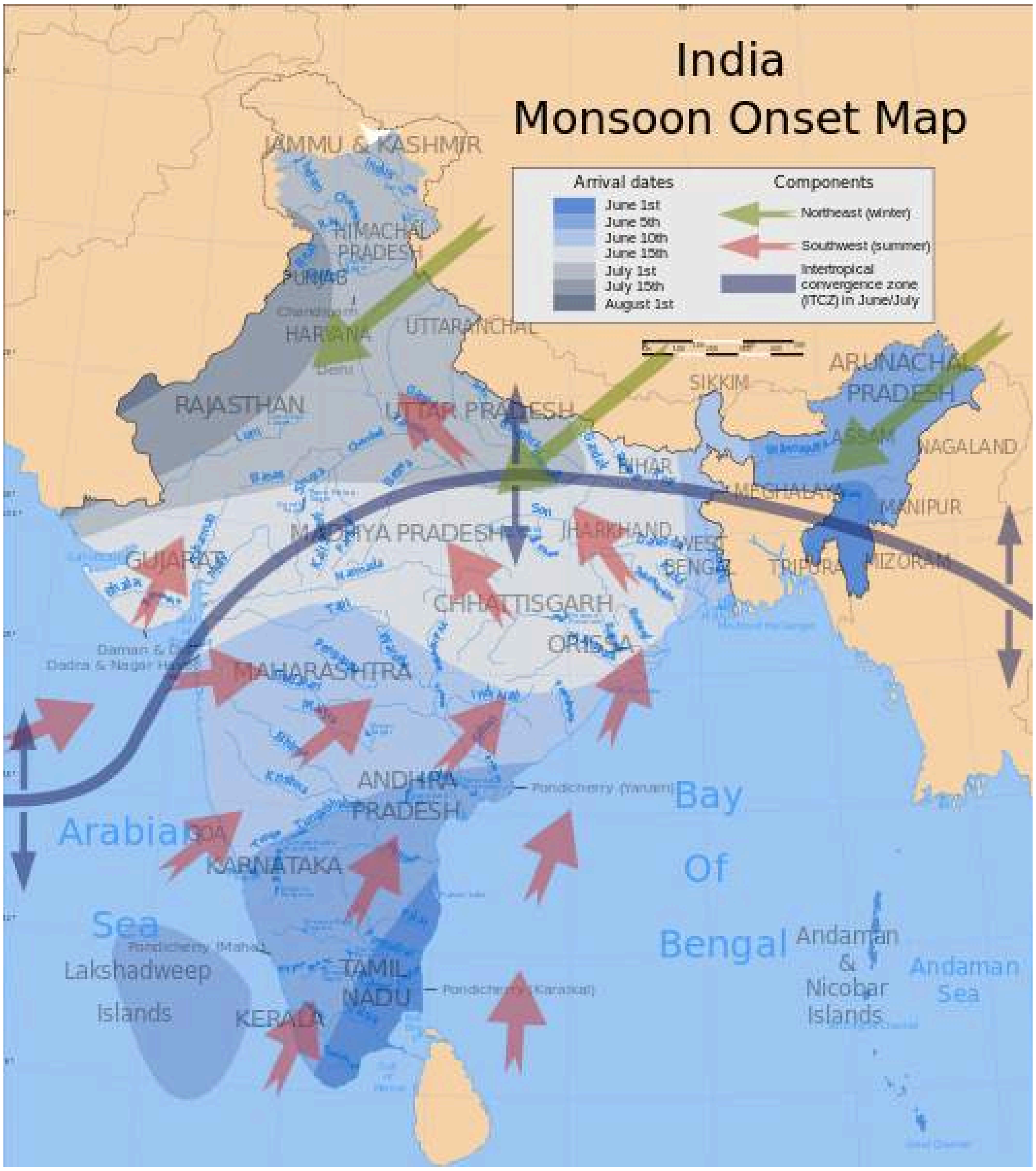


Figure 8.3.4 *Seasonal wind patterns and monsoons over India. In summer, moist air from the ocean moves over the continent and rises, creating rain and the summer monsoons (pink arrows). In winter, winds are blowing from land to the sea, leading to the dry season (green arrows) (By Saravask, based on work by Planemad and Nichalp [CC BY-SA 3.0], via Wikimedia Commons).*

8.4 Hurricanes

The most dramatic examples of low pressure systems leading to storms and rain are hurricanes, cyclones, and typhoons. All three of these terms describe the same atmospheric processes and the same types of storms; it's just that different terminology is used in different parts of the world. In the Atlantic and Northeast Pacific, the storms are called hurricanes, in the Indian and South Pacific Oceans they are referred to as cyclones, and they are called typhoons in the Northwest Pacific.

Hurricanes begin as low pressure systems formed over warm, tropical water. They only form in tropical regions because they need the heat from the warm water to fuel the storm. The warm, moist air rises, cools, and condenses, forming rain, and the condensation releases more latent heat into the atmosphere. This heat causes even more air to rise and condense, further fueling the storm.

As the air rises towards the center of the storm, more warm tropical air rushes in to replace it, causing very strong winds. But the air does not move directly towards the center of the storm. Because of the large size of hurricanes, the air rushing towards the center will be deflected by the Coriolis Effect, causing the entire storm to rotate. In the Northern Hemisphere that deflection is to the right, causing Northern Hemisphere hurricanes to rotate counterclockwise. In the Southern Hemisphere, the winds are deflected to the left, leading to a clockwise rotation (Figure 8.4.1).

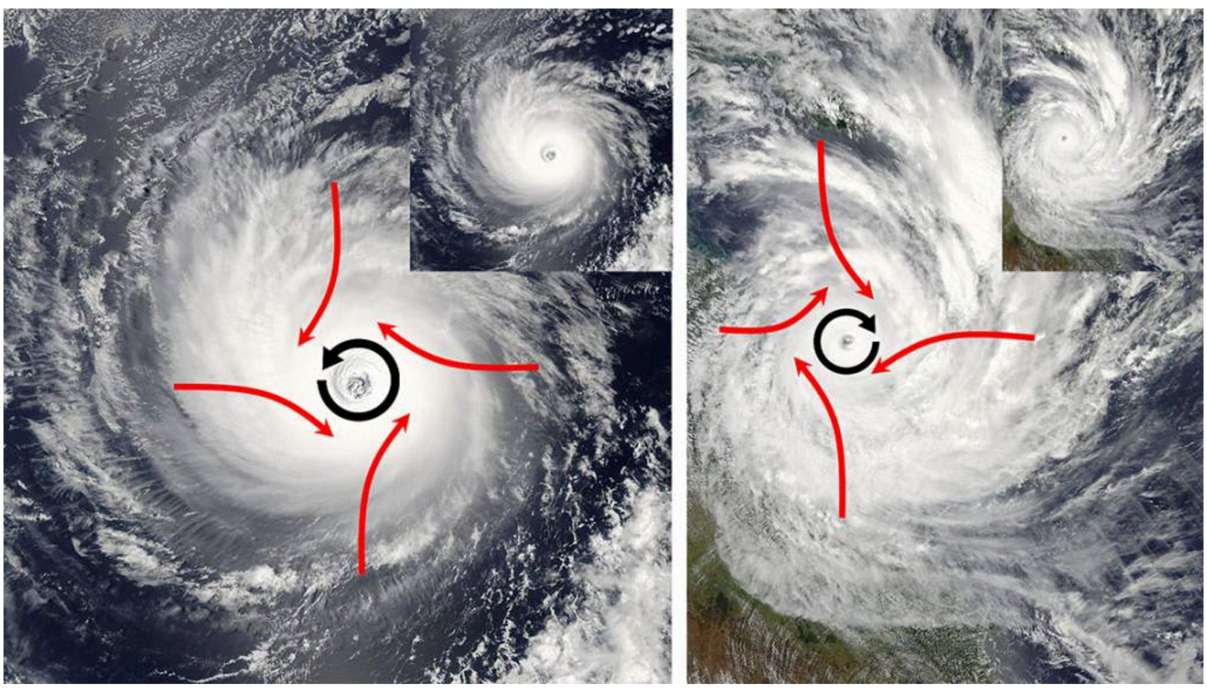


Figure 8.4.1 *Hurricanes in the Northern Hemisphere rotate counterclockwise (left, Hurricane Daniel, 2006), as air rushes towards the center and is deflected to the right by the Coriolis Effect. In the Southern Hemisphere, hurricanes rotate clockwise as the Coriolis deflection is to the left (right, Cyclone Yasi, 2011) (Modified by PW; Daniel image by NASA image courtesy Jeff Schmaltz, MODIS Land Rapid Response Team at NASA GSFC; Yasi image by NASA; MODIS [Public domain], via Wikimedia Commons).*

The violent winds characteristic of hurricanes are the result of the spiraling air that is moving towards the center of the storm, and once its winds exceed 74 mph the storm officially becomes a hurricane. At the very center of the hurricane, the pressure is so low that cool, dry air from the upper atmosphere get sucked downwards, leading to a central region of calm, clear skies; the hurricane's eye (Figure 8.4.2).

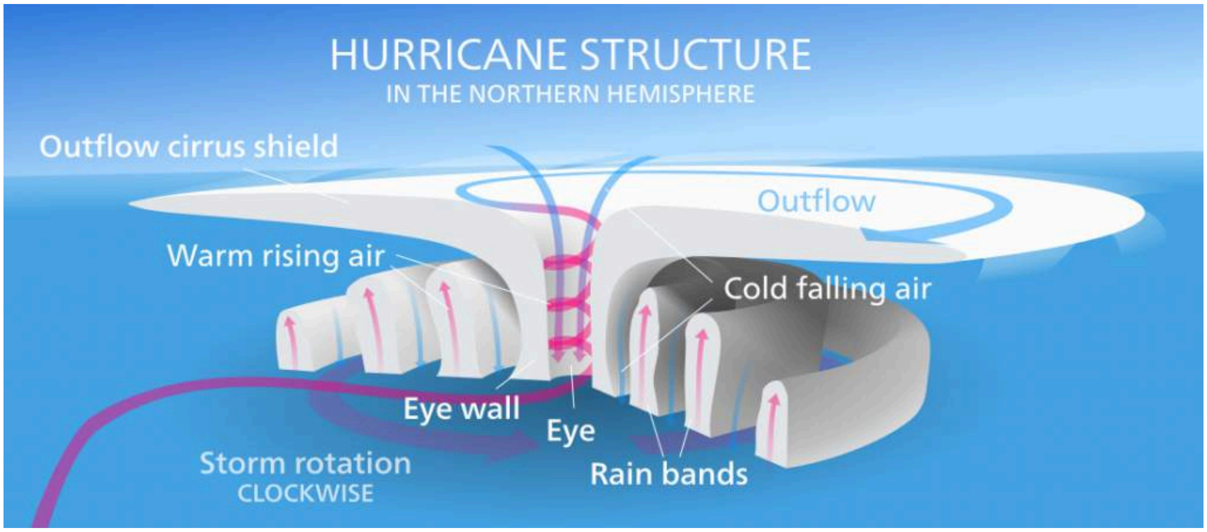


***Figure 8.4.2** Hurricane structure. Air rising in the center of the hurricane is replaced by warm air moving inwards, and the Coriolis Effect deflects the winds, causing the storm to rotate. In the eye, the extreme low pressure causes cool, dry air to sink, creating calm, clear conditions within the eye (By Kelvinsong (Own work) [CC BY 3.0], via Wikimedia Commons).*

Hurricanes in the North Atlantic form as tropical storms over the warm water off of the African coast, and are moved east to west by the trade winds (Figure 8.4.3). As the storms move west over the tropical ocean, their energy increases until they reach hurricane status. As they approach the Caribbean, the Coriolis Effect deflects their path to the right, causing them to move towards the north (Figure 8.4.3). Eventually hurricanes might make landfall, causing extensive damage to coastal areas through the high winds, rain, and flooding. However, hurricanes often die out fairly soon after reaching land. When a storm moves over land it becomes cut off from the warm moist ocean air that has sustained it. Without that fuel source, the storm loses power and begins to dissipate.

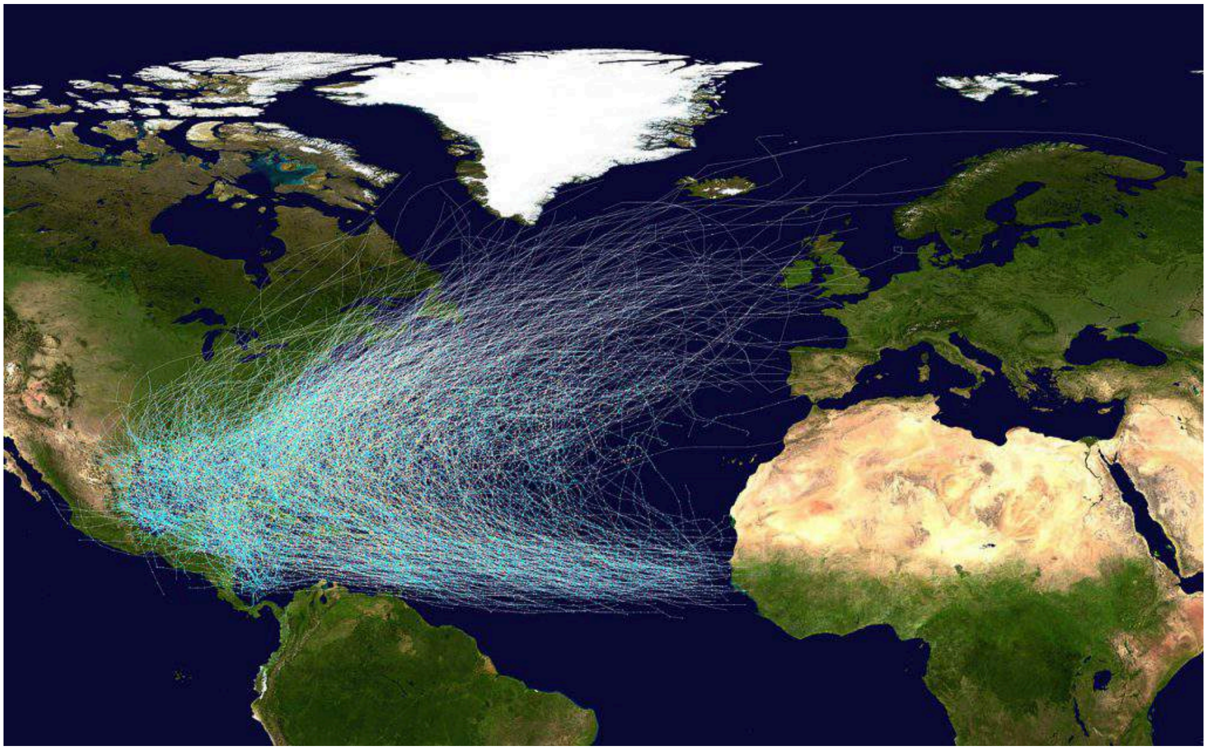


Figure 8.4.3 *Hurricane tracks in the North Atlantic from 1980-2005. Hurricanes begin near the coast of Africa and are blown westward by the trade winds. Coriolis deflection causes them to take a northward path as they approach the Caribbean (By Nilfanion [Public domain], via Wikimedia Commons).*

A similar pattern occurs in the Pacific and in the Southern Hemisphere. The trade winds move the storms from east to west, and they are deflected as they approach the coasts; to the right in the Northern Hemisphere and to the left in the Southern Hemisphere (Figure 8.4.4).

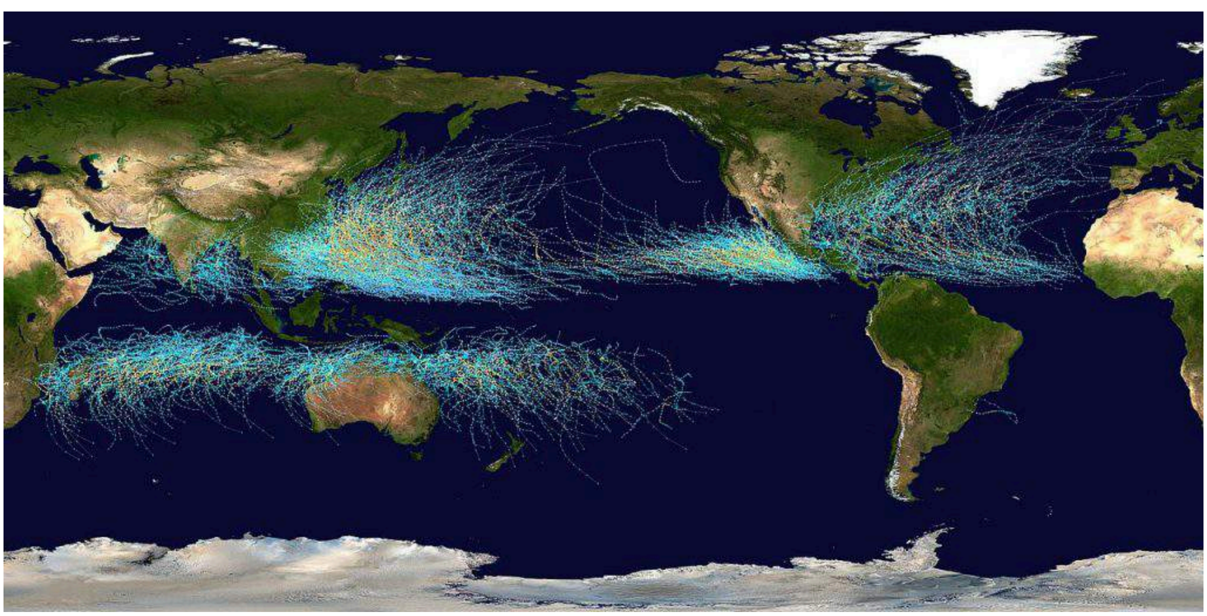


Figure 8.4.4 *Global hurricane/cyclone tracks 1985-2005. Hurricanes move west via the trade winds, and their paths are deflected away from the equator in both hemispheres as they approach the continents (Background image: NASA this version: Nilfanion [Public domain], via Wikimedia Commons).*

While the very high winds and intense rain of hurricanes can cause significant damage, in many cases it is the **storm surge** that leads to the most death and destruction. The storm surge is a “hill” of water that forms on the ocean surface below a hurricane. The surge is the result of two processes; a small hill is produced due to the extreme low pressure in the eye of a hurricane, which pulls water upwards towards the eye, creating a pressure surge. A larger surge is produced by the winds blowing and piling up water in the direction the storm is traveling (Figure 8.4.5). As the hurricane makes landfall, the effect of the storm surge is equivalent to a very large and sudden rise in sea level as the surge moves over the land, causing extensive flooding.

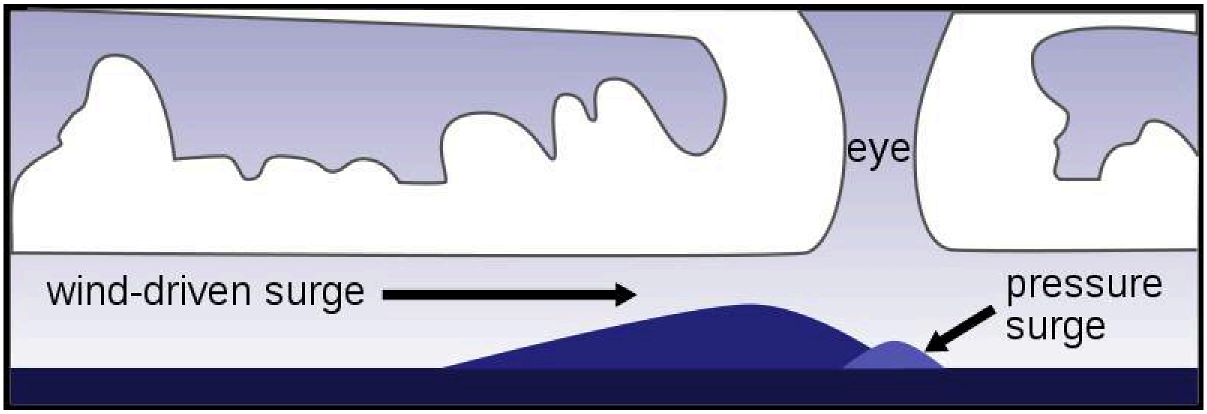


Figure 8.4.5 *Storm surges are created by hurricanes and move with the storm, causing a rapid rise in sea level when they reach shore. Pressure surges are due to the low pressure within the hurricane’s eye, wind-driven surges are a product of high winds piling up water (By Howcheng. Original graphic by Robert Simmon, NASA GSFC. [Public domain], via Wikimedia Commons).*

In 1970 the Bhola Cyclone struck Bangladesh with a 40 ft. storm surge, leading to the death of about 500,000 people, the deadliest hurricane in history. The east coast of the United States was hit by the New England Hurricane of 1938, which had a 16 ft. storm surge and left almost 700 people dead.

Preventing Storm Surge Damage

In response to the hurricane-related tragedies like those listed above, may cities have built hurricane barriers designed to reduce the flooding and damage associated with storm surges. Downtown Providence, Rhode Island, USA, was submerged under 13 feet of water during the Great New England hurricane of 1938, and was flooded again following Hurricane Carol in 1954. In the 1960s the Fox Point Hurricane Barrier was constructed at the mouth of the Providence River. It consists of a high wall with three “doors” that are left open under normal conditions, but can be closed during a hurricane to prevent a storm surge of up to 20.5 feet from inundating the city (Figure 8.4.6, left). A related concept is seen in the storm surge barrier on the Hollandse IJssel river in the Netherlands, where the barrier is lowered to prevent flooding (Figure 8.4.6, right).

Figure 8.4.6 *The Fox Point Hurricane Barrier in Providence, RI, USA (left) and the surge barrier on the Hollandse IJssel river in southern Holland. (Fox Point by Marcbela (Marc N. Belanger) (Own work) [Public domain], via Wikimedia Commons; Dutch barrier by Mark Voorendt (Own work) [CC BY-SA 4.0], via Wikimedia Commons).*

8.5 Climate Change

*Modified from "Physical Geology" by Steven Earle**

If one thing has been constant about Earth's climate over geological time, it is its constant change. In the geological record, we can see this in the evidence of glaciations in the distant past, and we can also detect periods of extreme warmth by looking at the isotope composition of seafloor sediments. Not only has the climate changed frequently, the temperature fluctuations have been very significant. Today's mean global temperature is about 15°C. However, during its coldest periods, the global mean was as cold as -50°C, while at various times during the Paleozoic and Mesozoic and during the Paleocene–Eocene thermal maximum, it was close to 30°C.

There are two parts to climate change, the first one is known as **climate forcing**, which is when conditions change to give the climate a little nudge in one direction or the other. The second part of climate change, and the one that typically does most of the work, is what we call a **feedback**. When a climate forcing changes the climate a little, a whole series of environmental changes take place, many of which either exaggerate the initial change (**positive feedback**), or suppress the change (**negative feedback**).

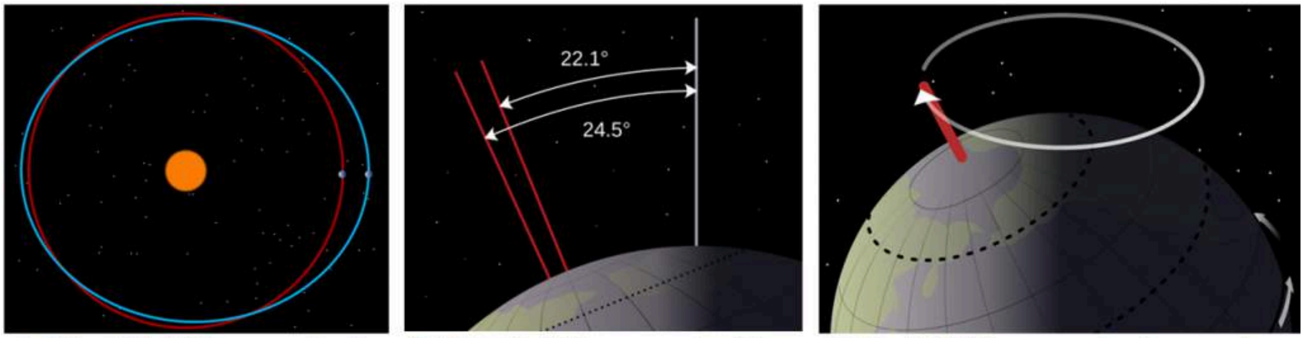
An example of a climate forcing mechanism is the increase in the amount of carbon dioxide (CO_2) in the atmosphere that results from our use of fossil fuels. CO_2 traps heat in the atmosphere and leads to climate warming. Warming changes vegetation patterns; contributes to the melting of snow, ice, and permafrost; causes sea level to rise; reduces the solubility of CO_2 in sea water; and has a number of other minor effects. Most of these changes contribute to more warming. Melting of permafrost, for example, is a strong positive feedback because frozen soil contains trapped organic matter that is converted to CO_2 and methane (CH_4) when the soil thaws. Both these gases accumulate in the atmosphere and add to the warming effect. On the other hand, if warming causes more vegetation growth, that vegetation should absorb CO_2, thus reducing the warming effect, which would be a negative feedback. Under our current conditions — a planet that still has lots of glacial ice and permafrost — most of the feedbacks that result from a warming climate are positive feedbacks and so the climate changes that we cause get naturally amplified by natural processes.

Natural Climate Forcing

Natural climate forcing has been going on throughout geological time. A wide range of processes has been operating at widely different time scales, from a few years to billions of years. The longest-term natural forcing variation is related to the evolution of the Sun. Like most other stars of a similar mass, our Sun is evolving. For the past 4.6 billion years, its rate of nuclear fusion has been increasing, and it is now emitting about 40% more energy (as light) than it did at the beginning of geological time. A difference of 40% is big, so it's a little surprising that the temperature on Earth has remained at a reasonable and habitable temperature for all of this time. The mechanism for that relative climate stability has been the evolution of our atmosphere from one that was dominated by CO_2, and also had significant levels of CH_4 — both greenhouse gasses — to one with only a few hundred parts per million of CO_2 and just under 1 part per million of CH_4. Those changes to our atmosphere have been no accident; over geological time, life and its metabolic processes have evolved (such as the evolution of photosynthetic bacteria that consume CO_2) and changed the atmosphere to conditions that remained cool enough to be habitable.

The position of the Earth relative to the Sun is another important component of natural climate forcing. Earth's orbit around the Sun is nearly circular, but like all physical systems, it has natural oscillations. First, the shape of the orbit changes on a regular time scale (close to 100,000 years) from being close to circular to being very slightly elliptical. But the circularity of the orbit is not what matters; it is the fact that as the orbit becomes more elliptical, the position of the Sun within that ellipse becomes less central or more **eccentric** (Figure 8.5.1a).

Eccentricity is important because when it is high, the Earth-Sun distance varies more from season to season than it does when eccentricity is low.



(a) The 100,000 year cycle of eccentricity of the Sun within the Earth's orbit. The blue orbit is more elliptical than the red one, and the Sun is offset from the centre of the ellipse

(b) The 41,000 year cycle of the angle of tilt (obliquity) of the Earth's axis (red lines) compared to a line perpendicular to the plane of the Earth's orbit (light blue line).

(c) the 20,000 year cycle of precession of the Earth's rotational axis

Figure 8.5.1 *Components of the Milankovitch cycles, which influence global climate over thousands of years (Steven Earle, "Physical Geology").*

Second, Earth rotates around an axis through the North and South Poles, and that axis is at an angle to the plane of Earth's orbit around the Sun (Figure 8.5.1b). The angle of tilt (also known as **obliquity**) varies on a time scale of 41,000 years. When the angle is at its maximum (24.5°), Earth's seasonal differences are accentuated. When the angle is at its minimum (22.1°), seasonal differences are minimized. The current hypothesis is that glaciation is favored at low seasonal differences as summers would be cooler and snow would be less likely to melt and more likely to accumulate from year to year. Third, the direction in which Earth's rotational axis points also varies, on a time scale of about 20,000 years (Figure 8.5.1c). This variation, known as **precession**, means that although the North Pole is presently pointing to the star Polaris (the pole star), in 10,000 years it will point to the star Vega. The importance of eccentricity, tilt, and precession to Earth's climate cycles (now known as **Milankovitch Cycles**) was first pointed out by Yugoslavian engineer and mathematician Milutin Milankovitch in the early 1900s. Milankovitch recognized that although the variations in the orbital cycles did not affect the total amount of insolation (light energy from the Sun) that Earth received, it did affect where on Earth that energy was strongest.

Volcanic eruptions don't just involve lava flows and exploding rock fragments; various particulates and gases are also released, the important ones being sulphur dioxide and CO_2. Sulphur dioxide is an aerosol that reflects incoming solar radiation and has a net cooling effect that is short lived (a few years in most cases, as the particulates settle out of the atmosphere within a couple of years), and doesn't typically contribute to longer-term climate change. Volcanic CO_2 emissions can contribute to climate warming but only if a greater-than-average level of volcanism is sustained over a long time (at least tens of thousands of years). It is widely believed that the catastrophic end-Permian extinction (at 250 Ma) resulted from warming initiated by the eruption of the massive Siberian Traps over a period of at least a million years.

Ocean currents are important to climate, and currents also have a tendency to oscillate. Glacial ice cores show clear evidence of changes in the Gulf Stream that affected global climate on a time scale of about 1,500 years during the last glaciation. The east-west changes in sea-surface temperature and surface pressure in the equatorial Pacific Ocean, known as the El Niño Southern Oscillation or ENSO (see section 9.6) varies on a much shorter time scale of between two and seven years. These variations tend to garner the attention of the public because they have significant climate implications in many parts of the world. The strongest El Niños in recent decades were in 1983, 1998, and 2015 and those were very warm years from a global perspective. During a strong El Niño, the equatorial Pacific sea-surface temperatures are warmer than normal and heat the atmosphere above the ocean, which leads to warmer-than-average global temperatures.

Climate Feedbacks

As already stated, climate feedbacks are critically important in amplifying weak climate forcings into full-blown climate changes. Since Earth still has a very large volume of ice, mostly in the continental ice sheets of Antarctica and Greenland, but also in alpine glaciers and permafrost, melting is one of the key feedback mechanisms. Melting of ice and snow leads to several different types of feedbacks, an important one being a change in albedo, or the reflectivity of a surface. Earth's various surfaces have widely differing albedos, expressed as the percentage of light that reflects off a given material. This is important because most solar energy that hits a very reflective surface is not absorbed and therefore does little to warm Earth. Water in the oceans or on a lake is one of the darkest surfaces, reflecting less than 10% of the incident light, while clouds and snow or ice are among the brightest surfaces, reflecting 70% to 90% of the incident light. When sea ice melts, as it has done in the Arctic Ocean at a disturbing rate over the past decade, the albedo of the area affected changes dramatically, from around 80% down to less than 10%. Much more solar energy is absorbed by the water than by the pre-existing ice, and the temperature increase is amplified. The same applies to ice and snow on land, but the difference in albedo is not as great. When ice and snow on land melt, sea level rises. (Sea level is also rising because the oceans are warming and that increases their volume; see section 13.7). A higher sea level means a larger proportion of the planet is covered with water, and since water has a lower albedo than land, more heat is absorbed and the temperature goes up a little more. Since the last glaciation, sea-level rise has been about 125 m; a huge area that used to be land is now flooded by heat-absorbent seawater. During the current period of anthropogenic climate change, sea level has risen only about 20 cm, and although that doesn't make a big change to albedo, sea-level rise is accelerating.

Most of northern Canada, Alaska, Russia, and Scandinavia has a layer of permafrost that ranges from a few centimeters to hundreds of meters in thickness. Permafrost is a mixture of soil and ice and it also contains a significant amount of trapped organic carbon that is released as CO_2 and CH_4 when the permafrost breaks down. Because the amount of carbon stored in permafrost is in the same order of magnitude as the amount released by burning fossil fuels, this is a feedback mechanism that has the potential to equal or surpass the forcing that has unleashed it. In some polar regions, including northern Canada, permafrost includes methane hydrate, a highly concentrated form of CH_4 trapped in solid form. Breakdown of permafrost releases this CH_4. Even larger reserves of methane hydrate exist on the seafloor, and while it would take significant warming of ocean water down to a depth of hundreds of meters, this too is likely to happen in the future if we don't limit our impact on the climate. There is strong isotopic evidence that the Paleocene–Eocene thermal maximum was caused, at least in part, by a massive release of sea-floor methane hydrate.

There is about 45 times as much carbon in the ocean (as dissolved bicarbonate ions, HCO_3^-) as there is in the atmosphere (as CO_2), and there is a steady exchange of carbon between the two reservoirs (see section 5.5). But the solubility of CO_2 in water decreases as the temperature goes up. In other words, the warmer it gets, the more oceanic bicarbonate that gets transferred to the atmosphere as CO_2. That makes CO_2 solubility another positive feedback mechanism. Vegetation growth responds positively to both increased temperatures and elevated CO_2 levels, and so in general, it represents a negative feedback to climate change because the more the vegetation grows, the more CO_2 is taken from the atmosphere. But it's not quite that simple, because when trees grow bigger and more vigorously, forests become darker (they have lower albedo) so they absorb more heat. Furthermore, climate warming isn't necessarily good for vegetation growth; some areas have become too hot, too dry, or even too wet to support the plant community that was growing there, and it might take centuries for something to replace it successfully. All of these positive (and negative) feedbacks work both ways. For example, during climate cooling, growth of glaciers leads to higher albedos, and formation of permafrost results in storage of carbon that would otherwise have returned quickly to the atmosphere.

Anthropogenic Climate Change

When we talk about anthropogenic climate change, we are generally thinking of the industrial era, which really got going when we started using fossil fuels (coal to begin with, and later oil and natural gas) to drive machinery and trains, and to generate electricity. That was around the middle of the 18th century. The issue with fossil fuels is that they involve burning carbon that was naturally stored in the crust over hundreds of millions of years as part of Earth's process of counteracting the warming Sun.

A rapidly rising population, the escalating level of industrialization and mechanization of our lives, and an increasing dependence on fossil fuels have driven the anthropogenic climate change of the past century. The trend of mean global temperatures since 1880 is shown in Figure 8.5.2. For approximately the past 55 years, the temperature has increased at a relatively steady and disturbingly rapid rate, especially compared to past changes. The average temperature now is approximately 0.8°C higher than before industrialization, and two-thirds of this warming has occurred since 1975.

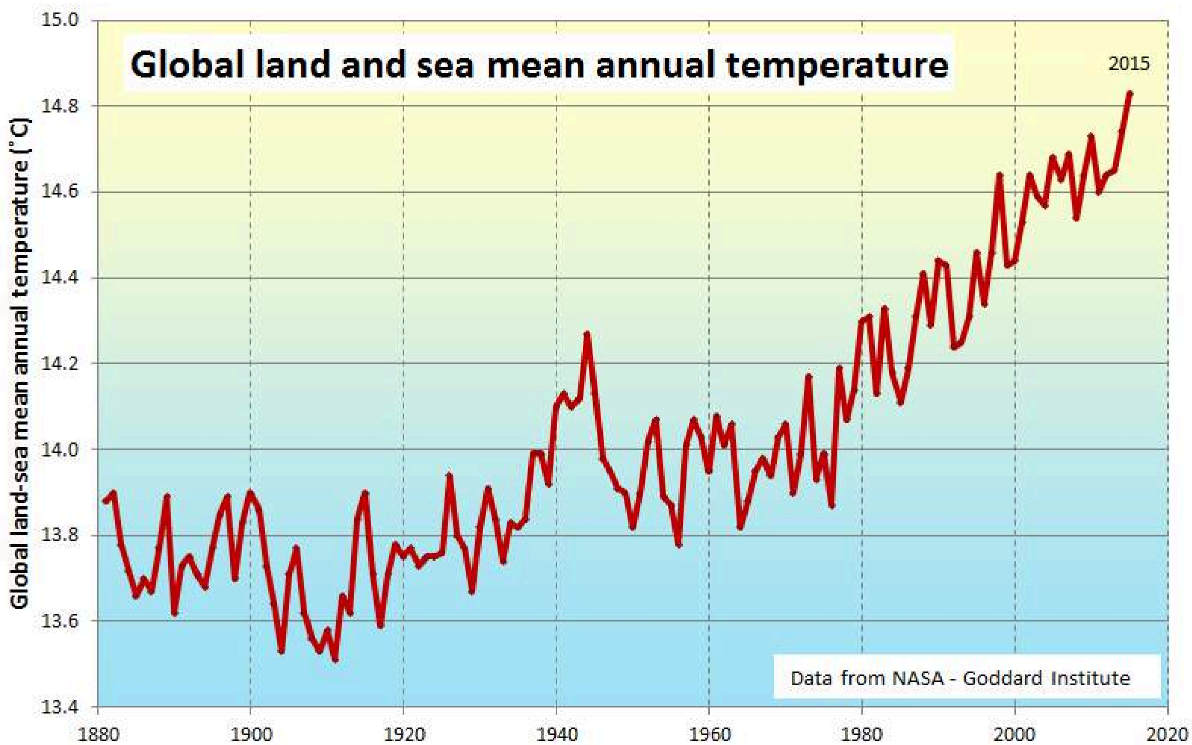


Figure 8.5.2 *Global mean annual temperatures for the period from 1880 to 2015 (Steven Earle, "Physical Geology", by SE from data at NASA at: http://data.giss.nasa.gov/gistemp/tabledata_v3/GLB.Ts+dSST.txt).*

The Intergovernmental Panel on Climate Change (IPCC), established by the United Nations in 1988, is responsible for reviewing the scientific literature on climate change and issuing periodic reports on several topics, including the scientific basis for understanding climate change, our vulnerability to observed and predicted climate changes, and what we can do to limit climate change and minimize its impacts. Figure 8.5.3, from the fifth report of the IPCC, issued in 2014, shows the relative contributions of various greenhouse gases and other factors to current climate forcing, based on the changes from levels that existed in 1750.

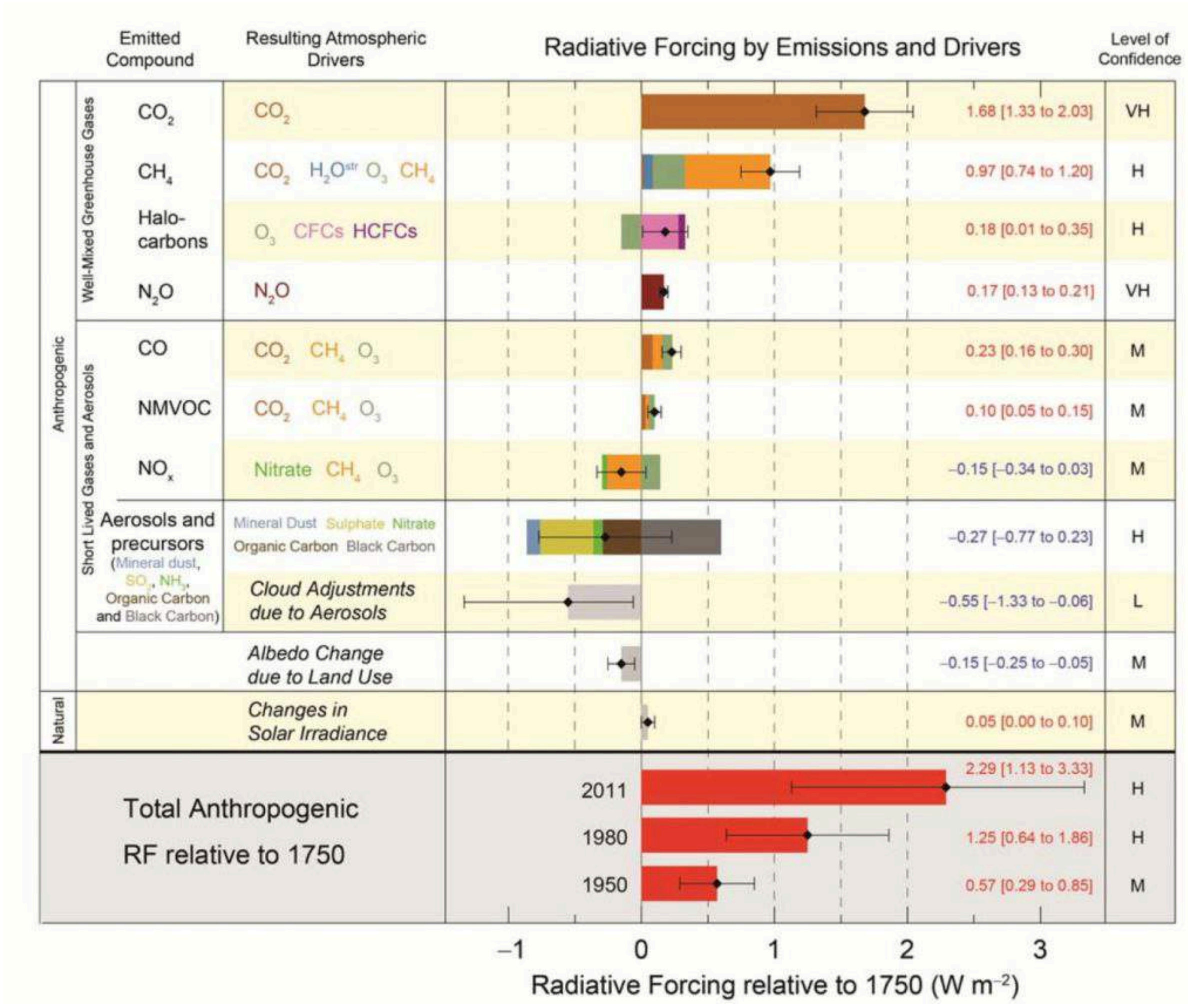


Figure 8.5.3 *The relative importance of factors that are contributing to anthropogenic warming (from http://www.ipcc.ch/report/graphics/index.php?t=Assessment%20Reports&r=AR5%20-%20WG1&f=SPM).*

The biggest anthropogenic contributor to warming is the emission of CO_2, which accounts for 50% of positive forcing. CH_4 and its atmospheric derivatives (CO_2, H_2O, and O_3) account for 29%, and the halocarbon gases (mostly leaked from air-conditioning appliances) and nitrous oxide (N_2O) (from burning fossils fuels) account for 5% each. Carbon monoxide (CO) (also produced by burning fossil fuels) accounts for 7%, and the volatile organic compounds other than methane (NMVOC) account for 3%. CO_2 emissions come mostly from coal- and gas-fired power stations, motorized vehicles (cars, trucks, and aircraft), and industrial operations (e.g., smelting), and indirectly from forestry. CH_4 emissions come from production of fossil fuels (escape from coal mining and from gas and oil production), livestock farming (mostly beef), landfills, and wetland rice farming. N_2O and CO come mostly from the combustion of fossil fuels. In summary, close to 70% of our current greenhouse gas emissions come from fossil fuel production and use, while most of the rest comes from agriculture and landfills. Figure 8.5.4 shows the IPCC's projections for temperature increases over the next 100 years as a result of these increasing greenhouse gases.

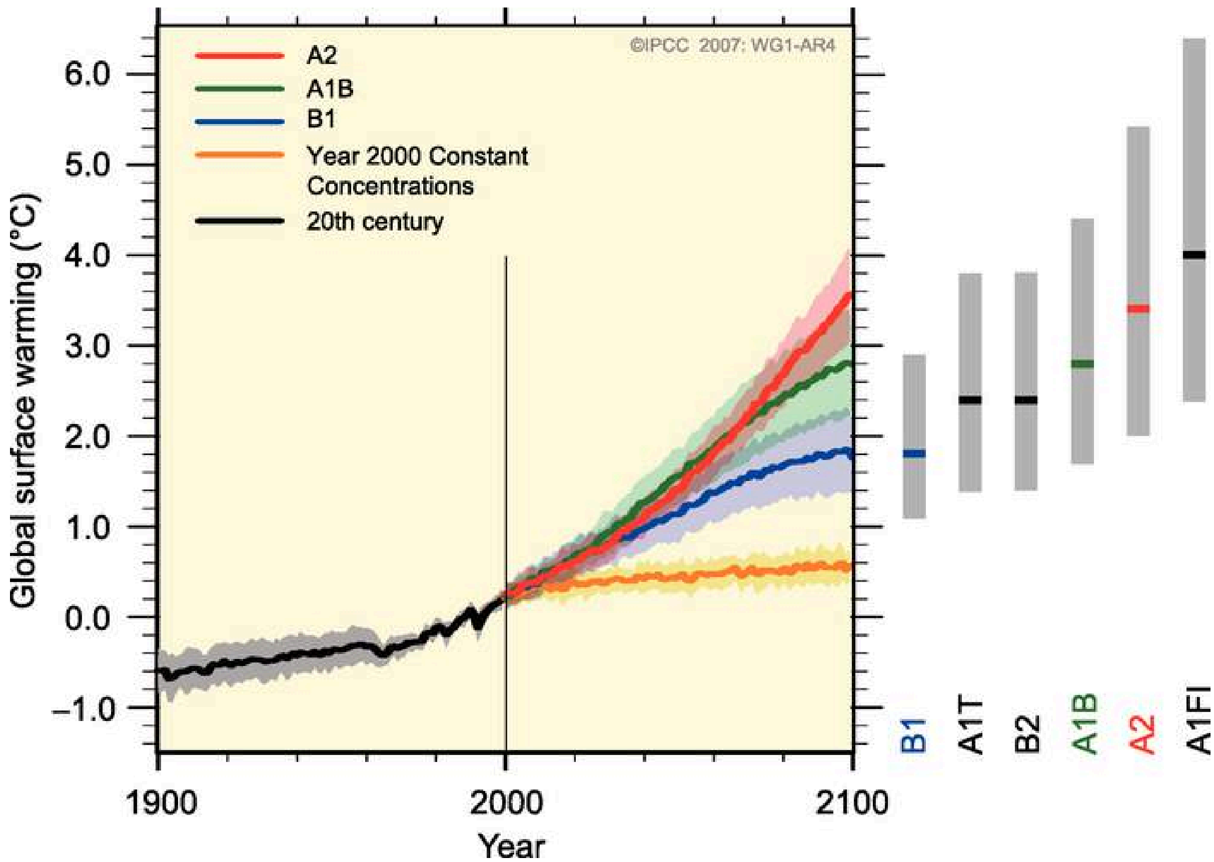


Figure 8.5.4 *Projected global temperature increases for the 21st century based on a range of different IPCC scenarios of future political and technological variables (from https://www.ipcc.ch/publications_and_data/ar4/wg1/en/fig/figure-spm-5-l.png).*

Impacts of Climate Change

We've all experienced the effects of climate change over the past decade. However, it's not straightforward for climatologists to make the connection between a warming climate and specific weather events, and most are justifiably reluctant to ascribe any specific event to climate change. In this respect, the best measures of climate change are those that we can detect over several decades, such as the temperature changes shown in Figure 8.5.2, or the sea level rise shown in Figure 8.5.5. As already stated, sea level has risen approximately 20 cm since 1750, and that rise is attributed to both warming (and therefore expanding) seawater and melting glaciers and other land-based snow and ice (melting of sea ice does not contribute directly to sea level rise as it is already floating in the ocean, see [section 13.7](#)).

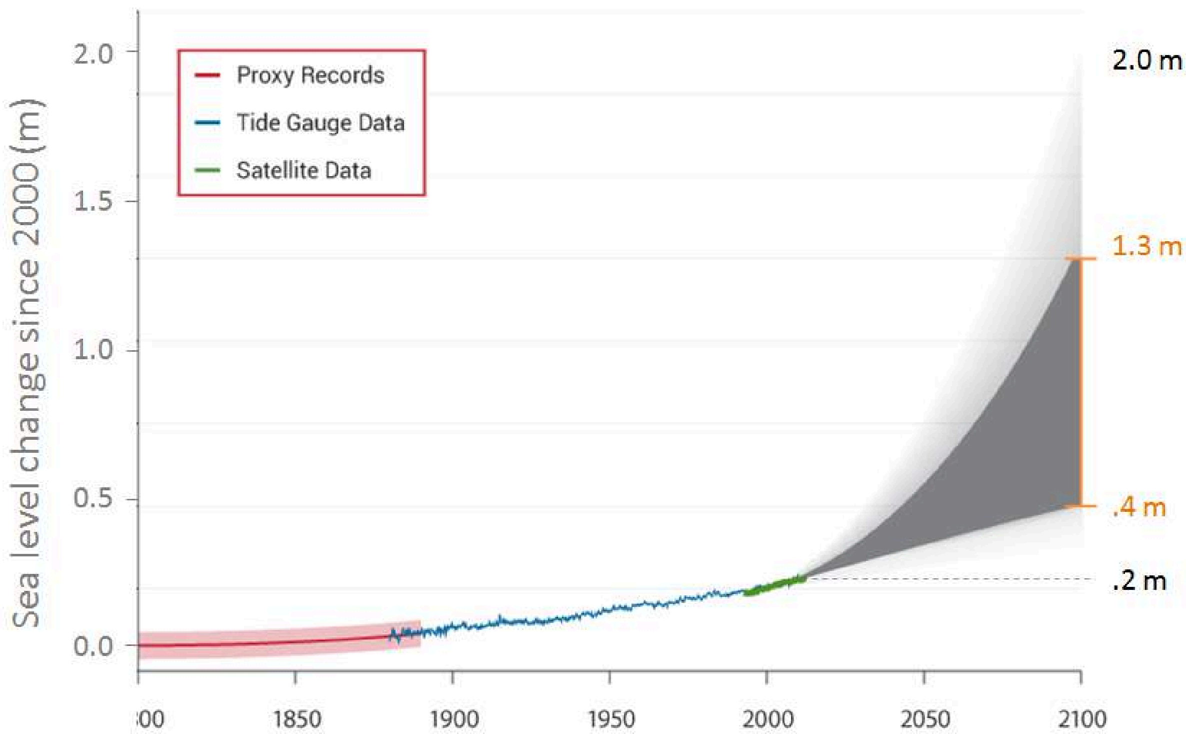


Figure 8.5.5 *Projected sea-level increases to 2100, showing likely range (grey) and possible maximum (Stephen Earle, "Physical Geology", adapted by SE from: http://nca2014.globalchange.gov/report/our-changing-climate/sea-level-rise#intro-section-2 based on data in Parris et al., 2012, NOAA).*

Projections for sea level rise to the end of this century vary widely. This is in large part because we do not know which of the above climate change scenarios (Figure 8.5.4) we will most closely follow, but many are in the range from 0.5 m to 2.0 m. One of the problems in predicting sea level rise is that we do not have a strong understanding of how large ice sheets, such as Greenland and Antarctica, will respond to future warming. Another issue is that the oceans don't respond immediately to warming. For example, with the current amount of warming, we are already committed to a future sea level rise of between 1.3 m and 1.9 m, even if we could stop climate change today. This is because it takes decades to centuries for the existing warming of the atmosphere to be transmitted to depth within the oceans and to exert its full impact on large glaciers. Most of that committed rise would take place over the next century, but some would be delayed longer. And for every decade that the current rates of climate change continue, that number increases by another 0.3 m. In other words, if we don't make changes quickly, by the end of this century we'll be locked into 3 m of future sea level rise. In a 2008 report, the Organization for Economic Co-operation and Development (OECD) estimated that by 2070 approximately 150 million people living in coastal areas could be at risk of flooding due to the combined effects of sea level rise, increased storm intensity, and land subsidence. The assets at risk (buildings, roads, bridges, ports, etc.) are in the order of $35 trillion ($35,000,000,000,000). Countries with the greatest exposure of population to flooding are China, India, Bangladesh, Vietnam, U.S.A., Japan, and Thailand. Some of the major cities at risk include Shanghai, Guangzhou, Mumbai, Kolkata, Dhaka, Ho Chi Minh City, Tokyo, Miami, and New York.

One of the other risks for coastal populations, besides sea level rise, is that climate warming is also associated with an increase in the intensity of tropical storms (e.g., hurricanes or typhoons; see section 8.4), which almost always bring serious flooding from intense rain and storm surges. Some recent examples are New Orleans in 2005 with Hurricane Katrina, and New Jersey and New York in 2012 with Hurricane Sandy. Tropical storms get their energy from the evaporation of warm seawater in tropical regions. In the Atlantic Ocean, this takes place between 8° and 20° N in the summer. Figure 8.5.6 shows the variations in the sea-surface temperature (SST) of the tropical Atlantic Ocean (in blue) versus the amount of power represented by Atlantic hurricanes between 1950 and 2008 (in red). Not only has the overall intensity of Atlantic hurricanes increased with the warming since 1975, but the correlation between hurricanes and sea-surface temperatures is very strong over that time period.

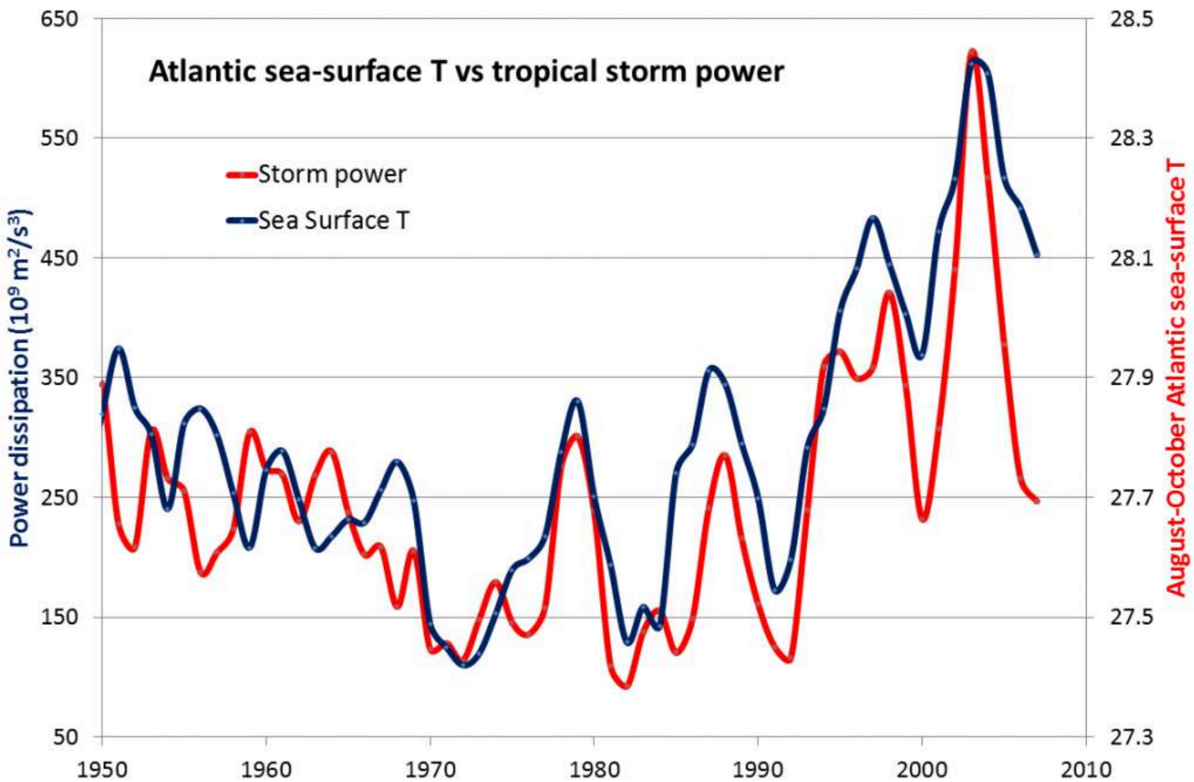


Figure 8.5.6 *Relationship between Atlantic tropical storm cumulative annual intensity and Atlantic sea-surface temperatures (Steven Earle, "Physical Geology", by SE from data at: http://wind.mit.edu/~emanuel/Papers_data_graphics.htm).*

The geographical ranges of diseases and pests, especially those caused or transmitted by insects, have been shown to extend toward temperate regions because of climate change. West Nile virus and Lyme disease are two examples that already directly affect North Americans, while dengue fever could be an issue in the future (dengue became a "nationally notifiable condition" in the United States in 2010). For several weeks in July and August of 2010, a massive heat wave affected western Russia, especially the area southeast of Moscow, and scientists have stated that climate change was a contributing factor. Temperatures soared to over 40°C, as much as 12°C above normal over a wide area, and wildfires raged in many parts of the country. Over 55,000 deaths are attributed to the heat and to respiratory problems associated with the fires. A summary of the impacts of climate change on natural disasters is given in Figure 8.5.7. The major types of disasters related to climate are floods and storms, but the health implications of extreme temperatures are also becoming a great concern. In the decade 1971 to 1980, extreme temperatures were the fifth most common natural disasters; by 2001 to 2010, they were the third most common.

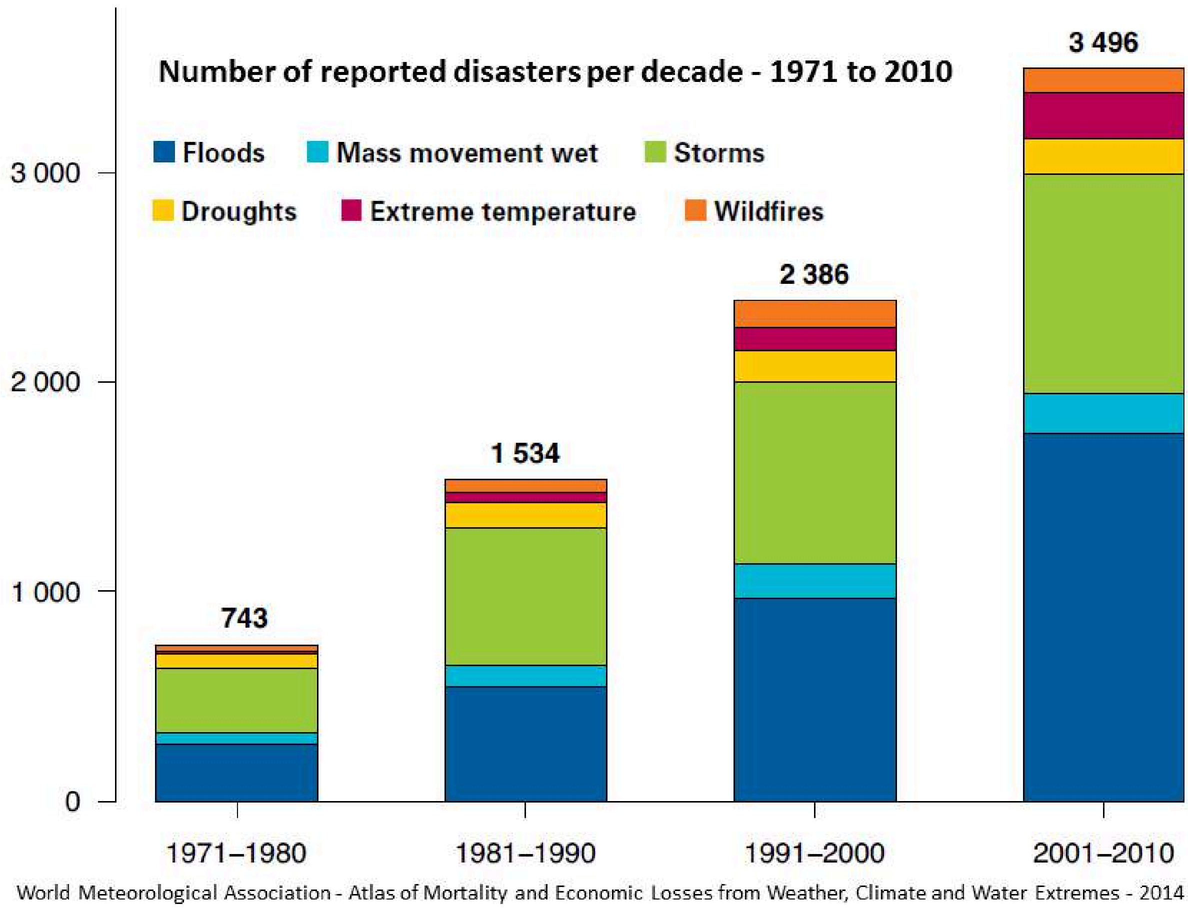


Figure 8.5.7 *Numbers of various types of disasters between 1971 and 2010 (From WMO atlas of mortality and economic Losses from weather, climate and water extremes, 2014).*

CHAPTER 9: OCEAN CIRCULATION

Chapter 9: Ocean Circulation

Learning Objectives

After reading this chapter you should:

- know the major ocean surface currents of the world (i.e. the gyres) and how they are created
- know the features of the Gulf Stream, including the formation of warm and cold core rings
- understand the causes and effects of the Ekman spiral
- understand geostrophic flow, and how it helps keep the gyres flowing even when wind dies down.
- understand why gyre currents are more intense on the western side of the oceans – i.e. western intensification
- know what causes upwelling and downwelling, and the impacts of these events on primary production
- know the locations of some of the major upwelling regions on Earth
- understand the causes and effects of ENSO events
- understand Langmuir cells
- understand the processes that drive thermohaline circulation
- understand how T-S diagrams work
- know the major global sites of deep water formation
- be familiar with the major global water masses
- understand how deep water circulates throughout the world ocean

Ocean waters are constantly in motion, from ocean-scale surface currents, to density-driven vertical turnover, to small rotating eddies. Oceanographers have an array of sophisticated tools to measure ocean currents, but from time to time, fortuitous accidents can also aid our understanding of ocean circulation. A great example is the case of the container ship *Ever Laurel*, which was on its way from Hong Kong to Tacoma, Washington, in January 1992, when 12 containers were washed overboard in a storm in the middle of the Pacific. One of the containers contained over 28,000 plastic bath toys, which were released into the ocean as the container hit the water. Ten months later, the bath toys began washing ashore, first near Sitka, Alaska, then elsewhere along the Alaskan coast, and by 1996, in Washington. Over the next two decades, some toys traveled as far as the Pacific coasts of South America and Australia, while others were found in the Arctic ice, and some even made it through the Arctic into the North Atlantic, washing up in Newfoundland and Scotland (Figure 9.1). There are still a few thousand of the toys floating around in the central North Pacific, and the paths taken by all of these toys have allowed oceanographers to study the movements of large-scale ocean surface currents.

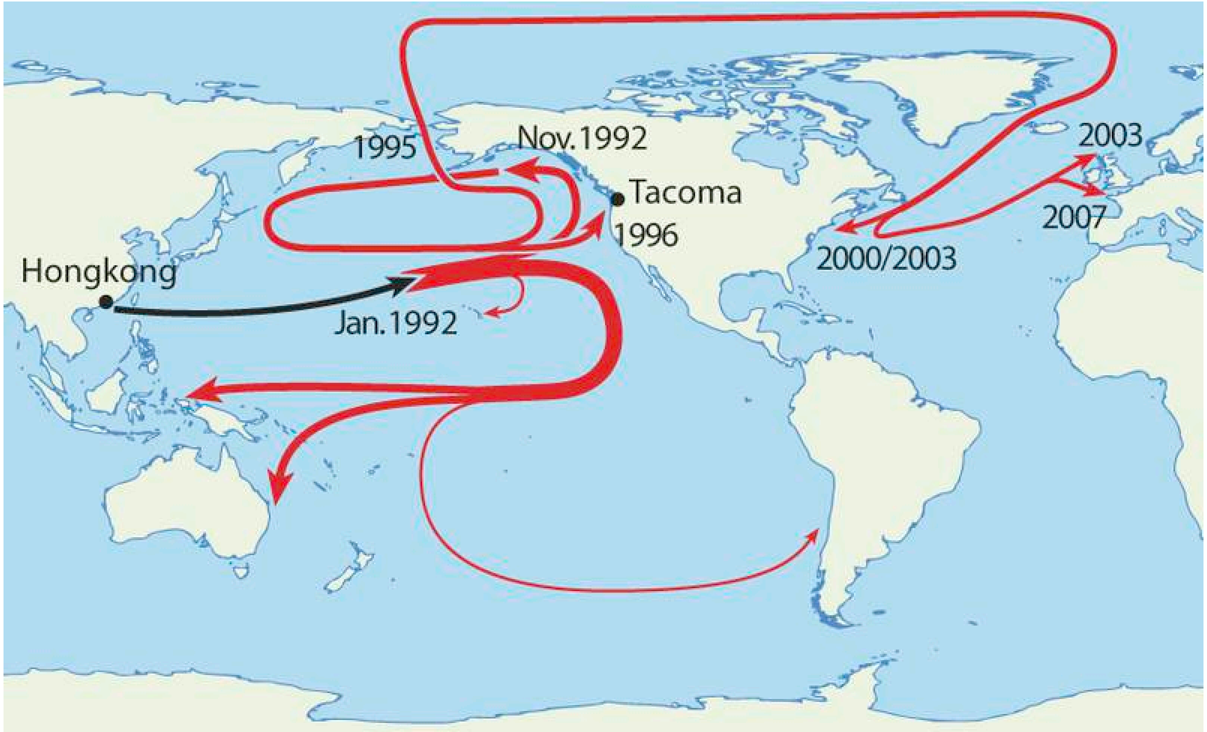


Figure 9.1 *Paths taken by the floating toys lost from the Ever Laurel in 1992, showing the years in which toys were found in different locations (NordNordWest [CC BY-SA 3.0] via Wikimedia Commons).*

9.1 Surface Gyres

In the previous chapter the major wind patterns on Earth were derived. It is these prevailing winds that blow across the water surface to create the major ocean surface currents. However, only about 2% of the wind energy is actually transferred to the water, so a 50 knot wind only creates a 1 knot current. Furthermore, wind-driven surface currents only affect the top 100-200m of water, meaning surface currents only involve about 10% of the world's ocean water. In section 9.8 we will examine deep, thermohaline circulation, which impacts around 90% of the ocean water.

Surface currents generally move in the same direction as the winds that created them. However, because of Coriolis deflection, the surface currents are offset approximately 45^o relative to the wind direction; 45^o to the right in the Northern Hemisphere, and 45^o to the left in the Southern Hemisphere. This creates a general circulation pattern where in both hemispheres, surface currents flow east to west between the equator and 30^o latitude, west to east between 30^o and 60^o, and east to west between 60^o and the poles (Figure 9.1.1).

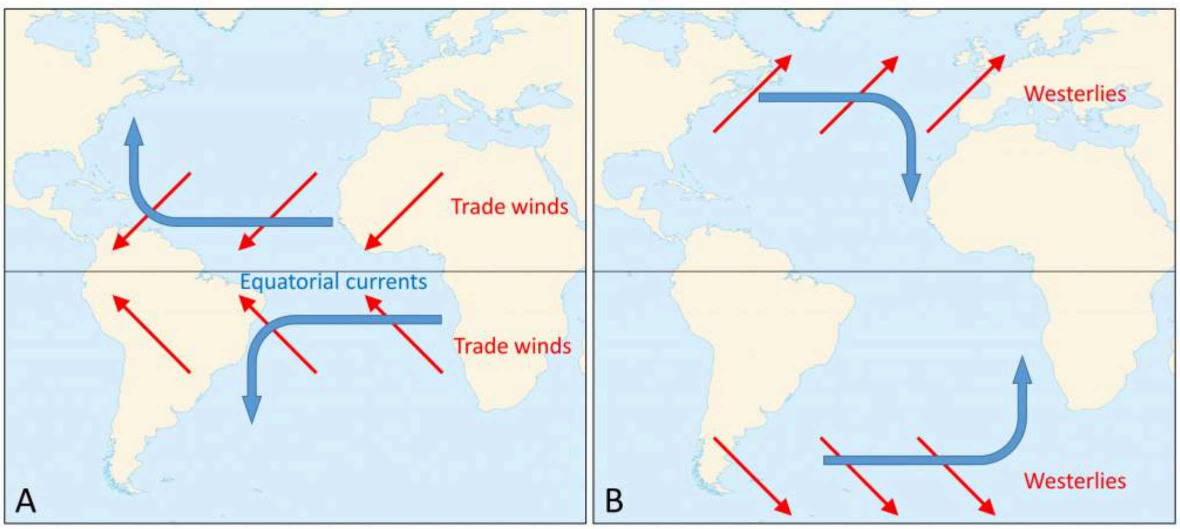


Figure 9.1.1 *Generalized surface currents in the Atlantic Ocean. A) Surface water moving at 45^o relative to the trade winds create the westward flowing equatorial currents. B) Between 30-60^o latitude, the westerlies form eastward flowing surface currents (PW, map by Catrin (Own work, Using GMT) [CC BY-SA 3.0], via Wikimedia Commons).*

The trade winds create the equatorial currents that flow east to west along the equator; the North Equatorial and South Equatorial currents. If there were no continents, these surface currents would travel all the way around the Earth, parallel to the equator. However, the presence of the continents prevents this unimpeded flow. When these equatorial currents reach the continents, they are diverted and deflected away from the equator by the Coriolis Effect; deflection to the right in the Northern Hemisphere and to the left in the Southern Hemisphere. These currents then become western boundary currents; currents that run along the western side of the ocean basin (i.e. the east coasts of the continents). Since these currents come from the equator, they are warm water currents, bringing warm water to the higher latitudes and distributing heat throughout the ocean.

At the same time, between 30-60^o latitude the westerlies move surface water towards the east. The Coriolis Effect and the presence of the continents deflect the currents towards the equator, creating eastern boundary currents (on the eastern side of the ocean basins). These currents come from high latitude areas, so they deliver cold water to the lower latitudes. Together, these currents combine to create large-scale circular patterns of surface circulation called **gyres**. In the Northern Hemisphere the gyres rotate to the right (clockwise), while in the Southern Hemisphere the gyres rotate to the left (counterclockwise).

There are five major gyres in the oceans; the North Atlantic, South Atlantic, North Pacific, South Pacific, and Indian (Figure 9.1.2). The North Pacific gyre is composed of the North Equatorial Current on its southern boundary, which turns into the Kuroshio Current (a.k.a. the Japan Current) bringing warm water north towards Japan. The Kuroshio flows into the North Pacific Current which moves east towards North America, where it becomes the California Current to complete the gyre. The North Atlantic gyre is formed by the North Equatorial Current flowing into the Gulf Stream along the east coast of the United States. The Gulf Stream merges into the North Atlantic Current to move water towards Europe, which then becomes the Canary Current as it moves south to join the North Equatorial Current.

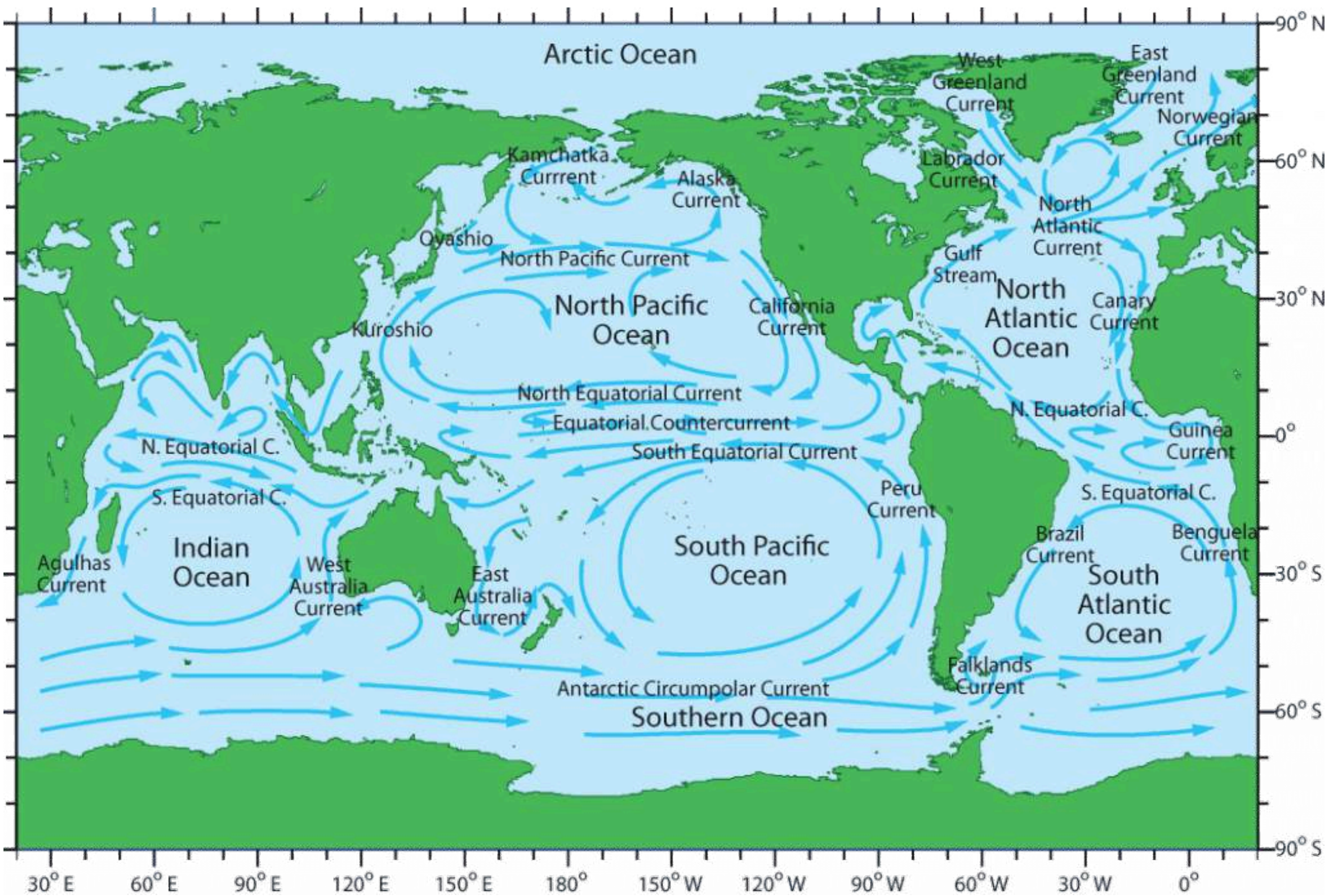


Figure 9.1.2 *The major global surface currents (http://www.seos-project.eu/modules/oceancurrents/oceancurrents-c02-p01.html, Source: NOC; CC BY-NC-SA 2.0).*

Near Antarctica the circulation is somewhat different. Because there is little in the way of continental land masses between 50-60^{o} south, the surface current created by the westerly winds can make its way completely around the Earth, creating the Antarctic Circumpolar Current (ACC) or West Wind Drift (WWD) that flows from west to east (Figure 9.1.2). The Antarctic Circumpolar Current is the only current that connects all of the major ocean basins, and in terms of the amount of water that it transports, it is the largest surface current on Earth. Above 60^{o} latitude the prevailing winds are the polar easterlies, which create a current flowing from east to west along the edge of the Antarctic continent, the East Wind Drift or the Antarctic Coastal Current.

The Antarctic Circumpolar Current creates the southern boundary for all of the Southern Hemisphere gyres. In the South Pacific gyre the ACC becomes the Peru Current (also known as the Humboldt Current) moving up the west coast of South America, before joining the South Equatorial Current. The South Equatorial Current flows southwards as the East Australia Current, before completing the gyre with the ACC. The South Atlantic gyre is composed of the South Equatorial Current, the Brazil Current, the ACC, and the Benguela Current. Finally, the currents making up the Indian gyre are the ACC, the West Australia Current, the South Equatorial Current, and the Agulhas Current.

Not all of the equatorial water that is moved westward by the trade winds and reaches the continents gets transported to higher latitudes in the gyres, because the Coriolis Effect is weakest along the equator. Instead, some of the water piles up along the western edge of the ocean, and then flows eastward due to gravity, creating narrow Equatorial Countercurrents between the North and South Equatorial Currents (Figure 9.1.2). Some of this water also moves east as equatorial undercurrents that flow at depths between 50-200 m, underneath the Equatorial Currents. These undercurrents are called the Lomonosov Current in the Atlantic, and the Cromwell Current in the Pacific.

9.2 The Gulf Stream

The primary surface current along the east coast of the United States is the Gulf Stream, which was first mapped by Benjamin Franklin in the 18th century (Figure 9.2.1). As a strong, fast current, it reduced the sailing time for ships traveling from the United States back to Europe, so sailors would use thermometers to locate its warm water and stay within the current.

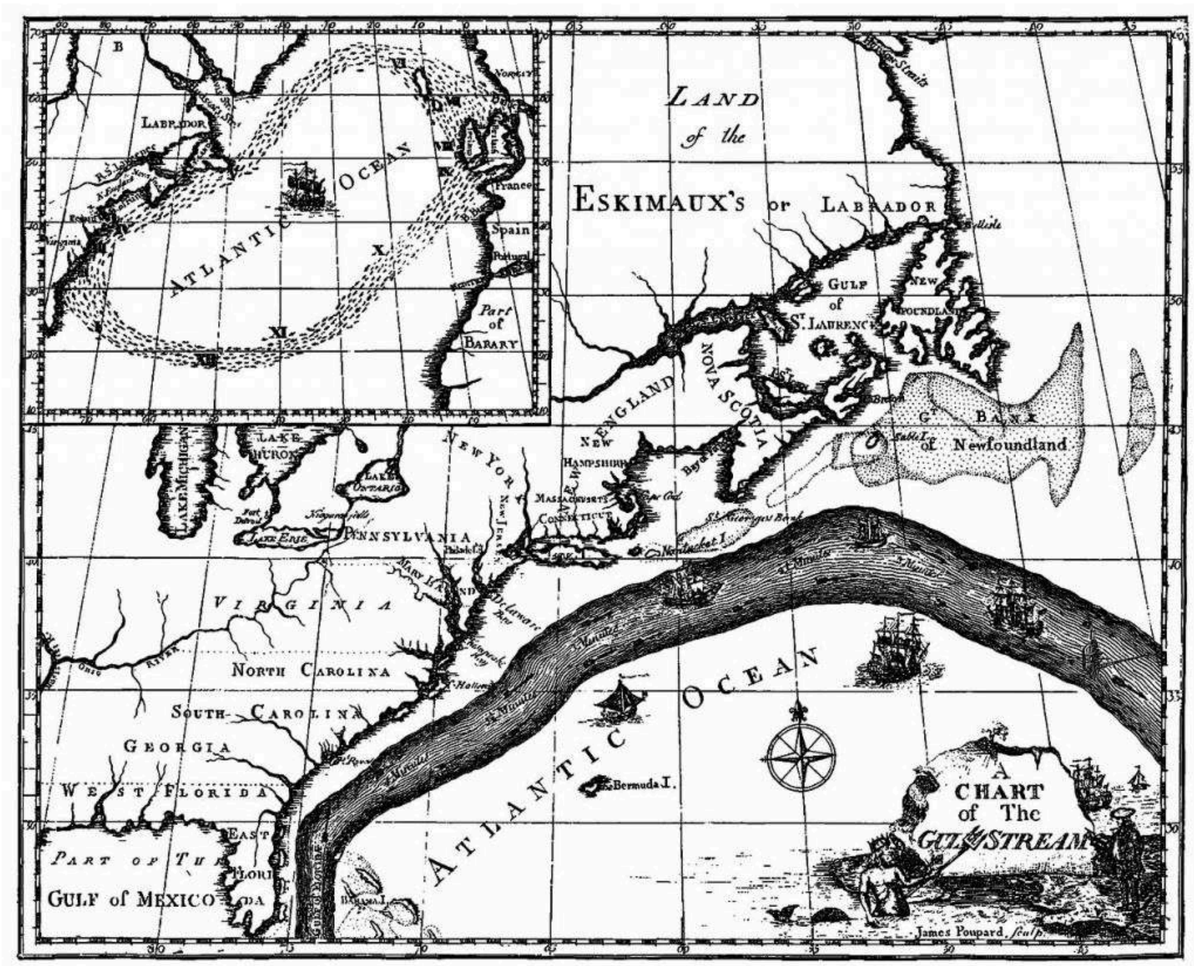


Figure 9.2.1 *Benjamin Franklin's original map of the Gulf Stream (Public domain, via Wikimedia Commons).*

The Gulf Stream is formed from the convergence of the North Atlantic Equatorial Current bringing tropical water from the east, and the Florida Current that brings warm water from the Gulf of Mexico. The Gulf Stream takes this warm water and transports it northwards along the U.S. east coast (Figure 9.2.2). As a western boundary current, the Gulf Stream experiences western intensification (section 9.4), making the current narrow (50-100 km wide), deep (to depths of 1.5 km) and fast. With an average speed of 6.4 km/hr, and a maximum speed of about 9 km/hr, it is the fastest current in the world ocean. It also transports huge amounts of water, more than 100 times greater than the combined flow of all of the rivers on Earth.

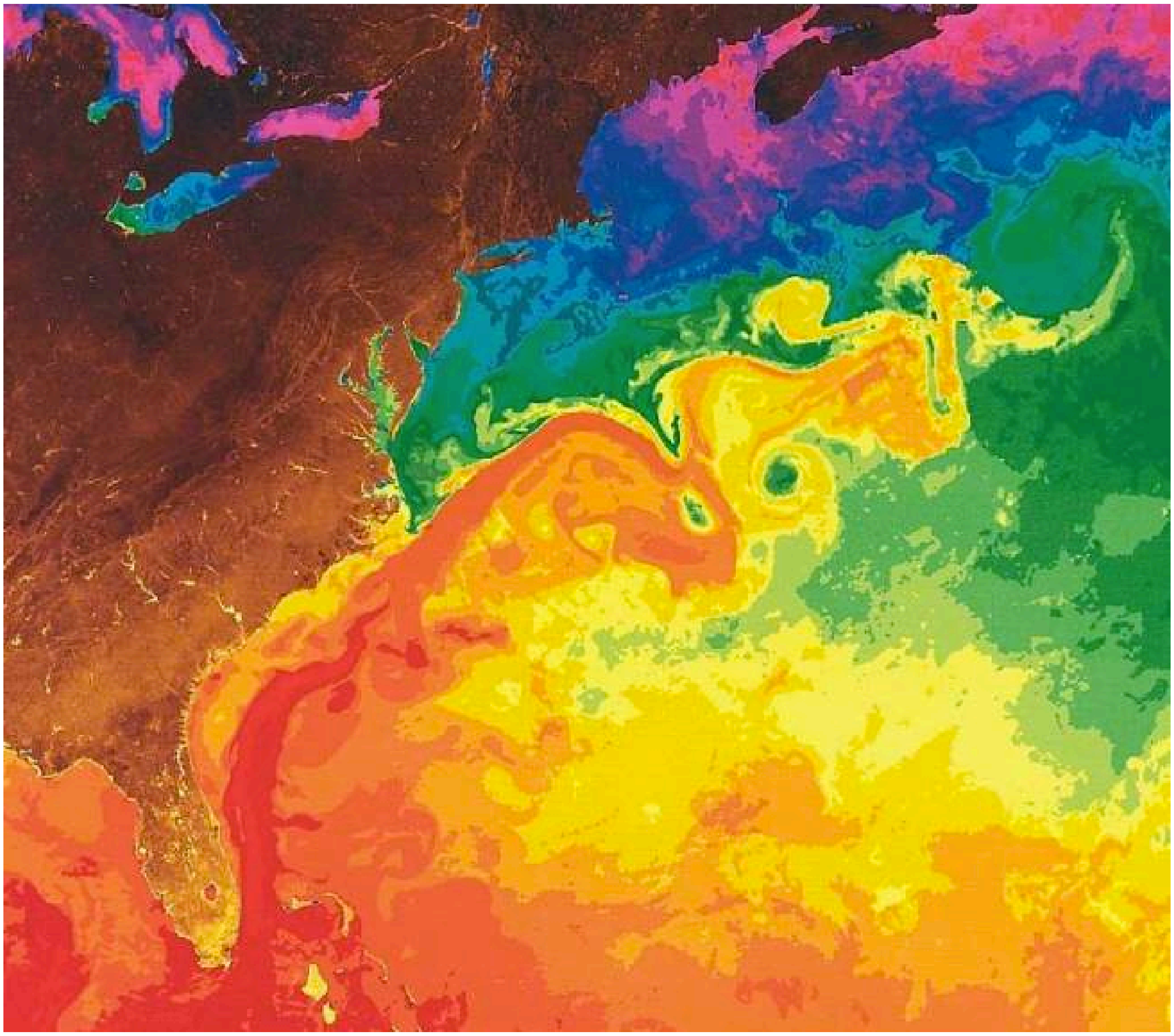


Figure 9.2.2 *Sea surface temperature map illustrating the Gulf Stream. Warmer water is shown in red, colder water in blue and violet. Meanders and eddies are visible where the current moves towards the northeast (NASA, Public Domain via Wikimedia Commons).*

As the Gulf Stream approaches Canada, the current becomes wider and slower as the flow dissipates and it encounters the cold Labrador Current moving in from the north. At this point, the current begins to meander, or change from a fast, straight flow to a slower, looping current (Figure 9.2.2). Often these meanders loop so much that they pinch off and form large rotating water masses called **rings** or **eddies**, that separate from the Gulf Stream. If an eddy pinches off from the north side of the Gulf Stream, it entraps a mass of warm water and moves it north into the surrounding cold water of the North Atlantic. These **warm core rings** are shallow, bowl-shaped water masses about 1 km deep, and about 100 km across, that rotate clockwise as they carry warm water in to the North Atlantic (Figure 9.2.3). If the meanders pinch off at the southern boundary of the Gulf Stream, they form **cold core rings** that rotate counterclockwise and move to the south. Cold core rings are cone-shaped water masses extending down to over 3.5 km deep, and may be over 500 km wide at the surface.

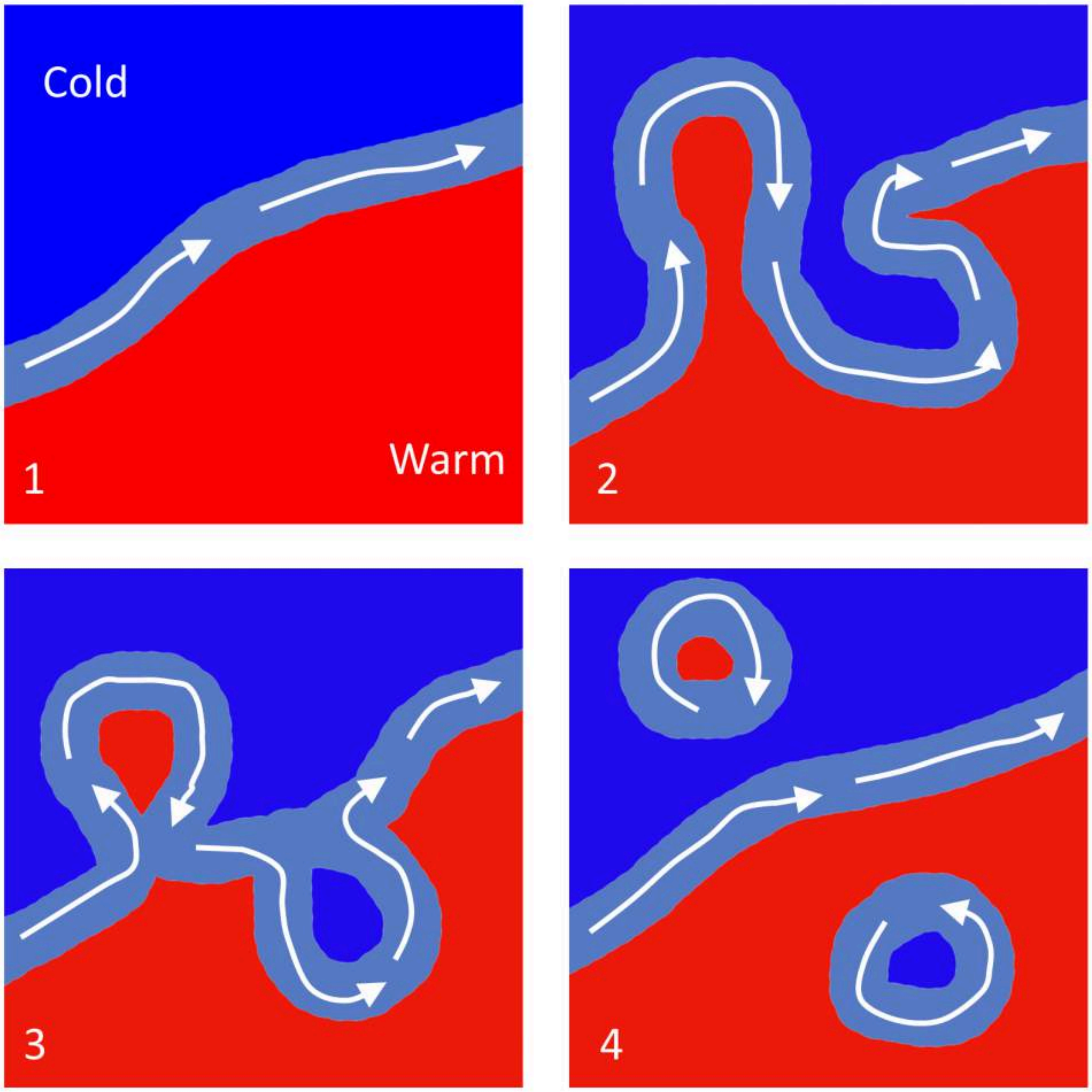


Figure 9.2.3 *Formation of warm and cold core rings from meanders in the Gulf Stream. As the Gulf Stream flows to the northeast (1), it starts to meander as it slows, forming warm or cold water extensions on either side of the current (2). If the meanders pinch off the extensions, they trap pockets of warm or cold water (3), that can separate from the Gulf Stream and travel north or south. Warm core rings rotate clockwise, while cold core rings rotate counterclockwise (4) (PW).*

After the Gulf Stream meets the cold Labrador Current, it joins the North Atlantic Current, which transports the warm water towards Europe, where it moderates the European climate. It is estimated that Northern Europe is up to 9^{o} C warmer than expected because of the Gulf Stream, and the warm water helps to keep many northern European ports ice-free in the winter.

In the east, the Gulf Stream merges into the Sargasso Sea, which is the area of the ocean within the rotation center of the North Atlantic gyre. The Sargasso Sea gets its name from the large floating mats of the marine algae *Sargassum* that are abundant on the surface (Figure 9.2.4). These *Sargassum* mats may play an important role in the early life stages of sea turtles, who may live and feed within the algae for many years before reaching adulthood.

Figure 9.2.4 *Map of the Sargasso Sea (left), and closeup photo of Sargassum algae (right) (Map by USFWS; photo by Bogdan Giușcă; Public Domain via Wikimedia Commons).*

9.3 The Ekman Spiral and Geostrophic Flow

Ekman spiral

Winds blowing over the ocean are ultimately what create the surface currents. However, not all of the water moved by the surface currents is transported in the same direction. The Coriolis Effect causes the surface water to move in a direction about 45 degrees offset from the wind direction, with the deflection to the right of the wind in the Northern Hemisphere and to the left in the Southern Hemisphere. The frictional movement of the topmost layer of water sets in motion the layer directly underneath it, which then sets in motion the next layer under that, and so on as the water gets deeper. Some energy is lost in each transition, so each successive layer of water will not move as far as the layer above it; in other words, there is decreasing energy with increasing depth. But at the same time, the Coriolis Effect deflects each layer relative to the layer above it (again, to the right in the Northern Hemisphere and to the left in the Southern Hemisphere). The movement of the successive layers therefore creates a spiraling pattern of water motion called the **Ekman spiral**, which usually penetrates to about 100 m deep before the motion ceases. If the magnitudes and directions of the movements of each layer are added together, the result is that the net movement of the upper 100 m of the water column is 90^{o} relative to the original wind direction (90^{o} to the right of the wind in the Northern Hemisphere, and 90^{o} to the left in the Southern Hemisphere). This net water movement is called **Ekman transport** (Figure 9.3.1).

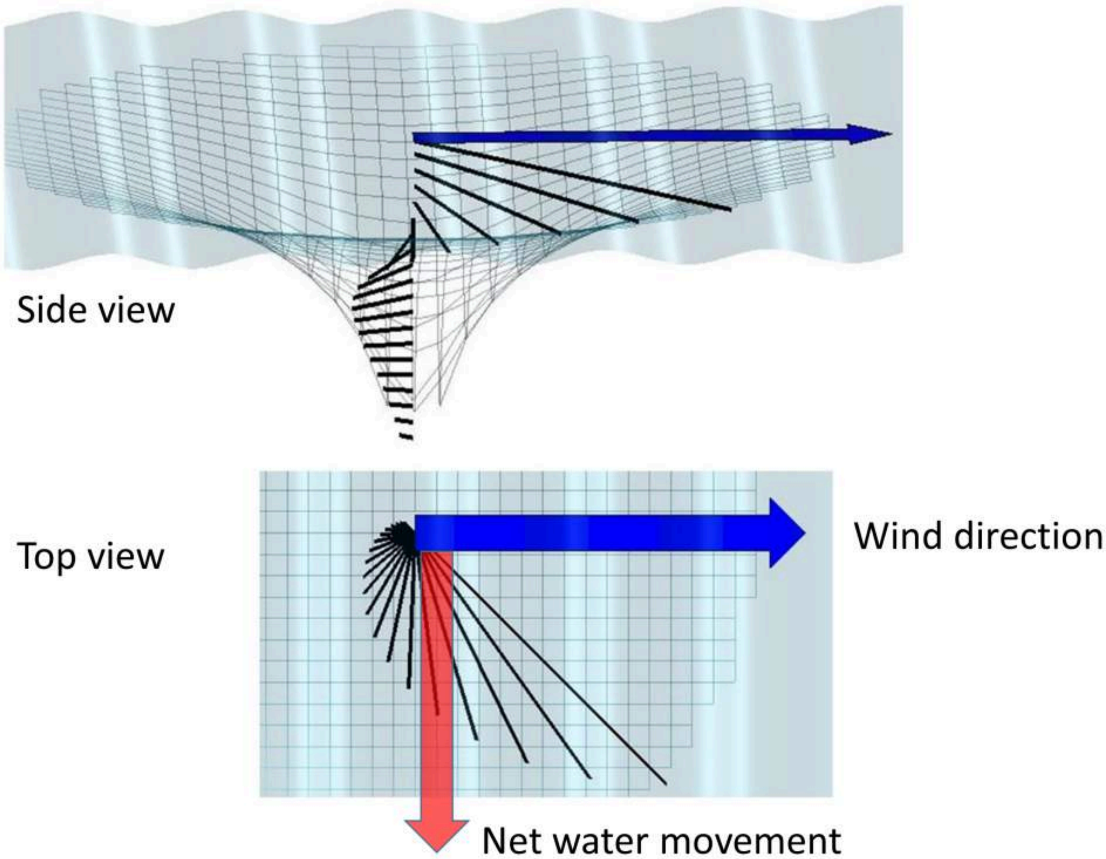


Figure 9.3.1 *The Ekman spiral, shown for the Northern Hemisphere. Wind blowing over the water (blue arrow) creates a surface current 45^{o} offset from the wind. Each successive layer of water is moved and deflected by the layer above, creating a spiraling pattern of water movement that diminishes with depth. The net movement of the water within the spiral is 90^{o} relative to the wind direction (red arrow) (Modified by PW from Ekman spiral By Schlusenbach (unbekannt) [Public domain], via Wikimedia Commons).*

Geostrophic flow

Gyre rotation is dependent on the wind and the Coriolis Effect impacting the surface currents (see section 9.1). But the rotation is also affected by movement below the surface due to Ekman transport. In the Northern Hemisphere, as the gyre rotates clockwise the net movement of the Ekman transport is 90^{o} to the right of the wind; in other words, towards the center of the gyre. The Ekman transport piles up water in the center of the gyre, making the water level higher in the gyre center than on the edges of the gyres. This pile of water then has a tendency to flow back "downhill" due to gravity. As the water flows "downhill" away from the gyre center, it is deflected to the right by the Coriolis force. This results in a clockwise current around the central "hill" called **geostrophic flow**, which moves in the same direction as the rotating gyre. Water is thus pushed into the "hill" by Ekman transport, and away from the "hill" by gravity, with both flows modified by the Coriolis Effect to create the rotation. As with the gyres, geostrophic flow is clockwise in the Northern Hemisphere, and counterclockwise in the Southern Hemisphere.

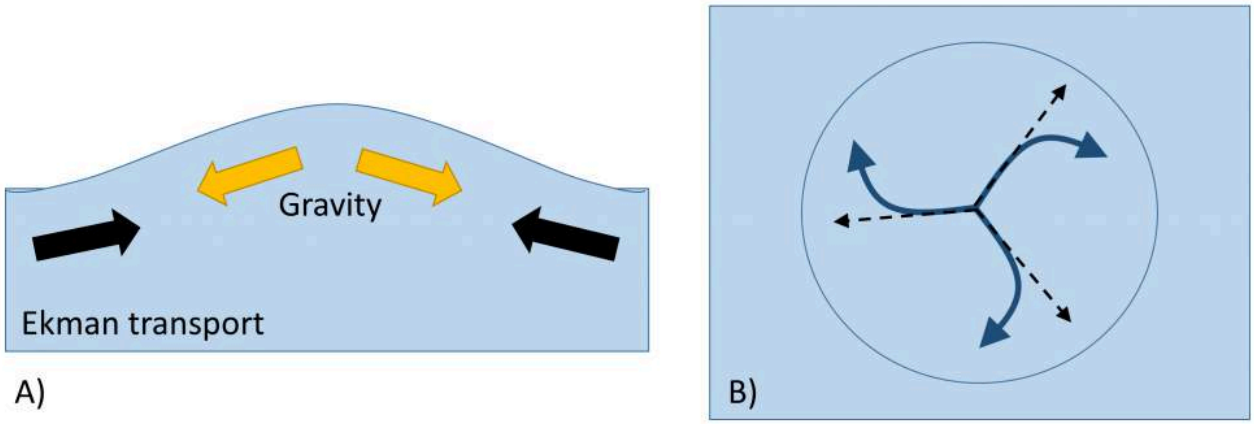


Figure 9.3.2 *Geostrophic flow in the Northern Hemisphere. A) Ekman transport moves water into the middle of the gyre, where it "piles up." Gravity causes the water to flow back "downhill." B) Viewed from above, as the water in the center flows "downhill" (dotted arrows) the Coriolis force deflects the movement to the right (solid arrows), causing the system to rotate clockwise (PW).*

Most major surface currents are a combination of wind-driven and geostrophic currents. Since winds can be variable, geostrophic flow ensure that the gyre currents keep moving at a fairly constant rate even when the wind dies down. The larger the area, and the higher the slope, the longer the geostrophic flow will continue to move and power the gyre after the wind subsides.

9.4 Western Intensification

In both hemispheres, the currents making up the western side of the gyre are much more intense than the currents on the eastern side. In other words, the currents off of the east coast of the continents are more intense than currents off of the west coast of the continents. This phenomenon is known as western intensification, and once again it is due to the Coriolis Effect.

As discussed in section 8.2, the Coriolis Effect is a result of the fact that different latitudes of the Earth are rotating at different speeds, and the apparent path taken by an object is deflected as it moves between areas of different rotation speeds. The greater the change in rotation speed, the stronger the Coriolis force. At the poles, the speed of rotation is 0 km/hr. The speed increases to about 800 km/hr at 60^{o} latitude, 1400 km/hr at 30^{o} latitude, and 1600 km/hr at the equator. Therefore there is an 800 km/hr difference between 60^{o} and 90^{o} latitude, while there is only a 200 km/hr difference between the equator and 30^{o}. Thus the speed of Earth's rotation changes more quickly with latitude near the poles than at the equator, making the Coriolis force strongest near the poles and weakest at the equator.

The high latitude surface currents of the major gyres experience a strong Coriolis force due to their proximity to the poles. As the currents move eastward, the strong Coriolis force begins to deflect the currents towards the equator relatively early. The currents on the eastern side of the gyre are therefore spread out over a wide area as they move towards the equator (Figure 9.4.1). Near the equator, the westward flowing currents experience a much weaker Coriolis force, so their deflection does not happen until the current is all the way over to the western side of the ocean basin. These western currents must therefore move through a much narrower area (Figure 9.4.1). This imbalance means that the center of rotation of the gyre is not in the center of the ocean basins, but it closer to the western side of the gyre.

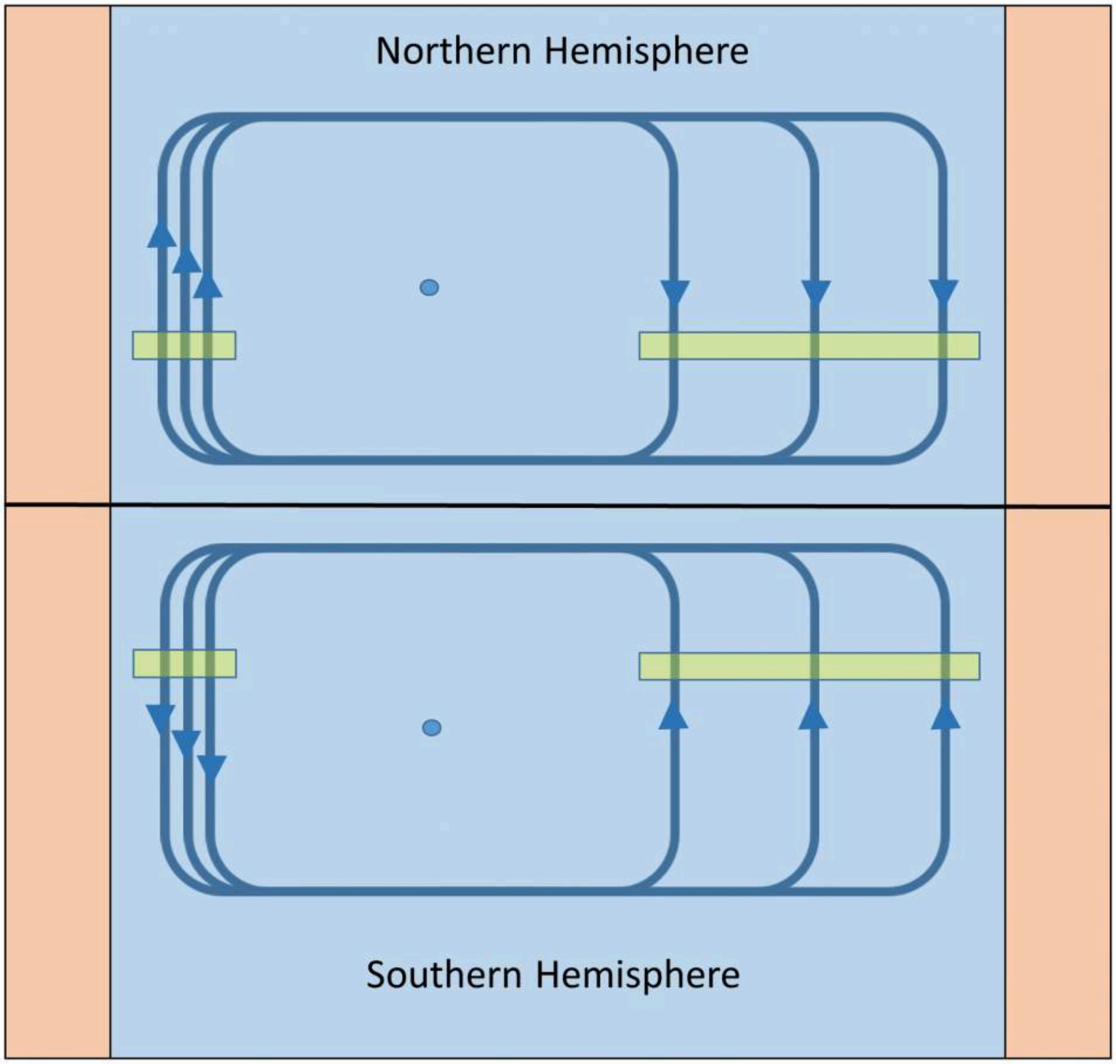


Figure 9.4.1 *Western intensification. In both hemispheres, currents on the western side of the gyres travel through a much narrower area than the currents on the eastern side (yellow rectangles). To move the same volume of water through each side, western boundary currents are faster, deeper, and narrower than eastern boundary currents. The center of rotation of the gyre is also closer to the western side of the gyre (blue dots) (PW).*

The same volume of water must pass through both the east and west sides of the gyre. In the western gyre currents, that volume is passing through a narrower area, so the current must travel faster in order to transport the same amount of water in the same amount of time. On the eastern side of the gyre the current is much wider, so the flow is slower. A simple analogy is the water flowing from a garden hose. You can make the water flow from the hose much faster and more strongly by covering part of the opening with your thumb. The same amount of water is exiting the hose whether the opening is covered or uncovered, but to get that water through the covered opening the flow has to be much faster and stronger. In the same way, western boundary currents are not only faster, but also deeper than eastern boundary currents, as they move the same volume through a narrower space. For example, the Kuroshio Current in the western Pacific is around 15 times faster, 20 times narrower, and 5 times deeper than the California Current in the eastern Pacific.

9.5 Currents, Upwelling and Downwelling

The movement of surface currents also plays a role in the vertical movements of deeper water, mixing the upper water column. **Upwelling** is the process that brings deeper water to the surface, and its major significance is that it brings nutrient-rich deep water to the nutrient-deprived surface, stimulating primary production (see section 7.3). **Downwelling** is where surface water is forced downwards, where it may deliver oxygen to deeper water. Downwelling leads to reduced productivity, as it extends the depth of the nutrient-limited layer.

Upwelling occurs where surface currents are diverging, or moving away from each other. As the surface waters diverge, deeper water must be brought to the surface to replace it, creating upwelling zones. The upwelled water is cold and rich in nutrients, leading to high productivity. Many of the most productive regions on Earth are found in upwelling zones. In the equatorial Pacific, the trade winds blow the North and South Equatorial Currents towards the west, while Ekman transport causes the upper layers to move to the north and south in their respective hemispheres. This creates a divergence zone, and a region of upwelling and high productivity (Figure 9.5.1).

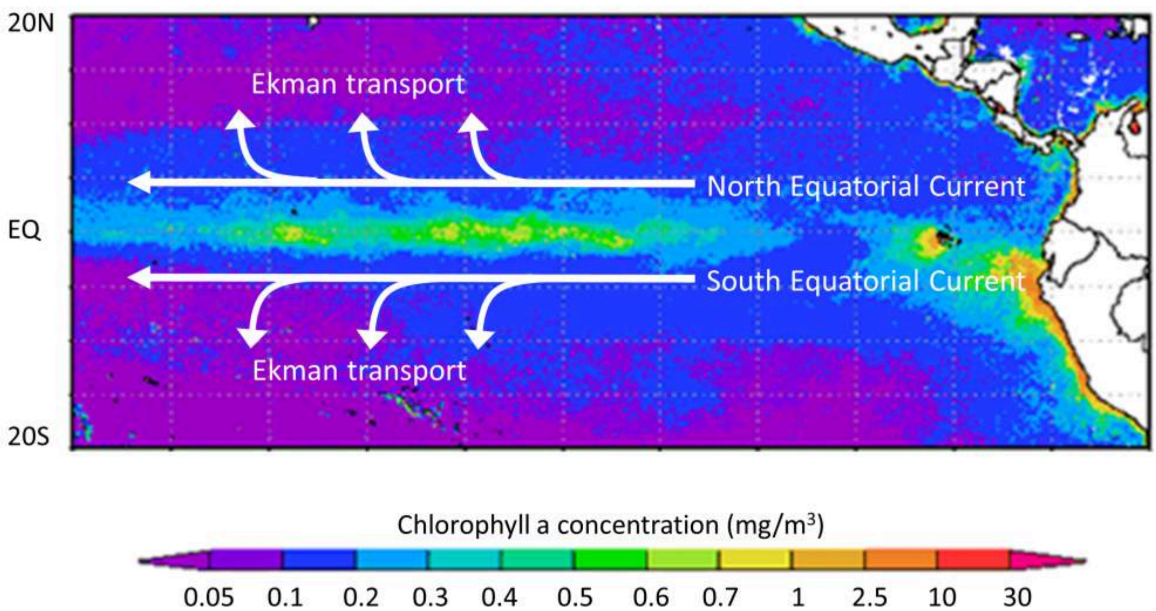


Figure 9.5.1 *Equatorial upwelling and increased productivity as a result of divergence between the north and south equatorial currents (Modified by PW from image by NASA [Public domain], via Wikimedia Commons).*

A similar process occurs near the Antarctic continent, creating one of the most productive regions on Earth, the Antarctic divergence. In this case, the West Wind Drift (Antarctic Circumpolar Current) is flowing parallel to, but in the opposite direction of the East Wind Drift. With both currents occurring in the Southern Hemisphere, Ekman transport will be to the left, so the eastward-flowing West Wind Drift water will be transported to the north, and the westward-flowing East Wind Drift water will be transported to the south, creating a highly productive divergence zone (Figure 9.5.2).

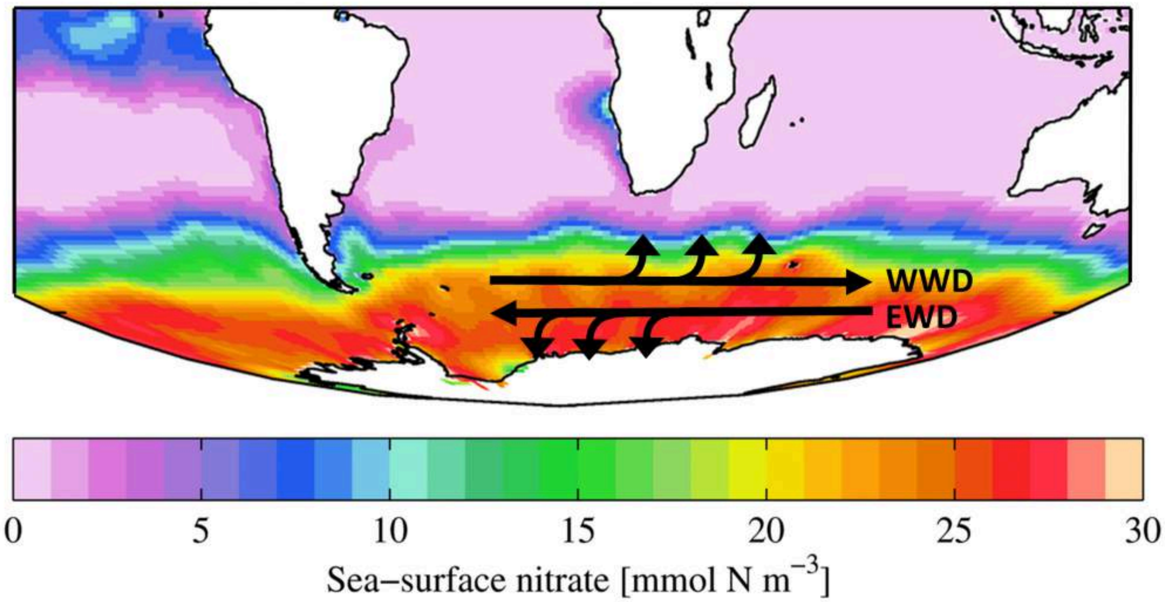


Figure 9.5.2 *High nutrient levels in the Antarctic divergence zone, as a result of the diverging West Wind Drift and East Wind Drift currents creating strong upwelling (Modified by PW from Plumbago (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

Downwelling occurs where surface currents converge. The converging water has nowhere to go but down, so the surface water sinks. Since surface water is usually low in nutrients, downwelling leads to low productivity zones. An example of a downwelling region is off of the Labrador coast in Canada, where the Gulf Stream, Labrador, and East Greenland Currents converge.

Coastal Upwelling

Upwelling and downwelling also occur along coasts, when winds move water towards or away from the coastline. Surface water moving away from land leads to upwelling, while downwelling occurs when surface water moves towards the land. Historically, some of the most productive commercial fishing grounds have been associated with coastal upwelling. Along the coast of California, the local prevailing winds blow towards the south. Ekman transport moves the surface layer 90^{o} to the right of the wind, meaning the net Ekman transport is in an offshore direction. The water displaced near the coast is replaced by cold, nutrient-rich deeper water that is brought to the surface through upwelling, leading to high productivity (Figure 9.5.3).

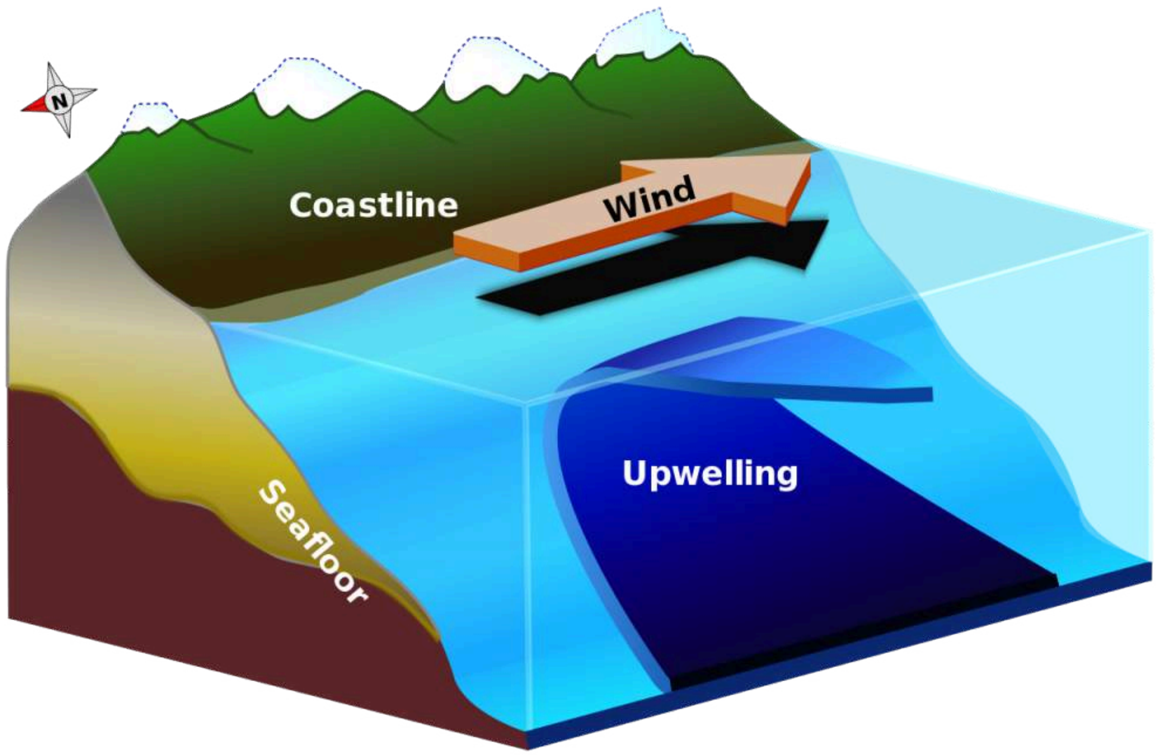


Figure 9.5.3 *Coastal upwelling. As wind blows along a coastline, Ekman transport moves the surface layer in a direction 90^o to the wind. In this figure, it is to the right of the wind, indicating a Northern Hemisphere location. As surface water moves offshore it it replaced by upwelled deeper water (By Lichtspiel [Public domain], via Wikimedia Commons).*

The same process happens off of the coast of Peru, which for a long time had the world's largest commercial fishery. Winds along the Peruvian coast blow towards the north, and since Peru is in the Southern Hemisphere, the Ekman transport is 90^o to the left of the wind, which causes the surface water to move offshore and leads to upwelling and productivity. In any coastal upwelling location, if the winds reverse, surface water moves towards the shore and downwelling is the result.

Upwelling can also occur due to geological features of the ocean floor. For example, as deep water currents encounter seamounts or other raised features, the water is forced upwards, bringing nutrient-rich water to the surface. This helps explain why productivity is often high in the water over seamounts.

9.6 El Niño and La Niña

As we saw in the previous section, coastal upwelling off of Peru makes that region one of the world's most productive fishing grounds. But every so often, the conditions in the region are very different. Every few years, the cold, nutrient-rich water is replaced by unusually warm water that is low in nutrients, leading to a decline in fish populations. In addition, the normally dry areas receive lots of rain. Because this phenomenon occurs in the northern winter close to Christmas, it is called **El Niño** (the child). More formally, the event is referred to as **El Niño-Southern Oscillation (ENSO)**. The Southern Oscillation portion refers to the fluctuating atmospheric conditions that lead to the localized ocean warming of El Niño. While the exact reasons for the oscillation events are unclear, it is easier to understand how they lead to an El Niño.

Under normal conditions in the equatorial Pacific, the trade winds blow towards the west, moving large amounts of warm surface water towards the western Pacific around Southeast Asia. As the surface water moves west, it is replaced by cold, nutrient-rich deep water through upwelling (Figure 9.6.1). The coastal upwelling leads to a shallow thermocline in the eastern Pacific. In terms of atmospheric conditions, the trade winds are part of a convection cell called the Walker Cell. There is low pressure over the western Pacific, leading to rising moist air and significant precipitation in the region. In the eastern Pacific near South America, there is high pressure, leading to drier conditions (Figure 9.6.1).

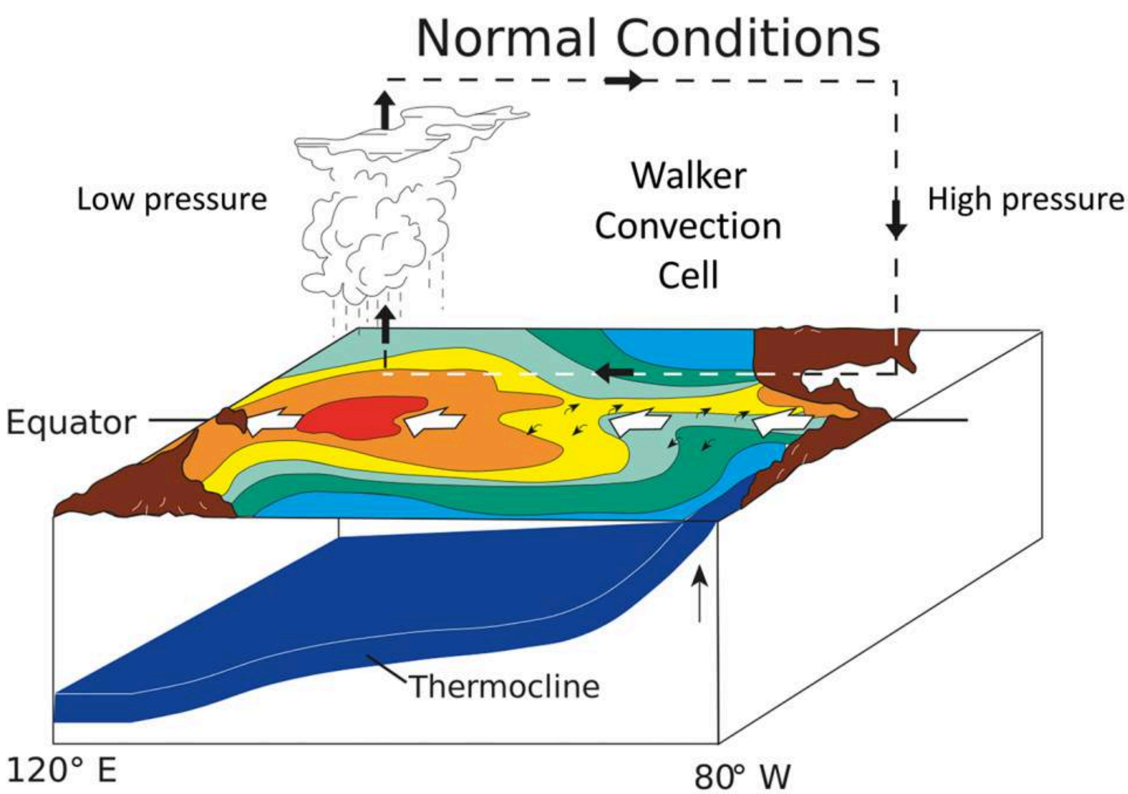


Figure 9.6.1 *Normal conditions in the equatorial Pacific. Low pressure in the western Pacific and high pressure in the eastern Pacific cause the trade winds to move surface water to the west, leading to upwelling and a shallow thermocline near South America (Modified by PW from Fred the Oyster (Own work) [Public domain], via Wikimedia Commons).*

During an El Niño-Southern Oscillation, the high pressure system over the eastern Pacific diminishes, so the trade winds are weakened, or in extreme cases will even reverse. When this happens, warm surface water begins to flow east across the Pacific towards South America (Figure 9.6.2), warming the coastal South American water by up to 8^{o} C in strong ENSO years. This influx of low density warm water deepens the thermocline and prevents upwelling, which dramatically reduces productivity and can devastate populations of fish and other marine life.

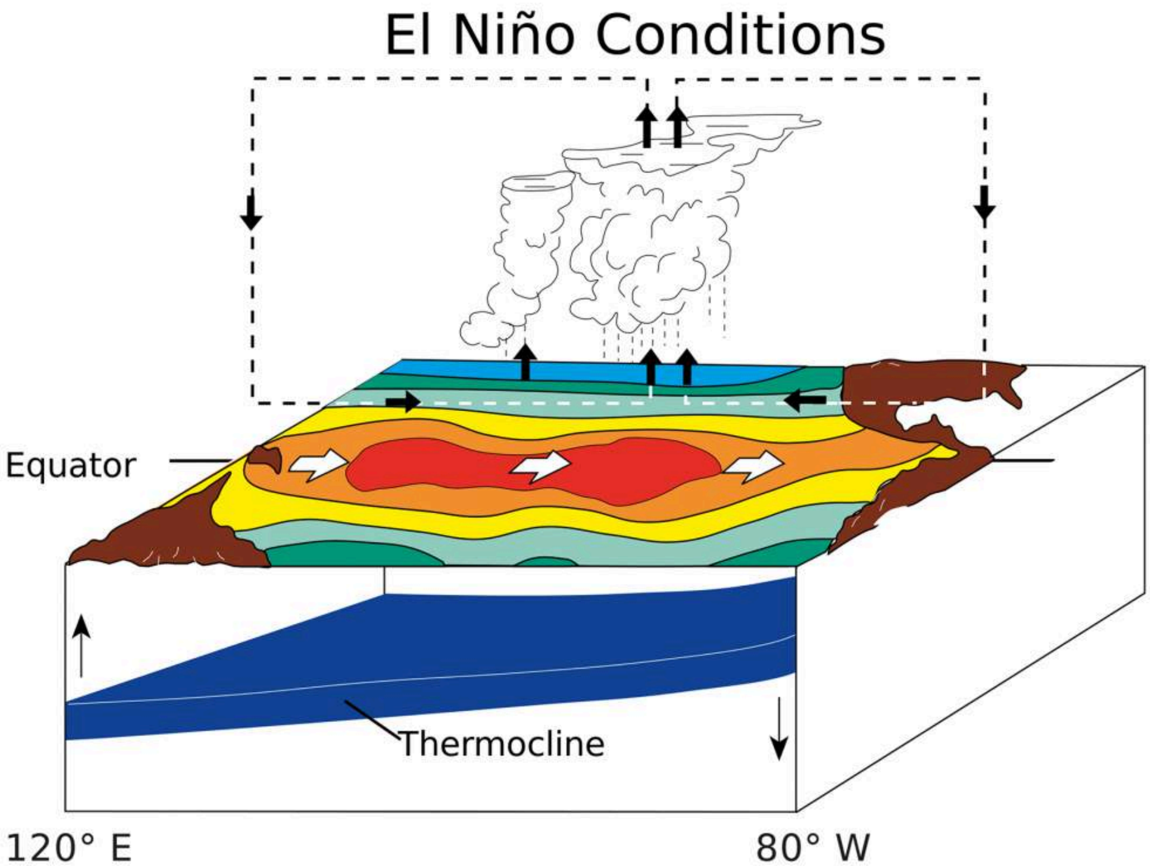


Figure 9.6.2 *El Niño conditions in the equatorial Pacific. Weakening or reversal of the trade winds transport warm surface water eastward towards South America, disrupting coastal upwelling (Public domain, via Wikimedia Commons).*

In the atmosphere, the low pressure system in the western Pacific is replaced by high pressure, bringing dry or even drought conditions to Southeast Asia and Australia. The low pressure system moves east across the Pacific, potentially reaching as far as South America in strong El Niño years. The low pressure over the eastern Pacific brings lots of rain and flooding to South America (Figure 9.6.2). But the effects of El Niño are not just limited to the Pacific; it can influence weather patterns throughout the globe (see box below).

Because the Southern Oscillation is a cyclic pattern, the eastern Pacific is not subject just to unusually warm conditions. There are also periods of abnormally cold water in the region known as **La Niña** events. During a La Niña the trade winds are unusually strong, leading to increased upwelling and transport of deep, cold water to the surface (Figure 9.6.3). The effects of a La Niña are essentially the opposite of an El Niño, bringing cooler and wetter conditions to the northwestern United States and Canada, while the southeastern US receives below-average precipitation. Monsoon seasons in Asia are drier during El Niños but wetter during La Niña events.

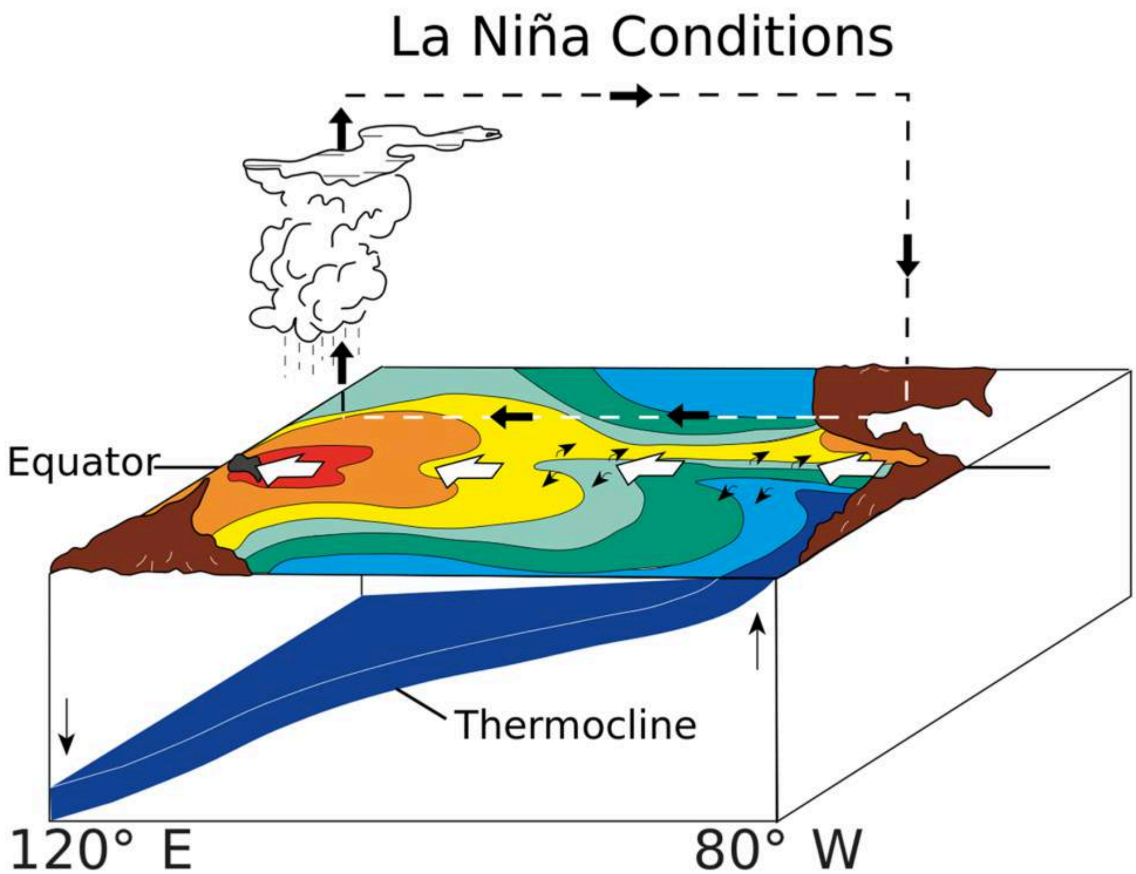


Figure 9.6.3 La Niña conditions. Stronger trade winds promote more intense upwelling in the eastern Pacific, leading to cooler than usual water temperatures (Public domain, via Wikimedia Commons).

El Niño and La Niña events alternate, although the presence of one does not always mean the other will automatically follow. El Niños occur roughly every 2-7 years, and each event may last from a few months to a year or more. Although we do not understand exactly why or when the ENSO events will occur, we can anticipate their arrival by monitoring a number of ocean and atmospheric phenomena that make up the Multivariate ENSO Index (MEI). Examination of the MEI over time demonstrates the cyclic nature of ENSO events (Figure 9.6.4).

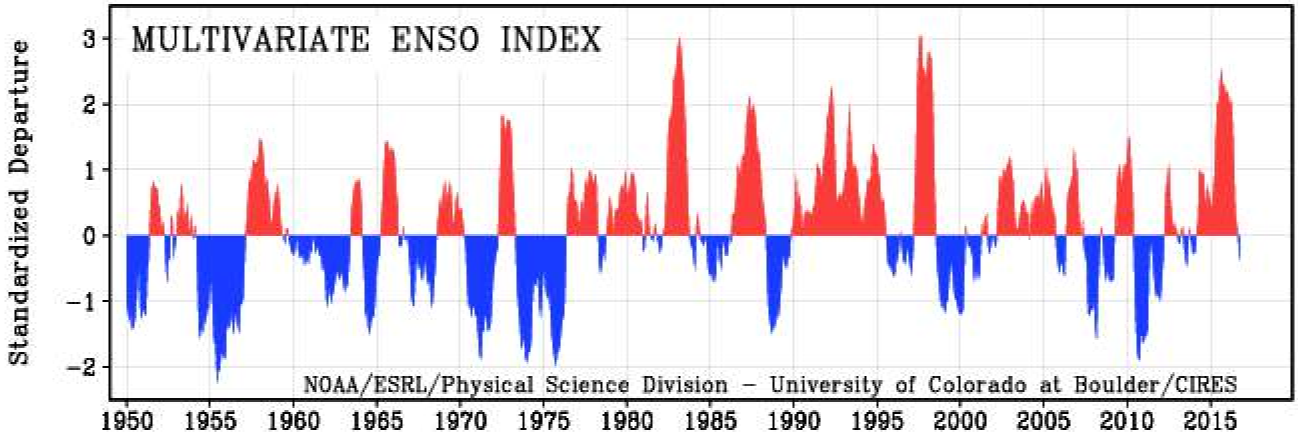


Figure 9.6.4 *Multivariate ENSO Index over time. Positive (red) values indicate warmer than normal conditions, while negative (blue) values represent conditions that are cooler than average. The greater the deviation from zero, the stronger the event. Note the intense El Niños in 1983, 1997-1998, and 2015 (NOAA, https://www.esrl.noaa.gov/psd/enso/mei/).*

Figure 9.6.5 shows a comparison of sea surface temperatures in the equatorial Pacific during normal, El Niño, and La Niña periods.

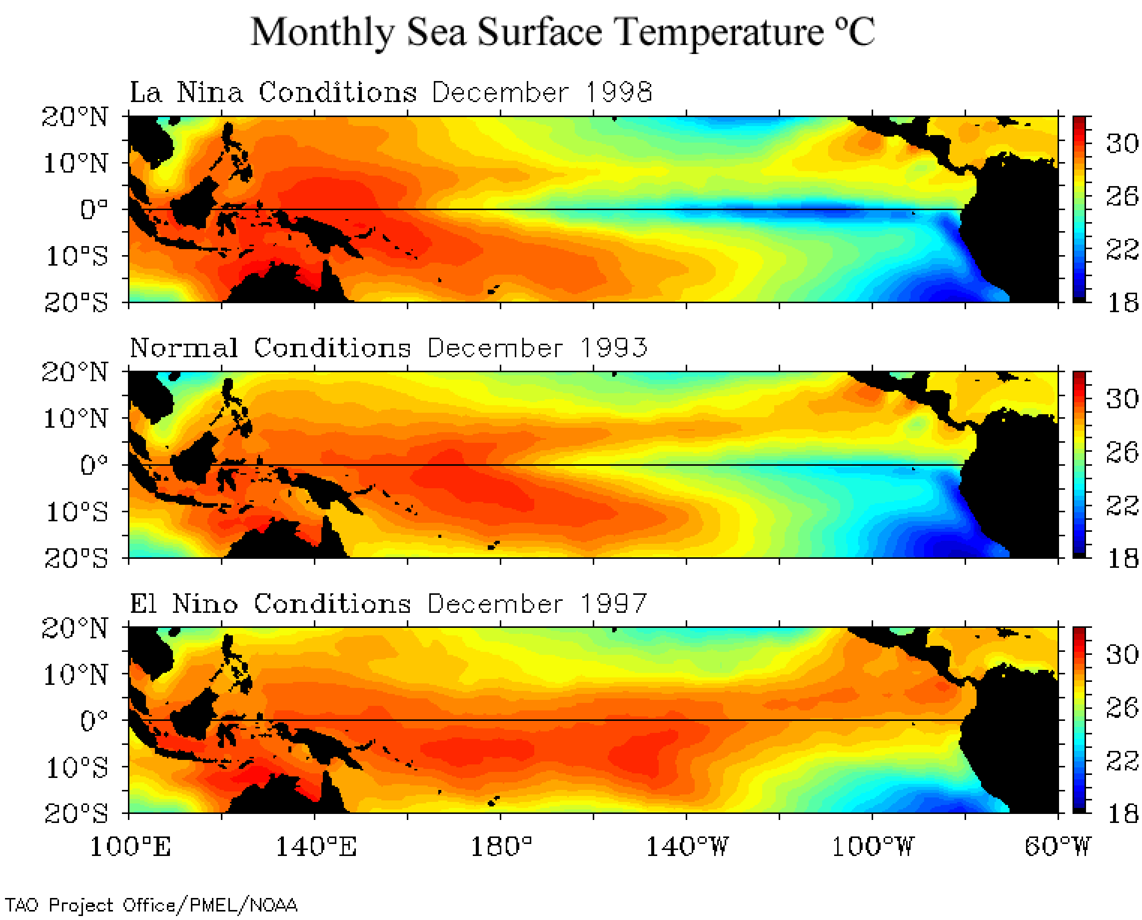


Figure 9.6.5 *Comparison of mean December sea surface temperatures in the equatorial Pacific during La Niña (top), normal (middle), and El Niño (bottom) conditions (NOAA, http://www.pmel.noaa.gov/elnino/sites/default/files/thumbnails/image/monthly-sst-lanina-normal-elnino.gif).*

Impacts of El Niño

The 2014-2016 El Niño was one of the strongest ENSO events on record (Figure 9.6.4). Some of the recorded global impacts of this El Niño included:

- Widespread droughts in the Philippines and many South Pacific island nations.
- Severe coral bleaching on the Great Barrier Reef in Australia.
- One of the most destructive bushfire seasons in Australia, in part due to low rainfall.
- High rainfall in the southeastern United States and parts of California, leading to flooding.
- Mild, low-precipitation winter in the New England region of the United States.
- Severe flooding in Peru and Argentina.
- Droughts in many portions of southern Africa.
- Nearly 100 million people worldwide suffered a lack of food or water from flooding and droughts.
- Peru suspended its second anchovy fishing season due to low biomass, and an anticipated 20% reduction in the yearly catch.

Additional links for more information

- What are the current ENSO conditions?: https://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/enso.shtml

9.7 Langmuir Circulation

One final type of wind-driven surface current occurs on a smaller scale when the winds blow in a consistent direction at relatively high speeds. Underneath these strong winds, water flows in a series of parallel corkscrew patterns called Langmuir cells (Figure 9.7.1 top). Each of the cells may be several meters wide and several meters deep, and adjacent cells rotate in opposite directions. While the overall direction of the corkscrew motion is in the direction of the wind, the rotation of the cells is roughly perpendicular to the wind direction. A divergence zone is created where the surface of the neighboring cells are rotating away from each other, and there is a degree of upwelling between the cells. Where the surface water of neighboring cells is rotating towards each other, a convergent zone is formed, with a region of downwelling between the cells. These alternating regions of divergence and convergence often take debris, foam, or algae floating at the surface and concentrate them along the convergence zones, creating long slicks running parallel to the wind direction (Figure 9.7.1 bottom).

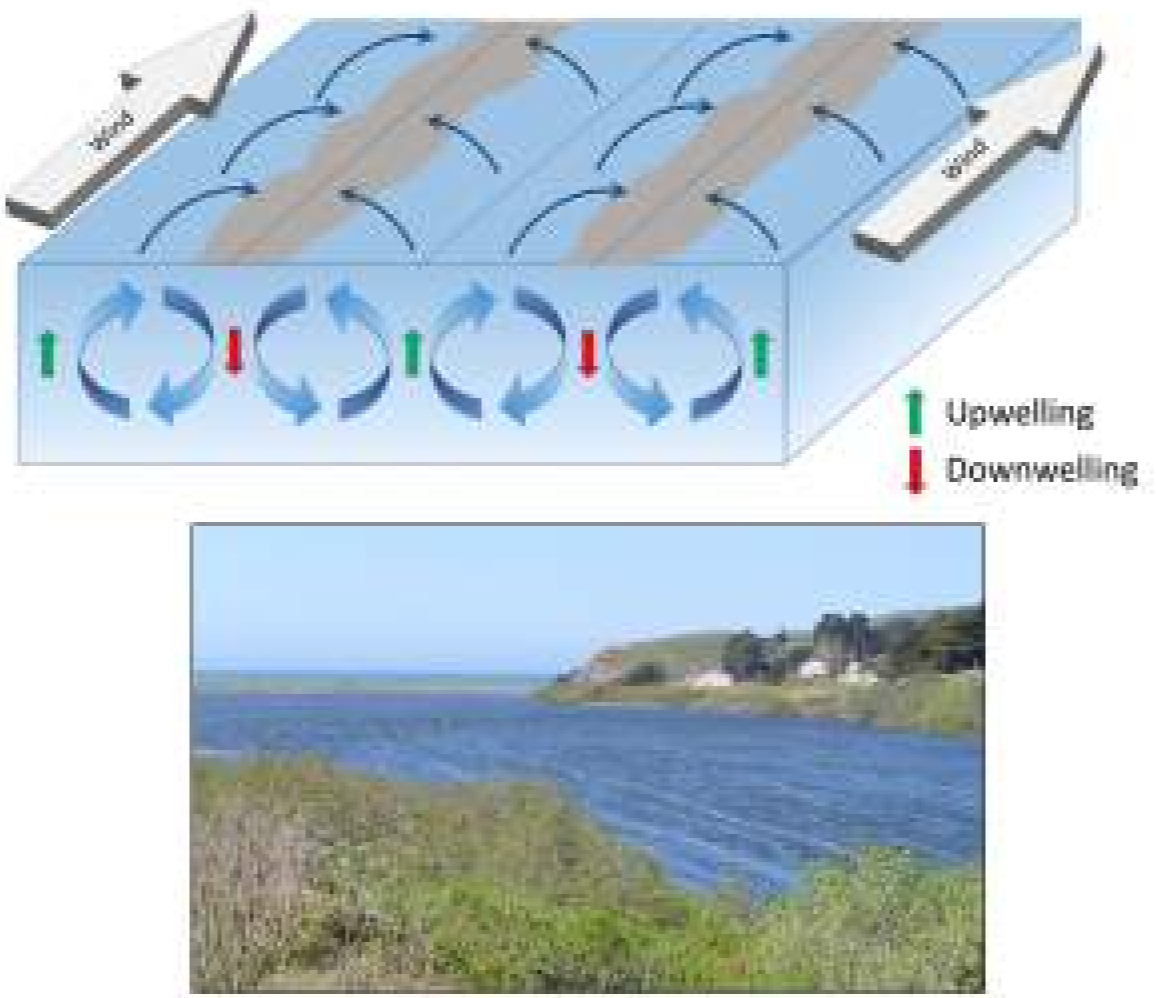


Figure 9.7.1 *(Top) Langmuir cells created by strong, sustained winds (gray arrows). Dark blue arrows show the movement of surface water as a result of the cells, producing downwelling convergence zones where floating material can accumulate (gray shading), and upwelling divergence zones. (Bottom) Parallel patterns of accumulated foam and debris indicating Langmuir circulation (Top: PW; Bottom: Photo By Mmelugin (Own work) [CC BY-SA 3.0 or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons).*

9.8 Thermohaline Circulation

The surface currents we have discussed so far are ultimately driven by the wind, and since they only involve surface water they only affect about 10% of the ocean's volume. However, there are other significant ocean currents that are independent of the wind, and involve water movements in the other 90% of the ocean. These currents are driven by differences in water density.

Recall that less dense water remains at the surface, while denser water sinks. Waters of different densities tend to stratify themselves into layers, with the densest, coldest water on the bottom and warmer, less dense water on top. It is the movement of these density layers that create the deep water circulation. Since seawater density depends mainly on temperature and salinity (section 6.3), this circulation is referred to as **thermohaline circulation**.

The main processes that increase seawater density are cooling, evaporation, and ice formation. Evaporation and ice formation cause an increase in density by removing fresh water, leaving the remaining seawater with greater salinity (see section 5.3). The main processes that decrease seawater density are heating, and dilution by fresh water through precipitation, ice melting, or fresh water runoff. Note that all of these processes exert their effects at the surface, but don't necessarily affect deeper water. However, changing the density of the surface water causes it to sink or rise, and these vertical, density-driven movements create the deep ocean currents. These thermohaline currents are slow, on the order of 10-20 km per year compared with surface currents that move at several kilometers per hour.

Water masses

A water mass is a volume of seawater with a distinctive density as a result of its unique profile of temperature and salinity. As stated above, the processes that affect seawater density really only happen at the surface. Once a water mass has reached its particular temperature and salinity profile due to these surface processes, it may sink below the surface, at which point its density properties won't really change. We can therefore distinguish particular water masses by taking salinity and temperature measurements at different depths, and looking for the unique combination of these variables that give it its characteristic density. This is often carried out using temperature-salinity diagrams (T-S diagrams, see box below).

There are several well-known water masses in the ocean, particularly in the Atlantic, that are distinguished by their temperature and salinity characteristics. The densest ocean water is formed in two primary locations near the poles, where the water is very cold and highly saline as a result of ice formation. The densest deep water mass is formed in the Weddell Sea of Antarctica, and becomes the **Antarctic Bottom Water (AABW)**. Similar processes in the North Atlantic produce the **North Atlantic Deep Water (NADW)** in the Greenland Sea (Figure 9.8.1).

Figure 9.8.1 *The primary sites of deep water formation; Antarctic Bottom Water is formed int he Weddell Sea, and North Atlantic Deep water is formed in the Greenland Sea (PW).*

This cold, dense water sinks, and once it is removed from the surface, its temperature and salinity remain unchanged, so it keeps the same characteristics as it moves throughout the ocean as part of the thermohaline circulation. AABW sinks to the bottom in the Weddell Sea and then moves north along the bottom into the Atlantic, and east through the Southern Ocean. At the same time NADW is sinking in the Greenland Sea. This water mass is less dense than AABW and tends to form a layer above the AABW as it flows across the equator to the south (Figure 9.8.2). As the NADW moves towards the Antarctic continent, it is brought to the surface. Recall that near Antarctica there is the Antarctic divergence, where surface waters move horizontally away from each other, and are replaced by deep water upwelling (bringing nutrients to the surface and leading to high productivity; see section 7.3). Since polar water has a weak thermocline, there isn't much of a density difference preventing the deep water from reaching the surface, so some NADW rises as part of the upwelling process (Figure 9.8.2).

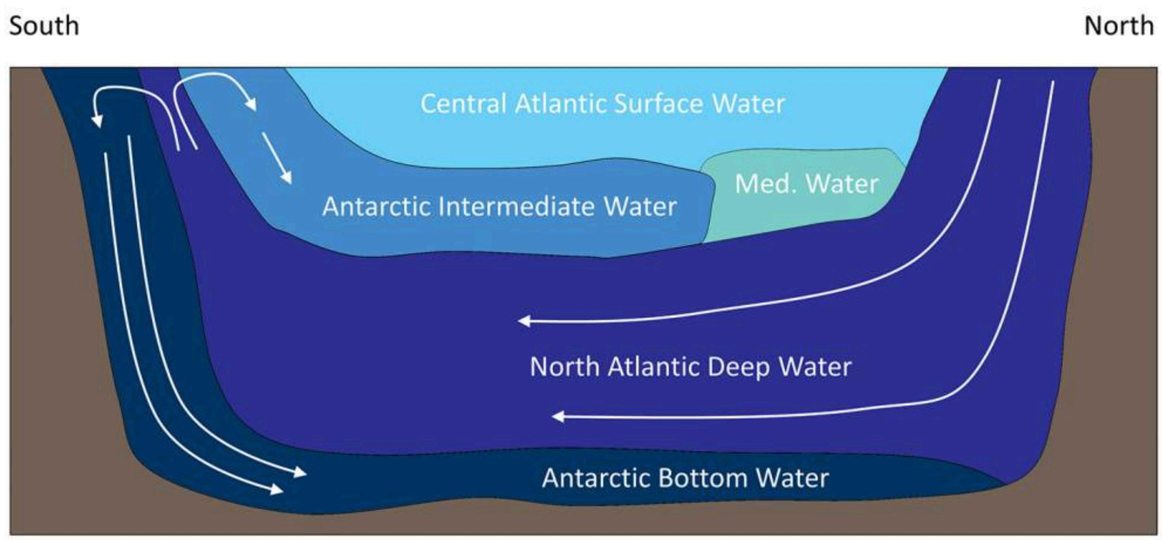


Figure 9.8.2 *The major water masses of the Atlantic Ocean (PW).*

As the rising NADW reaches the surface, some travels south where it will eventually contribute to the production of new AABW. The NADW that moves north encounters the Antarctic convergence, which produces downwelling. This sinking NADW becomes a new water mass; **Antarctic Intermediate Water (AAIW)**, which sinks and creates a layer between the surface water and the NADW (Figure 9.8.2). The surface water in the equatorial Atlantic, also called the **Central Atlantic Surface Water**, is very warm and low density, therefore it remains at the surface and does not contribute much to thermohaline circulation.

In the Atlantic, **Mediterranean Intermediate Water (MIW)** flows through the Straits of Gibraltar into the open ocean. This water is warm and salty from the warm temperatures and high evaporation characteristic of the Mediterranean Sea, so it is denser than the normal surface water and forms a layer about 1-1.5 km deep. Eventually this water will move north to the Greenland Sea, where it will be cooled and will sink, becoming the dense NADW.

T-S Diagrams

A temperature-salinity (T-S) diagram is used to examine how temperature, salinity, and density change with depth, and to identify the vertical structure of the water column, including the water masses it contains. Water temperature is on the y-axis, and salinity appears on the x-axis. Often, instead of the actual temperature of the water, oceanographers plot **potential temperature,** which is the temperature the water would achieve if it was brought to the surface and did not get any additional heat through compression at depth. A T-S diagram shows lines of equal density, or **isopycnals**, for various combinations of temperature and salinity (Figure 9.8.3). You can then plot the values of temperature and salinity on the diagram, and use their point of intersection to calculate the density of the water. In the example in Figure 9.8.3 a temperature of about 11^{o} C and a salinity of 34.6 PSU results in a density of 1.0265 g/cm^3.

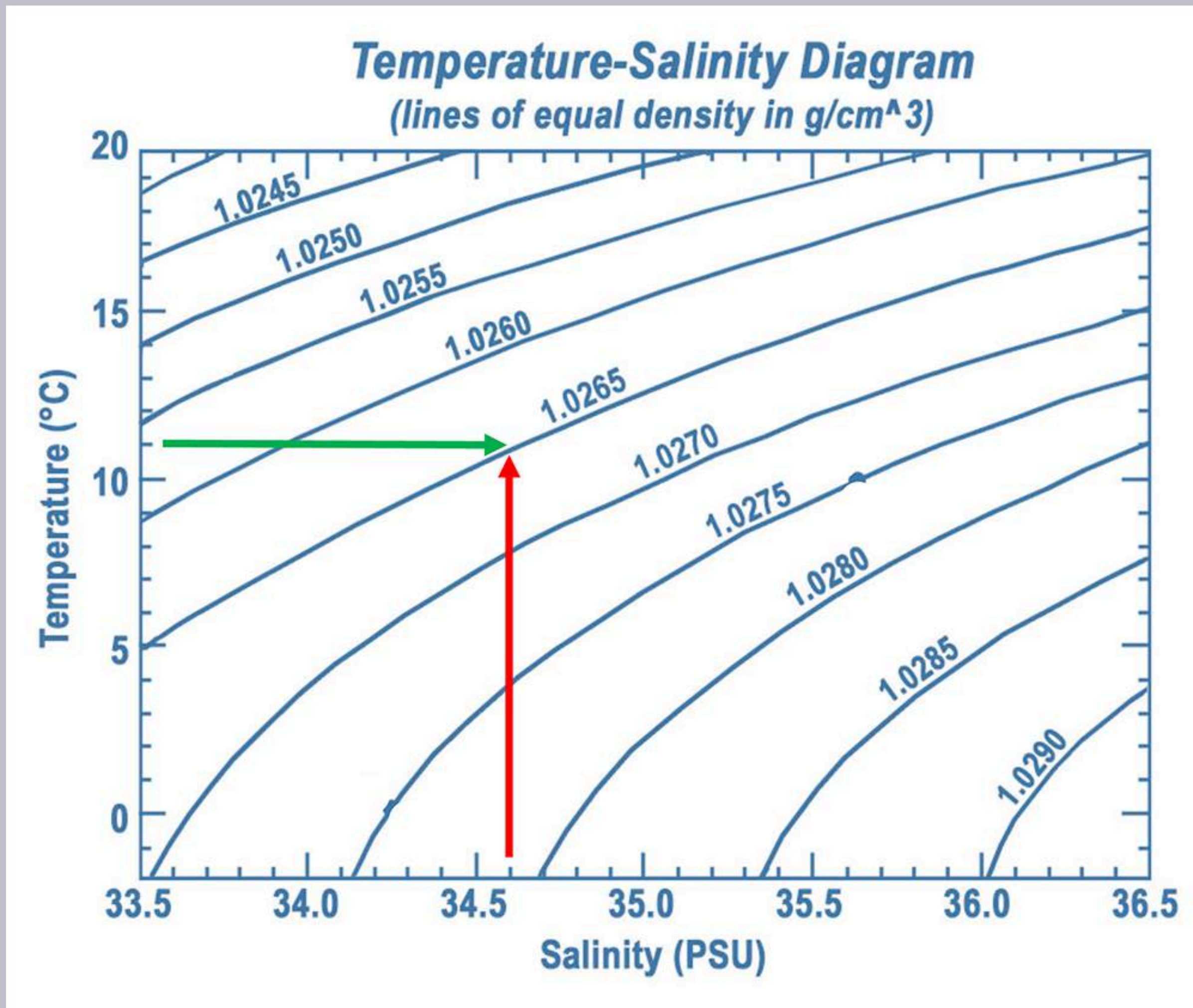


Figure 9.8.3 *Using a T-S diagram to determine density. A temperature of about 11° C (green arrow) and a salinity of 34.6 PSU (red arrow) results in a density of 1.0265 g/cm³.*

Since the range of densities in the ocean is rather small, often the density value is shortened and is expressed as sigma-t or σ_t. Sigma-t is calculated as: (density – 1) x 1000. So it essentially just looks at the last three decimal places of the density value. Thus a density of 1.0275 g/cm^3 would have a σ_t of 27.5.

T-S diagrams can be used to identify water masses. Since each major water mass has its own characteristic range of temperatures and salinities, a deep water sample that falls into that range can presumably have come from that water mass. Figure 9.8.4 shows the typical range of temperature and salinity for the major Atlantic water masses.

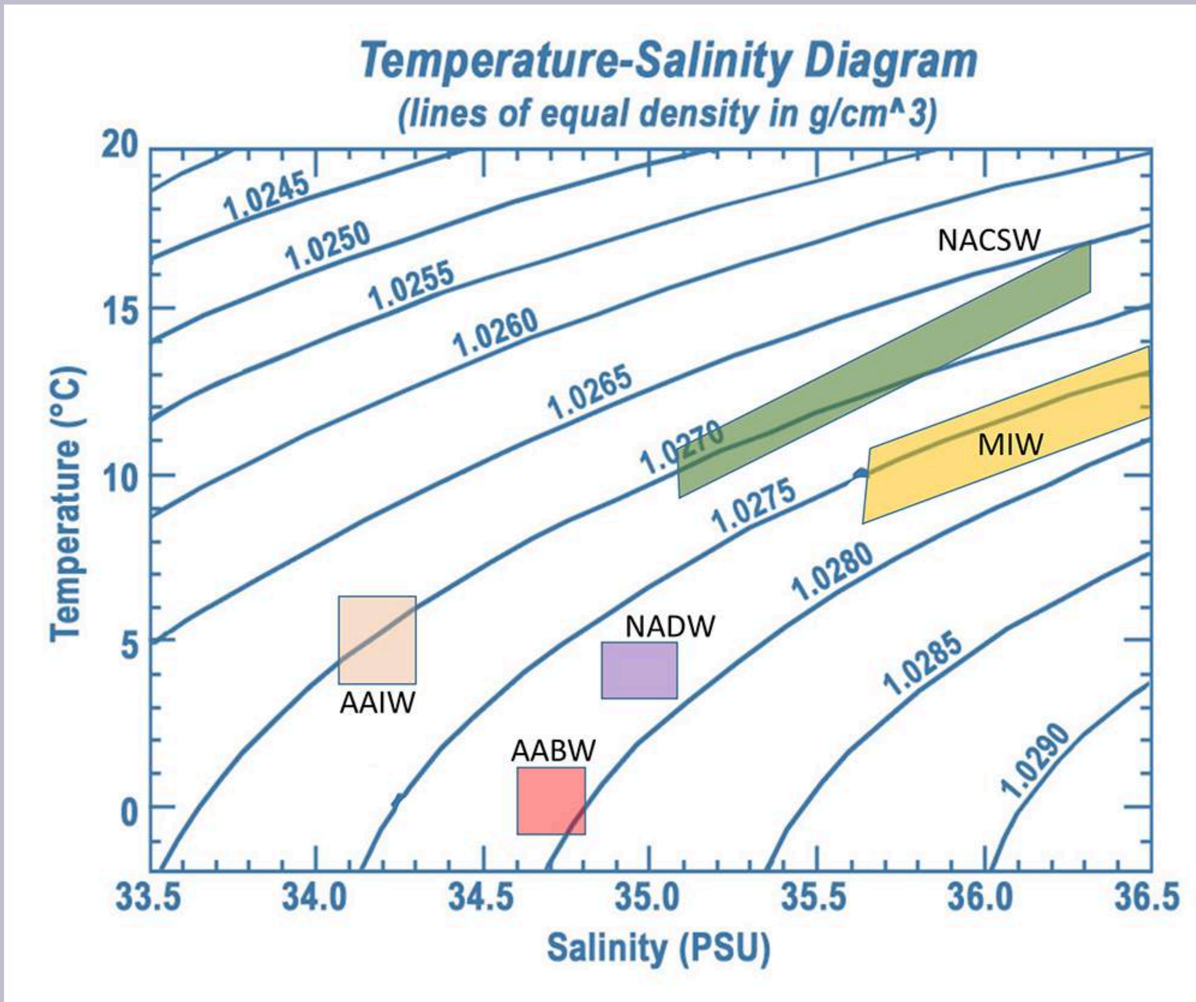


Figure 9.8.4 *Characteristic ranges of temperature and salinity for the major Atlantic water masses; North Atlantic Central Surface Water (NACSW), Mediterranean Intermediate Water (MIW), Antarctic Intermediate Water (AAIW), North Atlantic Deep Water (NADW), and Antarctic Bottom Water (AABW).*

To investigate water masses, oceanographers can take a series of temperature and salinity measurements over a range of depths at a particular location. If the water column was highly stratified and there was no mixing between or within the layers, as the probe was lowered your would get a series of constant temperature and salinity readings as you moved through the first water mass, followed by a sudden jump to another set of different but constant readings as you moved through the next water mass. Plotting temperature vs. salinity on a T-S diagram would result in a distinct and independent point for each water mass. However, in reality, the water masses will show some mixing within and between layers. So as the probes are lowered, they will encounter water that shows traits intermediate between the two points. Therefore, with increasing depth, the points on the T-S diagram will gradually move from one point to the other, creating a line connecting the two points, illustrating mixing between those two water masses.

In the example in Figure 9.8.5, NACSW is present at the surface (0 m depth), and between 0 and about 800 m there is a transition from NACSE into AAIW. Between about 800-2100 m there is a transition from AAIW into the NADW layer just beyond 2000 m. AABW is the deepest water mass, at depths of about 4000 m. The transition between NADW and AABW occurs between about 2100-4000 m.

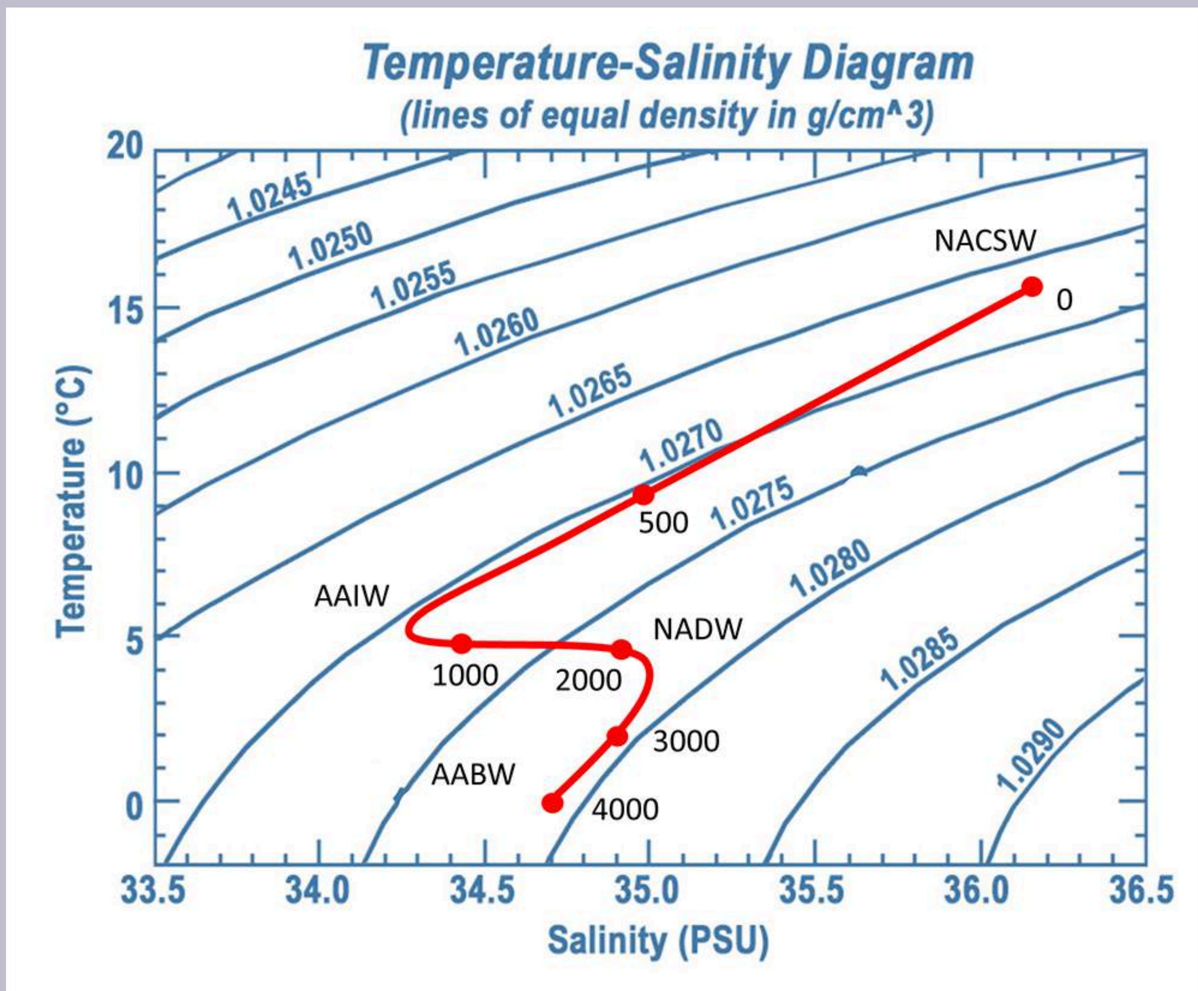


__Figure 9.8.5__ Hypothetical T-S diagram for the North Atlantic. Points represent readings taken at the corresponding depths (m). Moving from the surface to the bottom results in water of increasing density, passing through distinct water masses.

Notice that as the recordings get deeper in Figure 9.8.5, the density is always increasing (i.e. moving towards the bottom right corner). This is because the densest water should be located at the bottom, with the other layers stratified according to their density, otherwise the water column would be unstable.

The "Ocean Conveyor Belt"

The bottom water from the Weddell Sea and Greenland Sea does not just circulate through the Atlantic. NADW moves south through the western Atlantic before meeting the AABW north of the Weddell Sea. Together these water masses move eastwards into the Indian and Pacific Oceans. By this time the NADW and AABW have started mixing, to create what is called **Common Water**. The deep Common Water moves northwards into the Pacific and Indian Oceans and gradually mixes with the warmer water, causing it to eventually rise to the surface. As surface water, it makes its way back to the North Atlantic through the surface currents of the Pacific and Indian Oceans. Once back in the North Atlantic, it cools and once again forms NADW, starting the process anew. This cycle of rising and sinking water transporting water between the surface and deep circulation has been referred to as the global oceanic "conveyor belt", and may take about 1000-2000 years to complete (Figure 9.8.6).

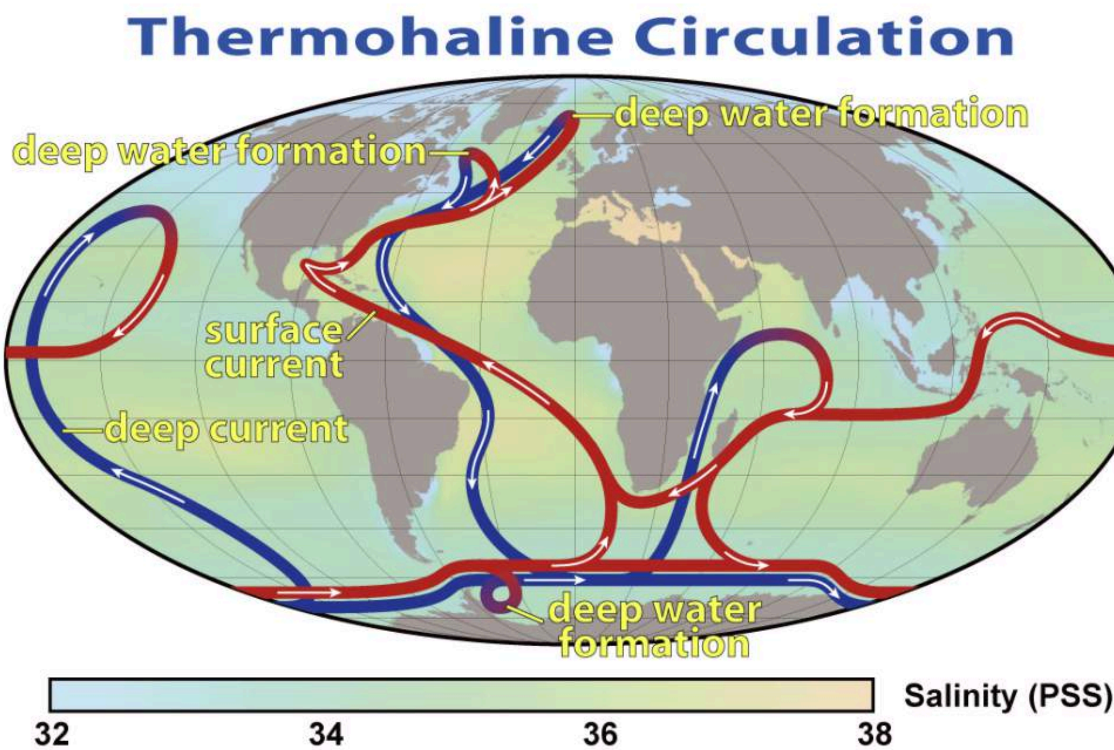


Figure 9.8.6 *The global ocean "conveyor belt." Cold, dense water sinks in the Greenland and Weddell Seas and circulates over the seafloor into the Indian and Pacific Oceans (blue paths). Eventually the water rises to the surface, and returns to the site of bottom water formation via surface currents (red paths), to start the cycle again (By Robert Simmon, NASA. Minor modifications by Robert A. Rohde also released to the public domain (NASA Earth Observatory) [Public domain], via Wikimedia Commons).*

This global circulation pattern has a number of important implications for Earth's environment. For one, it is vital to the transport of heat around the globe, bringing warm water towards the poles, and cold water to the tropics, stabilizing temperature in both environments.

The conveyor belt also helps deliver oxygen to deep water habitats. The deep water began as cold surface water that was saturated with oxygen, and when it sank it brought that oxygen to depth. Thermohaline circulation carries this oxygen-rich deep water throughout the oceans, where the oxygen will be used by deep water organisms. Bottom water in the Atlantic is relatively high in oxygen, as it still retains much of its original oxygen content, but as it travels over the seafloor the oxygen is used up, so that deep water in the Pacific Ocean has much less oxygen than deep Atlantic water, with Indian Ocean water somewhere in between. At the same time, deep water will accumulate nutrients as organic matter sinks and decomposes. Atlantic bottom water is low in nutrients because it has not had much time to accumulate them, and the original surface water was nutrient-poor. By the time this bottom water reaches the Indian Ocean, and after that the Pacific, it has been accumulating the sinking nutrients for centuries, so deep nutrient concentrations are greater in the Pacific than the Atlantic. We can therefore use the ratios of oxygen to nutrients in the deep water to tell the relative age of a water mass, i.e. how long it has been since it sank from the surface. Younger bottom water should be high in oxygen and low in nutrients, while the opposite would be expected for older bottom water.

The ocean conveyor belt may be significantly impacted by climate change disrupting thermohaline circulation. Increased warming, particularly in the Arctic, could lead to continuing melting of the polar ice caps, adding a large amount of fresh water to the polar surface water. This input of fresh water could create a low density, low salinity surface layer of water that no longer sinks, thus disrupting the deep circulation conveyor belt and preventing oxygen and nutrient transport to bottom communities. The sinking of seawater in the Greenland Sea also helps drive the Gulf Stream; as water sinks, more surface water is pulled northwards in the Gulf Stream. If this polar water stops sinking the Gulf Stream could weaken, reducing heat transport to the poles and cooling the northern climate. It seems counter intuitive, but global warming could lead to colder conditions in Europe and the freezing of ports and cities that are usually ice-free due to the warming effects of the Gulf Stream. Recent evidence has already shown that the strength of the Gulf Stream is waning, likely due to the increased melting of Arctic ice.

CHAPTER 10: WAVES

Chapter 10: Waves

Learning Objectives

After reading this chapter you should:

- know the parts of a basic wave
- know the terminology used to describe the motion of a wave (i.e. period, frequency, speed etc.)
- understand the circular motion of water particles involved in wave motion
- understand the difference between deep water waves and shallow water waves
- know what factors influence wave speed in deep and shallow waves
- know the three factors that determine the energy of wind-generated waves
- understand the concept of restoring force
- understand the difference between seas and swell
- understand the concepts of destructive, constructive and mixed interference
- understand why waves break as they approach shore
- know the differences in the different types of breakers, and how the bottom topography impacts breaker type
- understand why waves always approach parallel to shore, and why waves are larger off of points and smaller in bays
- understand what causes tsunamis, and how they behave in the ocean

Waves come in many shapes and sizes; a 100 foot wave might be a surfer's dream, but a ship captain's nightmare. What was the largest wave ever recorded? 50 feet? 100 feet? Not even close. That record belongs to a wave created in Lituya Bay, Alaska, on July 9, 1958 (Figure 10.1). On that day, a magnitude 7.8 earthquake caused a massive rockslide that slid down a mountainside and into the headwaters of the bay. The rockslide created a splash wave that was high enough to flatten vegetation up to 1722 ft (525 m) above sea level! The wave then moved through the narrow bay towards the sea, destroying a number of fishing boats along the way. Miraculously, a father and son on one fishing boat were carried above the trees by the wave, and survived to tell the story. This is by far the largest wave, a megatsunami, ever reliably recorded. The waves we will discuss in this chapter may not be quite that dramatic, but it is still important to know how they form, how they are propagated, and what happens to them as they interact with the shore.

Figure 10.1 *A view of Lituya Bay taken a few weeks after the 1958 megatsunami. The rockslide occurred in the mountains at the head of the bay, producing the wave that them moved through the bay towards the sea (D.J. Miller, United States Geological Survey, [Public domain], via Wikimedia Commons).*

10.1 Wave Basics

Waves generally begin as a disturbance of some kind, and the energy of that disturbance gets propagated in the form of waves. We are most familiar with the kind of waves that break on shore, or rock a boat at sea, but there are many other types of waves that are important to oceanography:

- **Internal waves** form at the boundaries of water masses of different densities (i.e. at a pycnocline), and propagate at depth. These generally move more slowly than surface waves, and can be much larger, with heights exceeding 100 m. However, the height of the deep wave would be unnoticeable at the surface.
- **Tidal waves** are due to the movement of the tides. What we think of as tides are basically enormously long waves with a wavelength that may span half the globe (see section 11.1). Tidal waves are not related to tsunamis, so don't confuse the two.
- **Tsunamis** are large waves created as a result of earthquakes or other seismic disturbances. They are also called seismic sea waves (section 10.4).
- **Splash waves** are formed when something falls into the ocean and creates a splash. The giant wave in Lituya Bay that was described in the introduction to this chapter was a splash wave.
- **Atmospheric waves** form in the sky at the boundary between air masses of different densities. These often create ripple effects in the clouds (Figure 10.1.1).

Figure 10.1.1 *Wake patterns in cloud cover over Possession Island, East Island, Ile aux Cochons, Ile de Pingouins. The ripple pattern is a result of internal waves in the atmosphere (NASA [Public domain], via Wikimedia Commons).*

There are several components to a basic wave (Figure 10.1.2):

- **Still water level:** where the water surface would be if there were no waves present and the sea was completely calm.
- **Crest:** the highest point of the wave.
- **Trough:** the lowest point of the wave.
- **Wave height:** the distance between the crest and the trough.
- **Wavelength:** the distance between two identical points on successive waves, for example crest to crest, or trough to trough.
- **Wave steepness:** the ratio of wave height to length (H/L). If this ratio exceeds 1/7 (i.e. height exceeds 1/7 of the wavelength) the wave gets too steep, and will break.

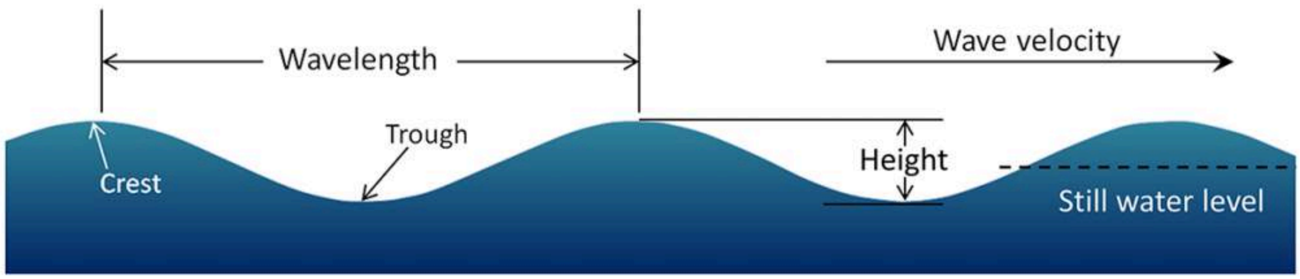


***Figure 10.1.2** Components of a basic wave (Modified by PW from Steven Earle "Physical Geology").*

There are also a number of terms used to describe wave motion:

- **Period:** the time it takes for two successive crests to pass a given point.
- **Frequency:** the number of waves passing a point in a given amount of time, usually expressed as waves per second. This is the inverse of the period.
- **Speed**: how fast the wave travels, or the distance traveled per unit of time. This is also called celerity (c), where

$$c = \text{wavelength} \times \text{frequency}$$

Therefore, the longer the wavelength, the faster the wave.

Although waves can travel over great distances, the water itself shows little horizontal movement; it is the *energy* of the wave that is being transmitted, not the water. Instead, the water particles move in circular orbits, with the size of the orbit equal to the wave height (Figure 10.1.3). This orbital motion occurs because water waves contain components of both longitudinal (side to side) and transverse (up and down) waves, leading to circular motion. As a wave passes, water moves forwards and up over the wave crests, then down and backwards into the troughs, so there is little horizontal movement. This is evident if you have ever watched an object such as a seabird floating at the surface. The bird bobs up and down as the wave pass underneath it; it does not get carried horizontally by a single wave crest.

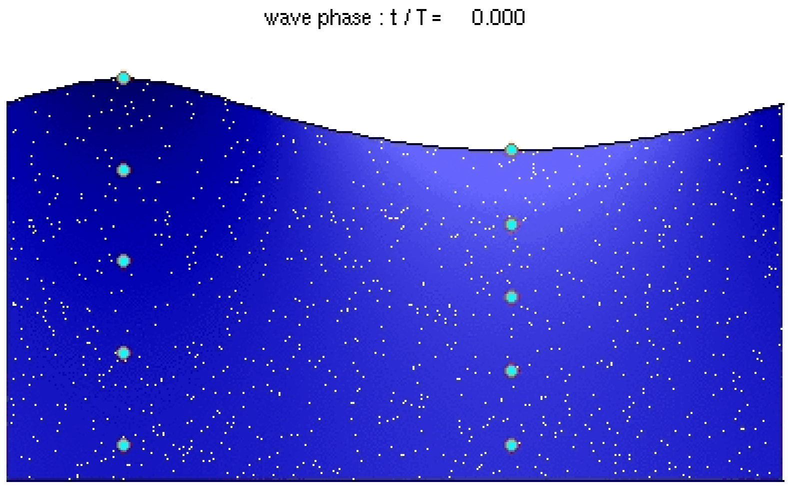


Figure 10.1.3 *Animation showing the orbital motion of particles in a surface wave (By Kraaiennest (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 4.0], via Wikimedia Commons).*

The circular orbital motion declines with depth as the wave has less impact on deeper water and the diameter of the circles is reduced. Eventually at some depth there is no more circular movement and the water is unaffected by surface wave action. This depth is the **wave base** and is equivalent to half of the wavelength (Figure 10.1.4). Since most ocean waves have wavelengths of less than a few hundred meters, most of the deeper ocean is unaffected by surface waves, so even in the strongest storms marine life or submarines can avoid heavy waves by submerging below the wave base.

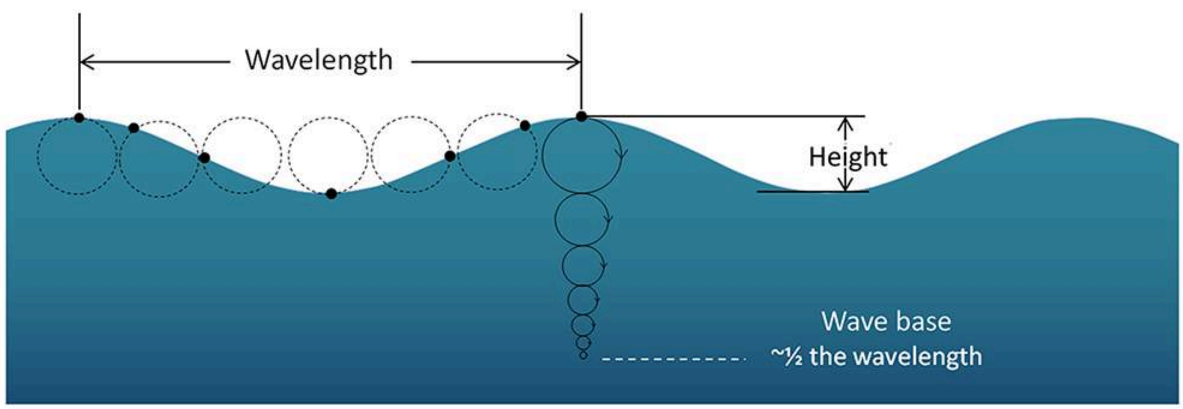


Figure 10.1.4 *Orbital motion of water within a wave, extending down to the wave base at a depth of half of the wavelength (Modified by PW from Steven Earle, "Physical Geology").*

When the water below a wave is deeper than the wave base (deeper than half of the wavelength), those waves are called **deep water waves**. Most open ocean waves are deep water waves. Since the water is deeper than the wave base, deep water waves experience no interference from the bottom, so their speed only depends on the wavelength:

$$\text{speed (m/s)} = \sqrt{\frac{gL}{2\pi}}$$

where *g* is gravity and *L* is wavelength in meters. Since *g* and *π* are constants, this can be simplified to:

$$\text{speed (m/s)} = 1.25\sqrt{L}$$

Shallow water waves occur when the depth is less than 1/20 of the wavelength. In these cases, the wave is said to "touch bottom" because the depth is shallower than the wave base so the orbital motion is affected by the seafloor. Due to the shallow depth, the orbits are flattened, and eventually the water movement becomes horizontal rather than circular just above the bottom. The speed of shallow water waves depends only on the depth:

$$\text{speed (m/s)} = \sqrt{gd}$$

where *g* is gravity and *d* is depth in meters. This can be simplified to:

$$\text{speed (m/s)} = 3.13\sqrt{d}$$

Intermediate or **transitional waves** are found in depths between ½ and 1/20 of the wavelength. Their behavior is a bit more complex, as their speed is influenced by both wavelength and depth. The speed of an intermediate wave is calculated as:

$$\text{speed (m/s)} = \sqrt{\frac{gL}{2\pi}\tanh(2\pi\frac{d}{L})}$$

which contains both depth and wavelength variables.

10.2 Waves at Sea

Most ocean waves are generated by wind. Wind blowing across the water's surface creates little disturbances called **capillary waves**, or ripples that start from gentle breezes (Figure 10.2.1). Capillary waves have a rounded crest with a V-shaped trough, and wavelengths less than 1.7 cm. These small ripples give the wind something to "grip" onto to generate larger waves when the wind energy increases, and once the wavelength exceeds 1.7 cm the wave transitions from a capillary wave to a wind wave. As waves are produced, they are opposed by a **restoring force** that attempts to return the water to its calm, equilibrium condition. The restoring force of the small capillary waves is surface tension, but for larger wind-generated waves gravity becomes the restoring force.

Figure 10.2.1 *Small capillary waves or ripples caused by winds blowing over the surface of calm water (By Blue Elf (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 3.0], via Wikimedia Commons).*

As the energy of the wind increases, so does the size, length and speed of the resulting waves. There are three important factors determining how much energy is transferred from wind to waves, and thus how large the waves will get:

- Wind **speed.**
- The **duration** of the wind, or how long the wind blows continuously over the water.
- The distance over which the wind blows across the water in the same direction, also known as the **fetch**.

Increasing any of these factors increases the energy of wind waves, and therefore their size and speed. But there is an upper limit to how large wind-generated waves can get. As wind energy increases, the waves receive more energy and they get both larger and steeper (recall from section 10.1 that wave steepness = height/wavelength). When the wave height exceeds 1/7 of the wavelength, the wave becomes unstable and collapses, forming whitecaps.

The ocean surface represents an irregular mixture of hundreds of waves of different speeds and sizes, all coming from different directions and interacting with each other. A histogram of wave heights within this mixture reveals a bell-shaped curve (Figure 10.2.2). In addition to basic statistics such as mode (most probable), median and mean wave height, wave heights are also reported in other ways. Marine weather forecasts and ship and buoy data often report **significant wave height (H_S)**, which is the mean height of the largest one-third of the waves. Mean wave height is approximately equal to two-thirds of the significant wave height. Finally, there is the minimum height of the highest 10% of waves (the 90th percentile of wave heights), often expressed as $H_{1/10}$.

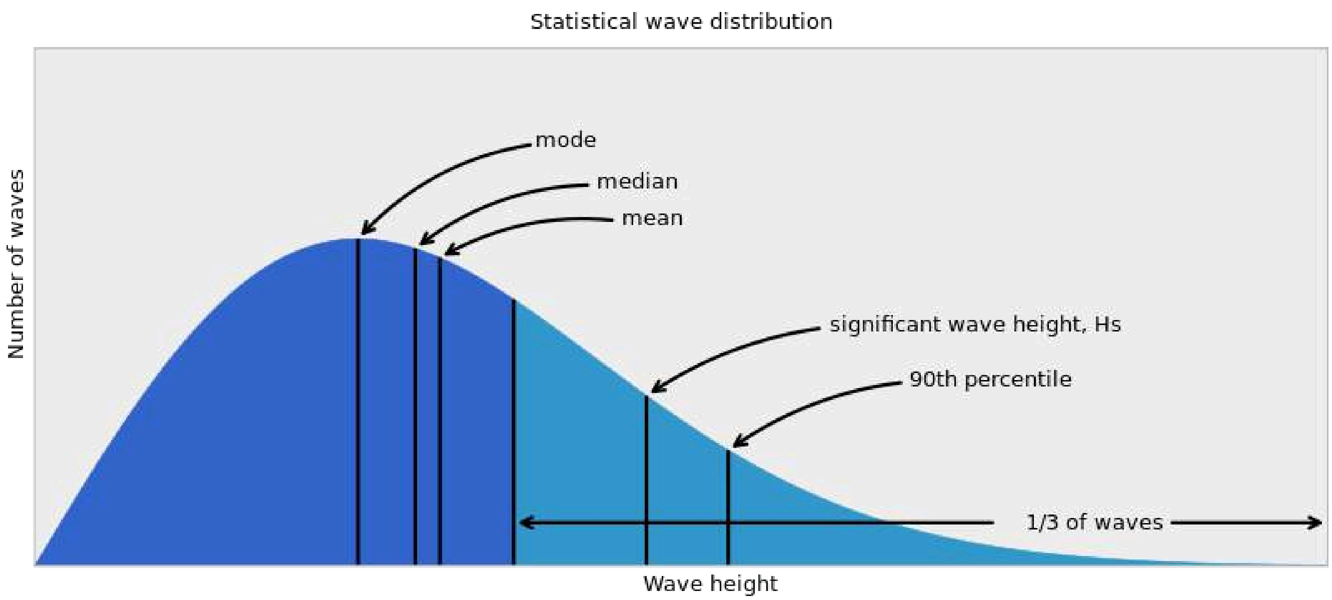
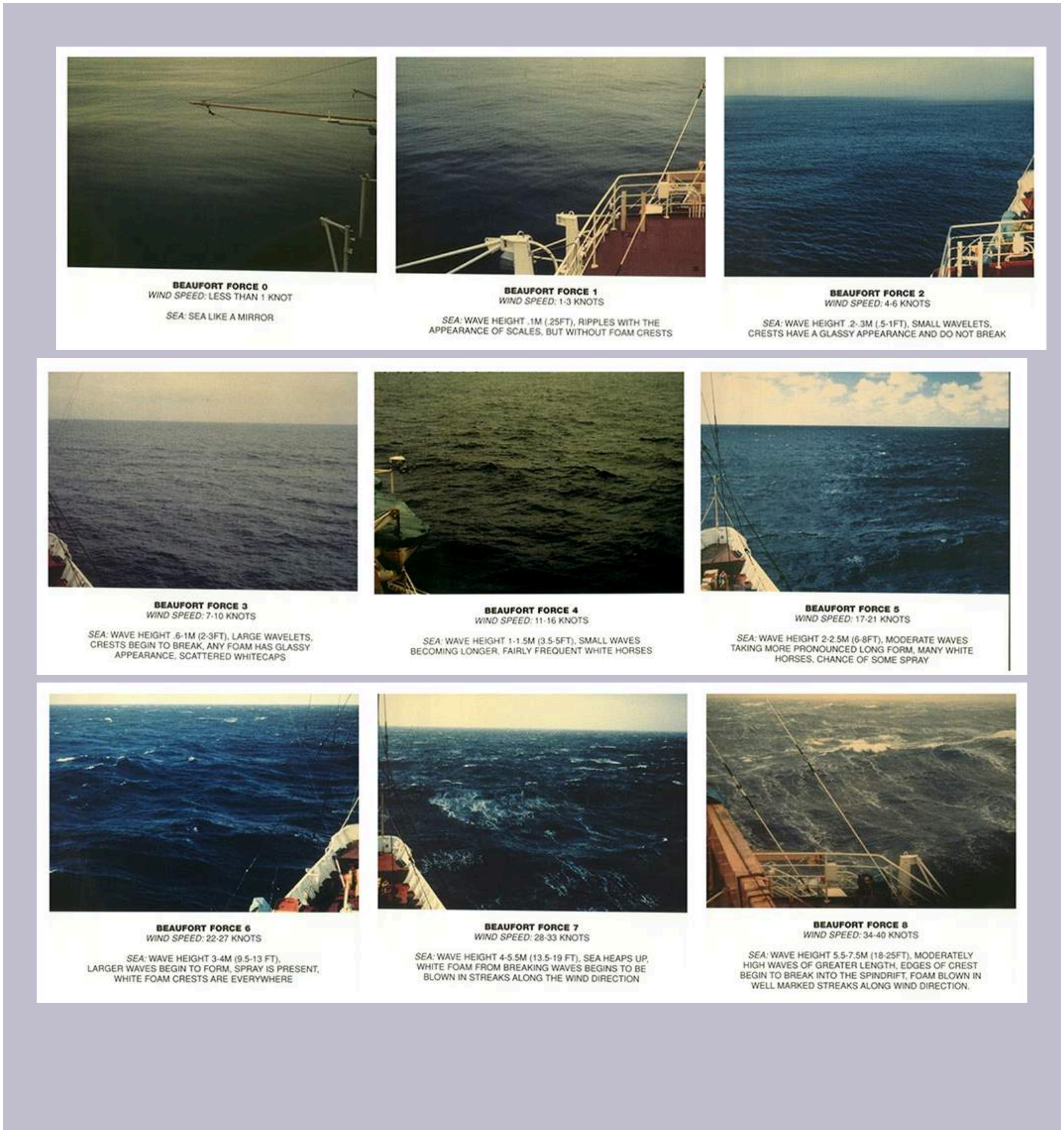


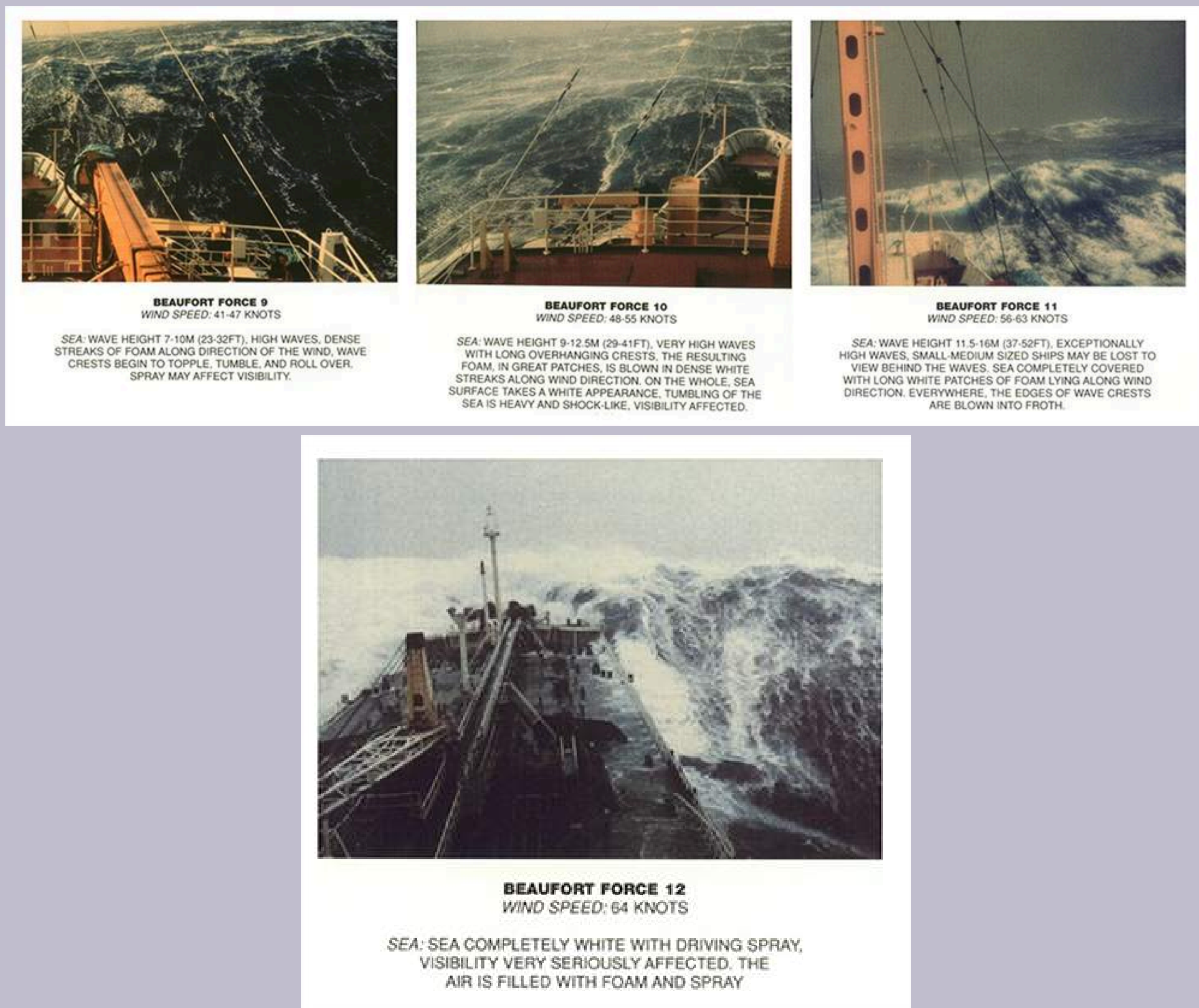
Figure 10.2.2 *Histogram of typical wave height distribution at sea, showing common statistical measurements (NOAA, Public Domain via Wikimedia Commons).*

Under strong wind conditions, the ocean surface becomes a chaotic mixture of choppy, whitecapped wind-generated waves. The term **sea state** describes the size and extent of the wind-generated waves in a particular area. When the waves are at their maximum size for the existing wind speed, duration, and fetch, it is referred to as a fully developed sea. The sea state is often reported on the **Beaufort scale**, ranging from 0-12, where 0 means calm, windless and waveless conditions, while Beaufort 12 is a hurricane (see box below).

The Beaufort Scale

The Beaufort scale is used to describe the wind and sea state conditions on the ocean. It is an observational scale based on the judgement of the observer, rather than one dictated by accurate measurements of wave height. Beaufort 0 represents calm, flat conditions, while Beaufort 12 represents a hurricane.


BEAUFORT FORCE 0
WIND SPEED: LESS THAN 1 KNOT
SEA: SEA LIKE A MIRROR
BEAUFORT FORCE 1
WIND SPEED: 1-3 KNOTS
SEA: WAVE HEIGHT .1M (.25FT), RIPPLES WITH THE APPEARANCE OF SCALES, BUT WITHOUT FOAM CRESTS
BEAUFORT FORCE 2
WIND SPEED: 4-6 KNOTS
SEA: WAVE HEIGHT .2-.3M (.5-1FT), SMALL WAVELETS, CRESTS HAVE A GLASSY APPEARANCE AND DO NOT BREAK
BEAUFORT FORCE 3
WIND SPEED: 7-10 KNOTS
SEA: WAVE HEIGHT .6-1M (2-3FT), LARGE WAVELETS, CRESTS BEGIN TO BREAK, ANY FOAM HAS GLASSY APPEARANCE, SCATTERED WHITECAPS
BEAUFORT FORCE 4
WIND SPEED: 11-16 KNOTS
SEA: WAVE HEIGHT 1-1.5M (3.5-5FT), SMALL WAVES BECOMING LONGER, FAIRLY FREQUENT WHITE HORSES
BEAUFORT FORCE 5
WIND SPEED: 17-21 KNOTS
SEA: WAVE HEIGHT 2-2.5M (6-8FT), MODERATE WAVES TAKING MORE PRONOUNCED LONG FORM, MANY WHITE HORSES, CHANCE OF SOME SPRAY
BEAUFORT FORCE 6
WIND SPEED: 22-27 KNOTS
SEA: WAVE HEIGHT 3-4M (9.5-13 FT), LARGER WAVES BEGIN TO FORM, SPRAY IS PRESENT, WHITE FOAM CRESTS ARE EVERYWHERE
BEAUFORT FORCE 7
WIND SPEED: 28-33 KNOTS
SEA: WAVE HEIGHT 4-5.5M (13.5-19 FT), SEA HEAPS UP, WHITE FOAM FROM BREAKING WAVES BEGINS TO BE BLOWN IN STREAKS ALONG THE WIND DIRECTION
BEAUFORT FORCE 8
WIND SPEED: 34-40 KNOTS
SEA: WAVE HEIGHT 5.5-7.5M (18-25FT), MODERATELY HIGH WAVES OF GREATER LENGTH, EDGES OF CREST BEGIN TO BREAK INTO THE SPINDRIFT, FOAM BLOWN IN WELL MARKED STREAKS ALONG WIND DIRECTION.



(Images by United States National Weather Service (http://www.crh.noaa.gov/mkx/marinefcst.php) [Public domain], via Wikimedia Commons).

A fully developed sea often occurs under stormy conditions, where high winds create a chaotic, random pattern of waves and whitecaps of varying sizes. The waves will propagate outwards from the center of the storm, powered by the strong winds. However, as the storm subsides and the winds weaken, these irregular seas will sort themselves out into more ordered patterns. Recall that open ocean waves will usually be deep water waves, and their speed will depend on their wavelength (section 10.1). As the waves move away from the storm center, they sort themselves out based on speed, with longer wavelength waves traveling faster than shorter wavelength waves. This means that eventually all of the waves in a particular area will be traveling with the same wavelength, creating regular, long period waves called **swell** (Figure 10.2.3). We experience swell as the slow up and down or rocking motion we feel on a boat, or with the regular arrival of waves on shore. Swell can travel very long distances without losing much energy, so we can observe large swells arriving at the shore even where there is no local wind; the waves were produced by a storm far offshore, and were sorted into swell as they traveled towards the coast.

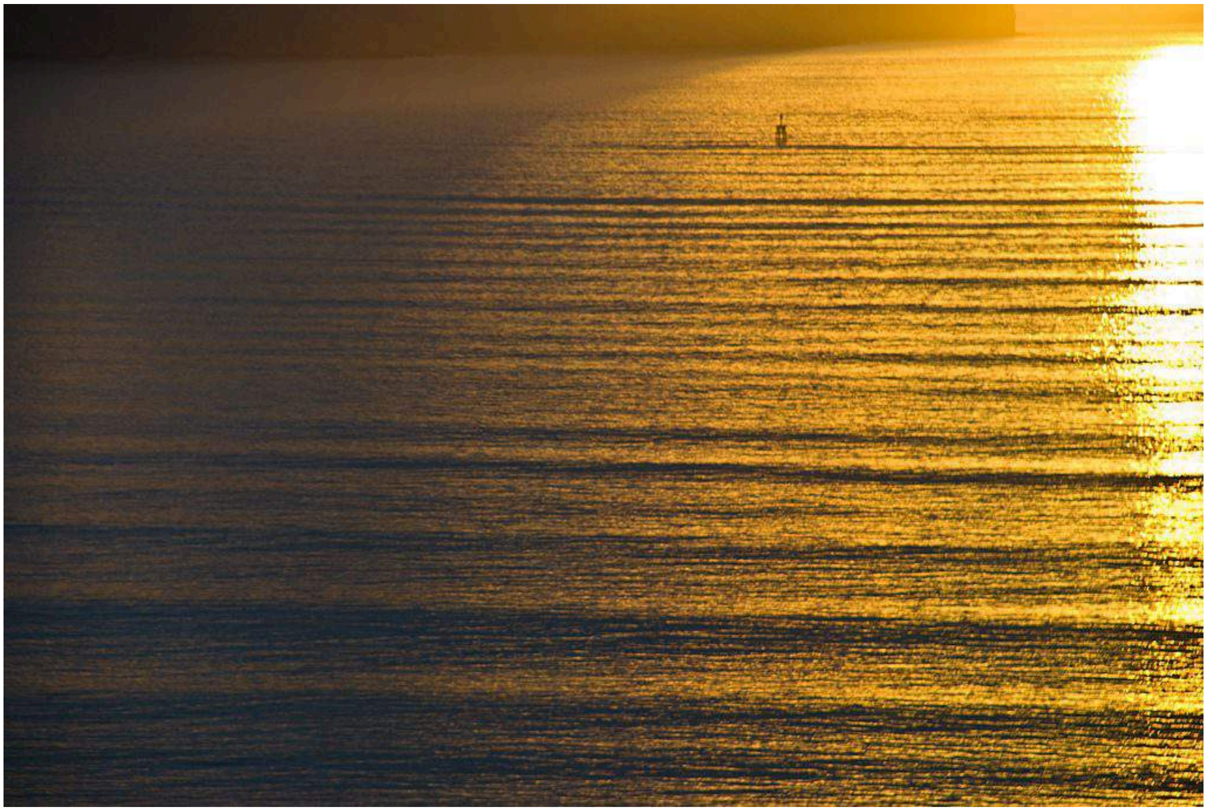


Figure 10.2.3 *Ocean swell, the regular pattern of waves of equal wavelength (Phillip Capper [CC BY 2.0], via Wikimedia Commons).*

Because swell travels such long distances, eventually swells coming from different directions will run into each other, and when they do they create interference patterns. The interference pattern is created by adding the features of the waves together, and the type of interference that is created depends on how the waves interact with each other (Figure 10.2.4). **Constructive interference** occurs when the two waves are completely in phase; the crest of one wave lines up exactly with the crest of the other wave, as do the troughs of the two waves. Adding the two crest together creates a crest that is higher than in either of the source waves, and adding the troughs creates a deeper trough than in the original waves. The result of constructive interference is therefore to create waves that are larger than the original source waves. In **destructive interference**, the waves interact completely out of phase, where the crest of one wave aligns with the trough of the other wave. In this case, the crest and the trough work to cancel each other out, creating a wave that is smaller than either of the source waves. In reality, it is rare to find perfect constructive or destructive interference as displayed in Figure 10.2.4. Most interference by swells at sea is **mixed interference**, which contains a mix of both constructive and destructive interference. The interacting swells do not have the same wavelength, so some points show constructive interference, and some points show destructive interference, to varying degrees. This results in an irregular pattern of both small and large waves, called **surf beat**.

It is important to point out that these interference patterns are only temporary disturbances, and do not affect the properties of the source waves. Moving swells interact and create interference where they meet, but each wave continues on unaffected after the swells pass each other.

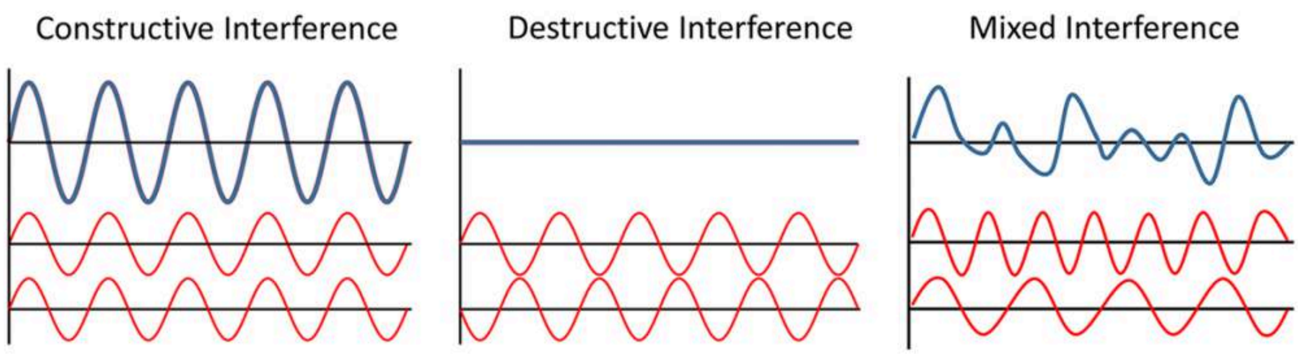


Figure 10.2.4 *Wave interference patterns. In constructive interference the source waves (red) are completely in phase, and when added together produce waves that are larger than the original waves (blue). In destructive interference the source waves are out of phase, so they cancel each other out and produce waves that are smaller than the originals. In mixed interference, constructive and destructive interference occur at various point, creating an irregular wave pattern. (Modified by PW from original version: Haade; vectorization: Wjh31, Quibik (Vecorized from File:Interference of two waves.png) [CC BY-SA 3.0 or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons).*

About half of the waves in the open sea are less than 2 m high, and only 10-15% exceed 6 m. But the ocean can produce some extremely large waves. The largest wind wave reliably measured at sea occurred in the Pacific Ocean in 1935, and was measured by the navy tanker the USS Ramapo. Its crew measured a wave of 34 m or about 112 ft high! Occasionally constructive interference will produce waves that are exceptionally large, even when all of the surrounding waves are of normal height. These random, large waves are called **rogue waves** (Figure 10.2.5). A rogue wave is usually defined as a wave that is at least twice the size of the significant wave height, which is the average height of the highest one-third of waves in the region. It is not uncommon for rogue waves to reach heights of 20 m or more.

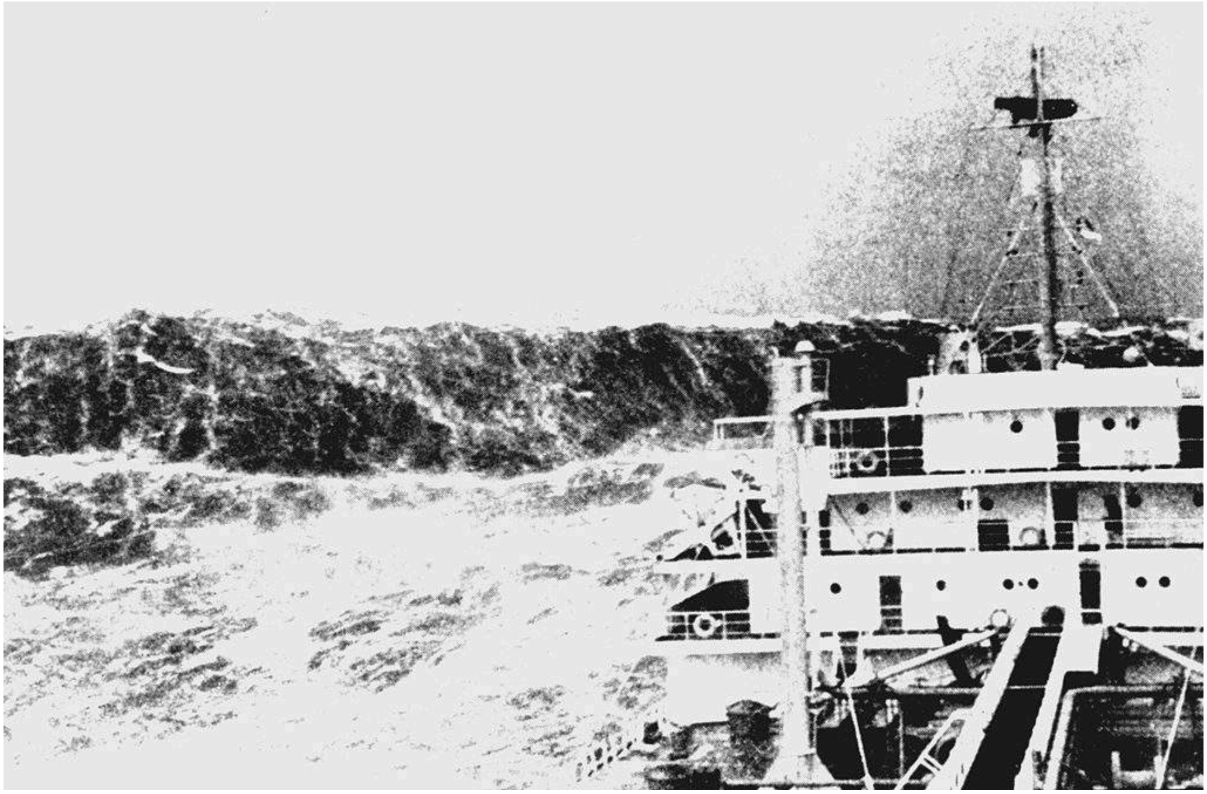


Figure 10.2.5 *A rogue wave in the Bay of Biscay, off of the French coast, ca. 1940 (NOAA, [Public domain], via Wikimedia Commons).*

Rogue waves are particularly common off of the southeast coast of South Africa, a region referred to as the “wild coast.” Here, Antarctic storm waves move north into the oncoming Agulhas Current, and the wave energy gets focused over a narrow area, leading to constructive interference. This area may be responsible for sinking more ships than anywhere else on Earth. On average about 100 ships are lost every year across the globe, and many of these losses are probably due to rogue waves.

Waves in the Southern Ocean are generally fairly large (the red areas in Figure 10.2.6) because of the strong winds and the lack of landmasses, which provide the winds with a very long fetch, allowing them to blow unimpeded over the ocean for very long distances. These latitudes have been termed the “Roaring Forties”, “Furious Fifties”, and “Screaming Sixties” due to the high winds.

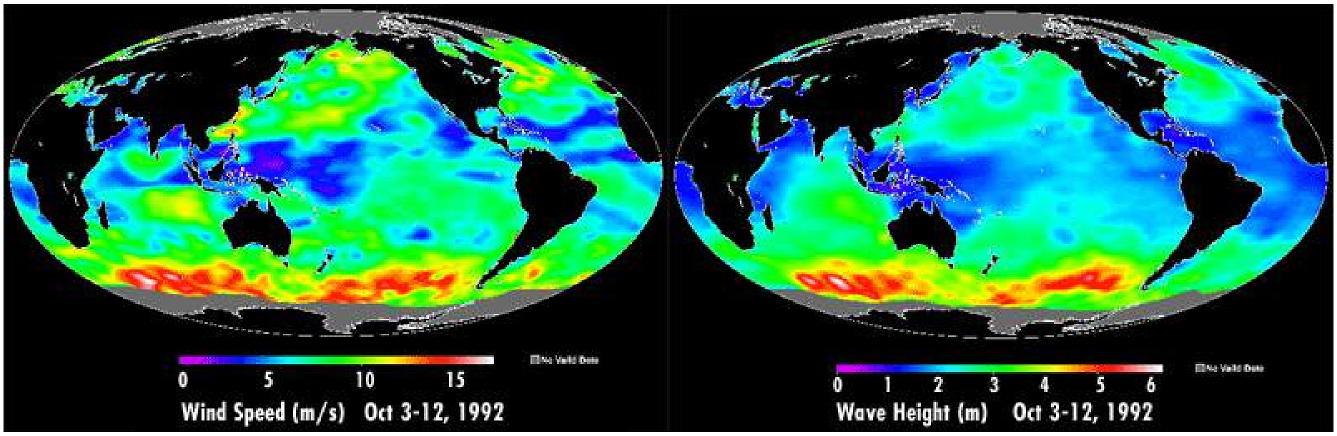


Figure 10.2.6 *Wind speed and wave height data for a 9-day period in 1992. The Southern Ocean is notorious for its high winds and large waves (NASA, Public Domain via Wikimedia Commons).*

10.3 Waves on the Shore

Most of the waves discussed in the previous section referred to deep water waves in the open ocean. But what happens when these waves move towards shore and encounter shallow water? Remember that in deep water, a wave's speed depends on its wavelength, but in shallow water wave speed depends on the depth (section 10.1). When waves approach the shore they will "touch bottom" at a depth equal to half of their wavelength; in other words, when the water depth equals the depth of the wave base (Figure 10.3.1). At this point their behavior will begin to be influenced by the bottom.

When the wave touches the bottom, friction causes the wave to slow down. As one wave slows down, the one behind it catches up to it, thus decreasing the wavelength. However, the wave still contains the same amount of energy, so while the wavelength decreases, the wave height increases. Eventually the wave height exceeds 1/7 of the wavelength, and the wave becomes unstable and forms a **breaker**. Often breakers will start to curl forwards as they break. This is because the bottom of the wave begins to slow down before the top of the wave, as it is the first part to encounter the seafloor. So the crest of the wave gets "ahead" of the rest of the wave, but has no water underneath it to support it (Figure 10.3.1).

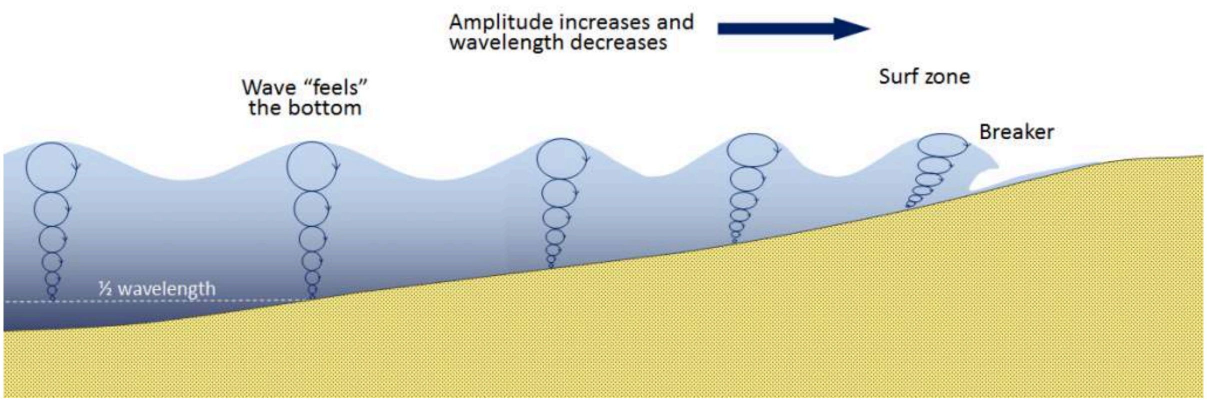


Figure 10.3.1 *As waves approach shore they "touch bottom" when the depth equals half of the wavelength, and the wave begins to slow down. As is slows, the wavelength decreases and the wave height increases, until the wave breaks (Steven Earle "Physical Geology").*

There are three main types of breakers: spilling, plunging, and surging. These are related to the steepness of the bottom, and how quickly the wave will slow down and its energy will get dissipated.

- **Spilling** breakers form on gently sloping or flatter beaches, where the energy of the wave is dissipated gradually. The wave slowly increases in height, then slowly collapses on itself (Figure 10.3.2). For surfers, these waves provide a longer ride, but they are less exciting.

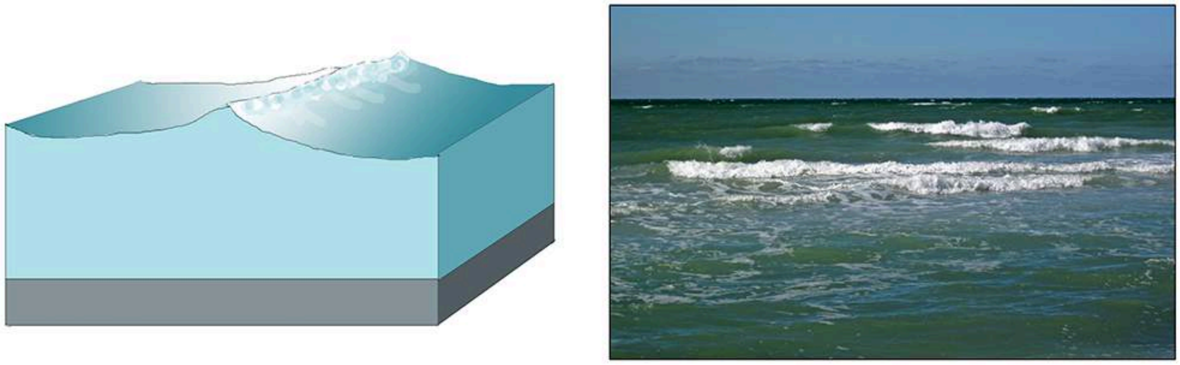


Figure 10.3.2 *A spilling breaker. The gentle slope of the bottom causes the wave height to slowly increase until the wave collapses on itself (left: JR, right: James St. John, [CC-BY-2.0], https://www.flickr.com/photos/jsjgeology/23769708334).*

- **Plunging** breakers form on more steeply-sloped shores, where there is a sudden slowing of the wave and the wave gets higher very quickly. The crest outruns the rest of the wave, curls forwards and breaks with a sudden loss of energy (Figure 10.3.3). These are the "pipeline" waves that surfers seek out.

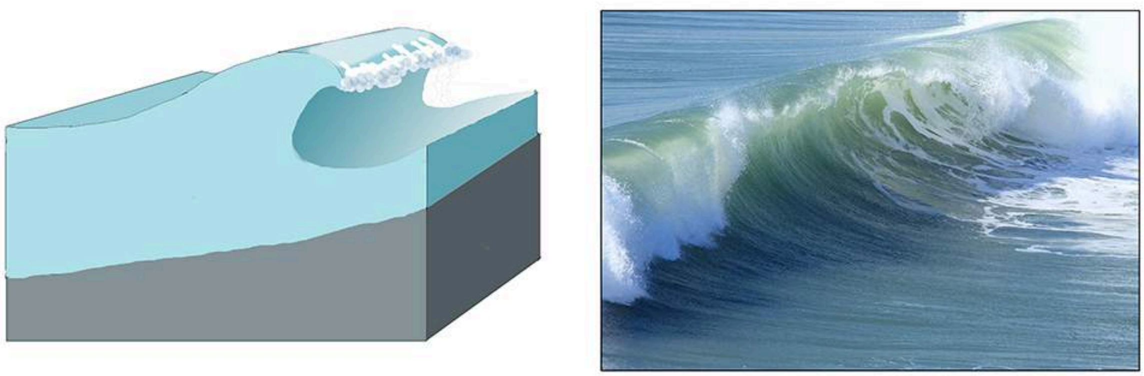


Figure 10.3.3 *A plunging breaker. The steeper slope causes the wave height to increase more rapidly, with the crest of the wave outrunning the base of the wave, causing it to curl as it breaks (left: JR, right: Andrew Schmidt, Public Domain [CC-0], publicdomainpictures.net).*

- **Surging** breakers form on the steepest shorelines. The wave energy is compressed very suddenly right at the shoreline, and the wave breaks right onto the beach (Figure 10.3.4). These waves give too short (and potentially painful) a ride for surfers to enjoy.

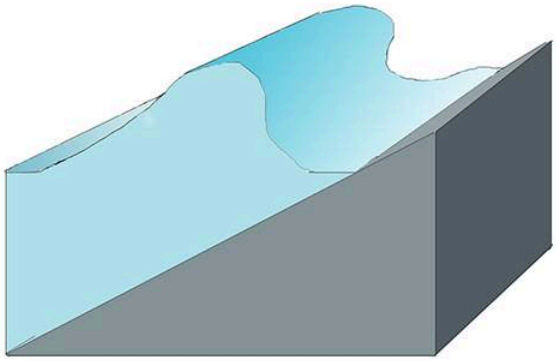


Figure 10.3.4 *A surging breaker. The very steep slope causes the wave height to increase suddenly and break right on the beach (left: JR, right: Tewy, [CC-BY-SA-3.0], via Wikimedia Commons).*

Wave Refraction

Swell can be generated anywhere in the ocean and therefore can arrive at a beach from almost any direction. But if you have ever stood at the shore you have probably noticed that the waves usually approach the shore somewhat parallel to the coast. This is due to wave refraction. If a wave front approaches shore at an angle, the end of the wave front closest to shore will touch bottom before the rest of the wave. This will cause that shallower part of the wave to slow down first, while the rest of the wave that is still in deeper water will continue on at its regular speed. As more and more of the wave front encounters shallower water and slows down, the wave font refracts and the waves tend to align themselves nearly parallel to the shoreline (they are refracted towards the region of slower speed). As we will see in section 13.2, the fact that the waves do not arrive perfectly parallel to the beach causes longshore currents and longshore transport that run parallel to the shore.

Refraction can also explain why waves tend to be larger off of points and headlands, and smaller in bays. A wave front approaching shore will touch the bottom off of the point before it touches bottom in a bay. Once again, the shallower part of the wave front will slow down, and cause the rest of the wave front to refract towards the slower region (the point). Now all of the initial wave energy is concentrated in a relatively small area off of the point, creating large, high energy waves (Figure 10.3.6). In the bay, the refraction has caused the wave fronts to refract away from each other, dispersing the wave energy, and leading to calmer water and smaller waves. This makes the large waves of a “point break” ideal for surfing, while water is calmer in a bay, which is where people would launch a boat. This difference in wave energy also explains why there is net erosion on points, while sand and sediments get deposited in bays (see section 13.3).

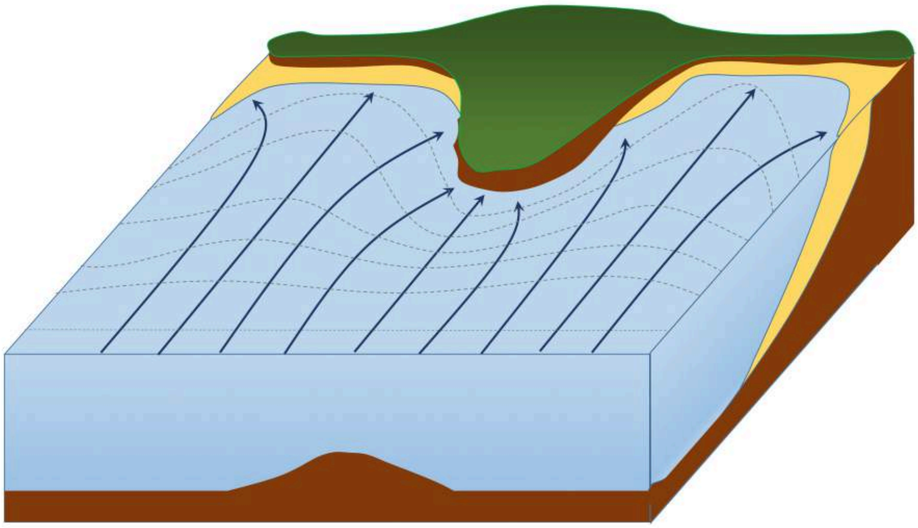


Figure 10.3.6 *Waves approaching shore (blue lines) touch bottom sooner off of points and are refracted towards the points, concentrating their wave energy. Wave energy is spread out in bays, causing smaller waves. Dotted lines represent the bottom contours (PW).*

10.4 Tsunamis

Tsunamis loom large in popular culture, but there are a number of misconceptions about these large waves. First, tsunamis have nothing to do with the tides, so it is a misnomer to refer to them as “tidal waves.” There are actual tidal waves (see section 11.1), but they are not related to tsunamis. Second, the giant, curling wave that is taller than skyscrapers and destroys cities in science fiction movies is also a fabrication, as tsunamis do not behave that way, as described below.

Tsunamis are large waves that are usually the result of seismic activity, such as the rising or falling of the seafloor due to earthquakes, although volcanic activity and landslides can also cause tsunamis in the form of splash waves (see section 10.1). As the seafloor rises or falls, so does the water column above it, creating waves. Only vertical seismic disturbances cause tsunamis, not horizontal movements. These vertical seafloor movements are usually less than 10 m high, so the resulting wave will be of an equal or lesser height at sea. While the tsunamis have a relatively small height at the point of origin, they have very long wavelengths (100-200 km). Because of the long wavelength, they behave as shallow water waves throughout the entire ocean; the depth of the ocean is always shallower than half of their wavelength. As shallow water waves, their speed depends on water depth, but they can still travel at speeds over 750 km/hr (Figure 10.4.1)!

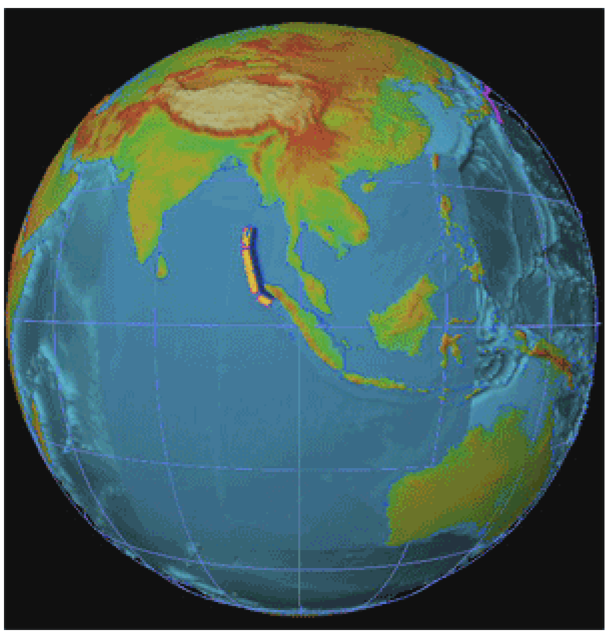


Figure 10.4.1 *Animation of the spread of tsunamis created during the 2004 Indonesia earthquake (NOAA Center for Tsunami Research (NCTR) [Public domain]).*

When tsunamis approach land, they behave just like any other wave; as the depth becomes shallower, the waves slow down and the wave height begins to increase. However, contrary to popular belief, tsunamis do not arrive on shore as giant, cresting waves. Since their wavelength is so long, it is impossible for their height to ever exceed 1/7 of their wavelength, so the waves don’t actually curl or break. Instead, they usually hit the shore as sudden surges of water causing a very rapid increase in sea level, like that of an enormous rise in tide. It may take several minutes for the wave to pass, during which time sea level can rise to 40 m higher than usual.

Large tsunamis occur every 2-3 years, with very large, damaging events happening every 15-20 years. The most devastating tsunami in terms of loss of life resulted from a magnitude 9 earthquake in Indonesia in 2004 (Figure 10.4.2), which created waves up to 33 m tall and left about 230,000 people dead in Indonesia, Thailand, and Sri Lanka. In 2011 a 9.0 magnitude earthquake in Japan triggered a tsunami up to 40.5 m high, which resulted in over 18,000 deaths. This earthquake also caused the Fukishima nuclear accident, and moved Japan about 8 inches closer to the U.S.

Figure 10.4.2 *A village in Sumatra following the Indonesia Tsunami in December 2004 (U.S. Navy photo by Photographer's Mate 2nd Class Philip A. McDaniel [Public domain via Wikimedia Commons]).*

CHAPTER 11: TIDES

Chapter 11: Tides

Learning Objectives

After reading this chapter you should:

- understand that tides are just very long waves, with crests and troughs
- understand Newton's Law of Universal Gravitation and how it applies to tides
- understand why most places on Earth experience two tides per day, not just the one predicted from gravitational attraction between the Earth and moon (i.e. inertial force)
- understand how the Earth, sun and moon interact to create spring and neap tides
- understand why the gravitational pull of the sun on tides is less than the pull of the moon
- understand why tides do not occur at the same time every day
- understand why amphidromic circulation occurs as a result of tides
- know the difference between diurnal, semi-diurnal, and mixed tides
- know the phases of a tidal current
- know what causes a tidal bore

The previous chapter discussed various types of waves at sea and along the shore. However, at least in terms of wavelength, the largest waves in the ocean are the tides, where one wavelength stretches halfway around the Earth. The crests of these long waves represent the high tides, while the troughs create low tides.

You probably learned when you were younger that the basic cause of the tides is the gravitational attraction between the Earth and moon. This is a very old idea, as the Greek scientist Pytheas first made the connection between the tides and the moon back in 330 B.C.E. Isaac Newton's gravitational work the 1600s led to our modern understanding of tidal cycles, however, we now know that the tides involve a lot more than just the Earth and the moon. There are many variables that influence the tides, yet despite this complexity, we are able to create accurate tide charts predicting the heights and timing of tides months or even years in advance.

11.1 Tidal Forces

Our modern understanding of tide formation stems from Isaac Newton's Law of Universal Gravitation, which states that any two objects have a gravitational attraction to each other. The magnitude of the force is proportional to the masses of the objects, and inversely proportional to the square of the distance between the objects, according to the equation in Figure 11.1.1.

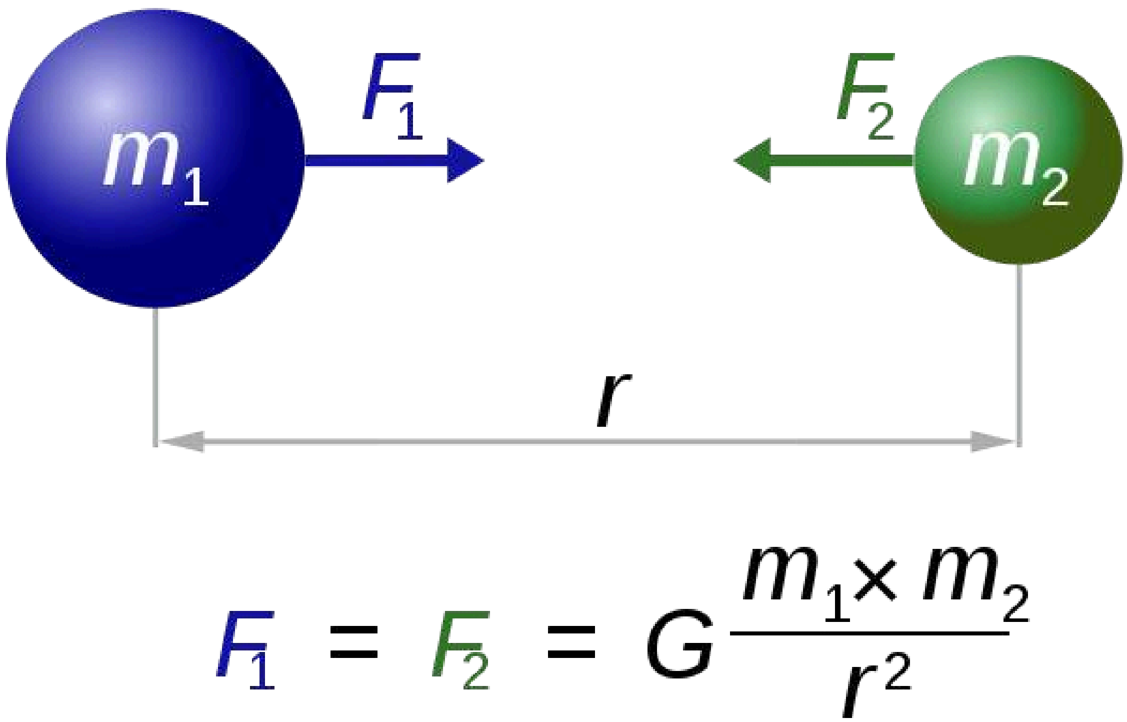


Figure 11.1.1 *Newton's Law of Universal Gravitation. The gravitational force between two objects (F) is calculated as the product of the two masses (m_1 and m_2) divided by the distance between them (r) squared. G is the universal gravitational constant; 6.67408 × 10^{-11} m^3 kg^{-1} s^{-2} (By I, Dennis Nilsson [CC BY 3.0], via Wikimedia Commons).*

In the case of tides, there are a few other factors that modify this equation so that the distance (r) is cubed rather than squared, giving distance an even greater impact on tidal forces. But for our purposes, the important lesson is that the greater the masses of the objects, the greater the gravitational force, and the farther the objects are from each other, the weaker the force.

Such a gravitational force exists between the Earth and moon, attempting to pull them towards each other. Since the water covering Earth is fluid (unlike the solid land that is more resistant to tidal forces), this gravitational force pulls water towards the moon, creating a "bulge" of water on the side of the Earth facing the moon (Figure 11.1.2). This bulge always faces the moon, while the Earth rotates through it; the regions of Earth moving through the bulge experience a high tide, while those parts of the Earth away from the bulge experience a low tide.

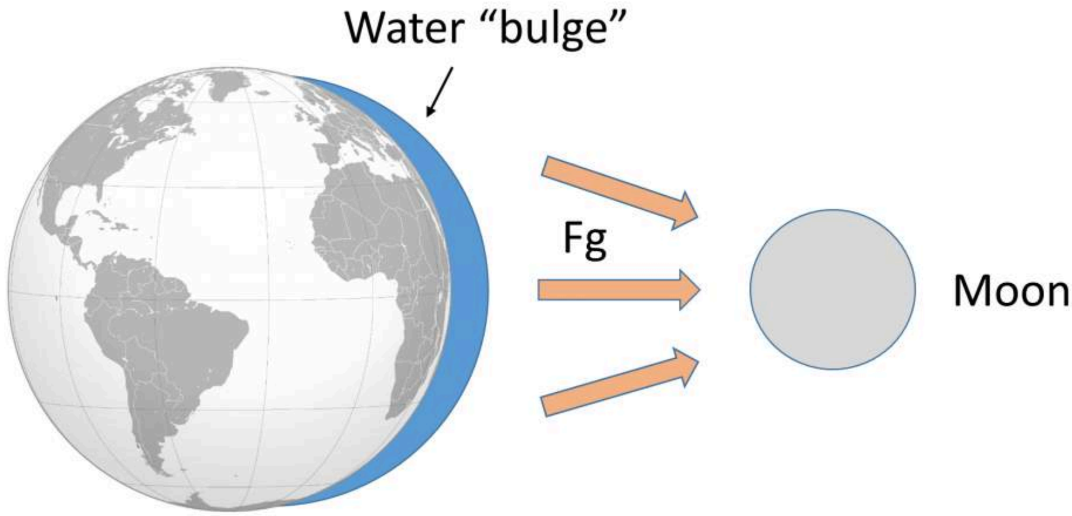


Figure 11.1.2 *Gravitational forces between the Earth and moon cause a bulge of water to appear on the side of the Earth facing the moon (PW).*

If the tides were this simple, everywhere on Earth would see one high tide per day, as there would only be a bulge of water on the side closest to the moon. However, if you have ever looked at tide charts, or lived near the ocean, you probably know that in most places there are two high tides and two low tides per day. Where is this second high tide "bulge" coming from?

The gravitational force between the Earth and moon might be expected to draw the two objects closer together, however, this is not happening. This is because the inward gravitational force is opposed by outward forces that keep the Earth and moon apart. The outward force is an **intertial force** created by the rotation of the Earth and moon. Contrary to popular belief, the moon is not simply rotating around the Earth; in fact, the Earth and moon are both rotating around each other. Imagine the Earth and moon as equal-sized objects revolving around a point at their center of mass. If both objects had the same mass, the center of rotation would be a point equidistant between the two objects. But since the mass of the Earth is 82 times greater than the mass of the moon, the center of revolution must be closer to the Earth. As an analogy, think about two people on a see-saw. If the people are of roughly equal size, they can sit on either end of the see-saw at it will rotate around a point at equal distance between them. But if the two people have very different masses, such as a large adult and a small child, the larger person must move closer to the pivot point for the see-saw to rotate effectively. In the same way, the center of rotation between the Earth and the moon (the **barycenter**) must be located closer to the Earth. In fact, the center of rotation lies *within* the Earth, about 1600 km below the surface. As the Earth and moon rotate around the barycenter, the moon travels much farther than the Earth, giving the impression that the moon is rotating around Earth (Figure 11.1.3).

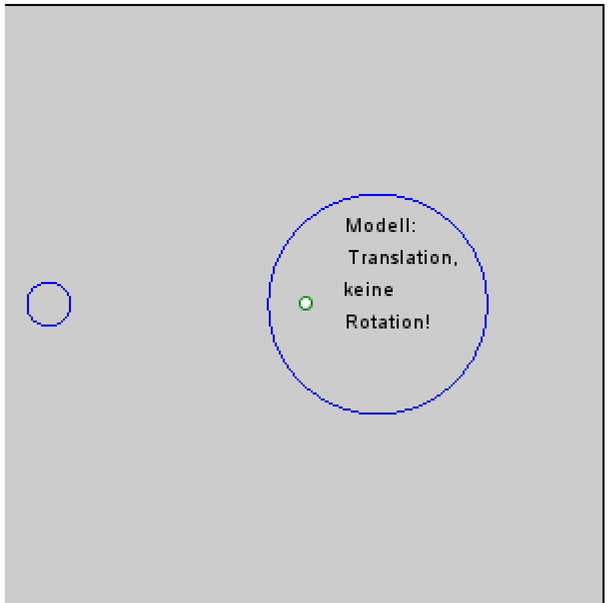


Figure 11.1.3 *Rotation of the Earth and moon around the barycenter (white dot). Because of its larger mass, the center of rotation lies closer to the Earth than to the moon (By Modalanalytiker (Own work) [CC BY-SA 3.0 or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons).*

The rotation of the Earth-moon system creates an outward inertial force, which balances the gravitational force to keep the two bodies in their orbits. The inertial force has the same magnitude everywhere on Earth, and is always directed away from the moon. Gravitational force, on the other hand, is always directed towards the moon, and is stronger on the side of the Earth closest to the moon. Figure 11.1.4 describes how these forces combine to create the tidal forces. At point O in the center of the Earth, the gravitational force (F_g) and the inertial force (F_r) are equal, and cancel each other out. On the side of Earth closest to the moon, the inward gravitational force (F_g) is greater than the outward inertial force (F_r); the net resulting force (A) is directed towards the moon, and creates a bulge of water on the side facing the moon. On the side of Earth opposite the moon, the outward inertial force is greater than the inward gravitational force; the net resulting force (C) is directed away from the moon, creating a water bulge directed away from the moon.

Now, as the Earth rotates through a 24 hour day, each region passes through two bulges, and experiences two high tides and two low tides per day. This represents Newton's **Equilibrium Theory of Tides,** where there are two high tides and two low tides per day, of similar heights, each six hours apart. But as with everything else in oceanography, reality is much more complex than this idealized situation.

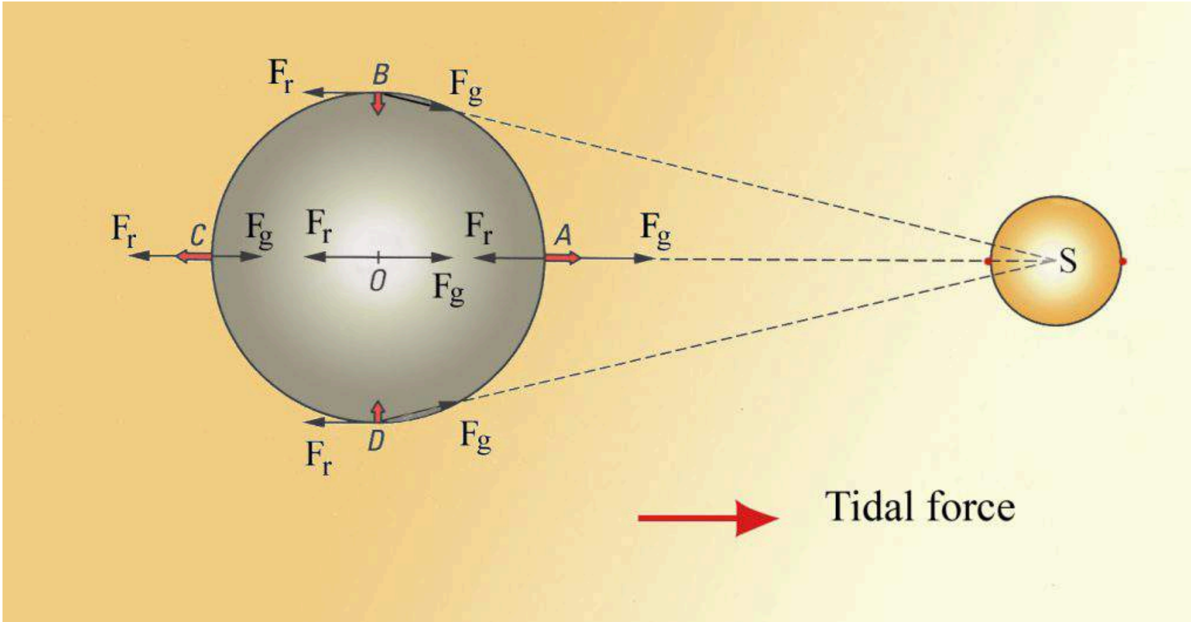


Figure 11.1.4 *Gravitational force (F_g) are strongest closer to the moon and weaker opposite the moon. Inertial forces (F_r) are equal throughout the Earth and directed away from the moon. The tidal forces A and C are the result of the interaction between F_g and F_r and create water bulges on both sides of the Earth, leading to two high tides per day (Vitold Muratov (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

Some of the additional complexity is because in addition to the moon, the sun also exerts tide-affecting forces on Earth. The solar gravitational and inertial forces arise for the same reasons described above for the moon, but the magnitudes of the forces are different. The sun is 27 million times more massive than the moon, but it is 387 times farther away from the Earth. Despite its larger mass, because the sun is so much farther away than the moon, the sun's gravitational forces are only about half as strong as the moon's (remember that distance is cubed in the gravity equation). The sun thus creates its own, smaller water bulges, independent of the moon's, that contribute to the creation of tides.

When the sun, Earth and moon are aligned, as occurs during new and full moons, the solar and lunar bulges are also aligned, and add to each other (constructive interference; see section 10.2) creating an especially high tidal range; high high tides and low low tides (Figure 11.1.5). This period of maximum tidal range is called a **spring tide**, and they occur every two weeks.

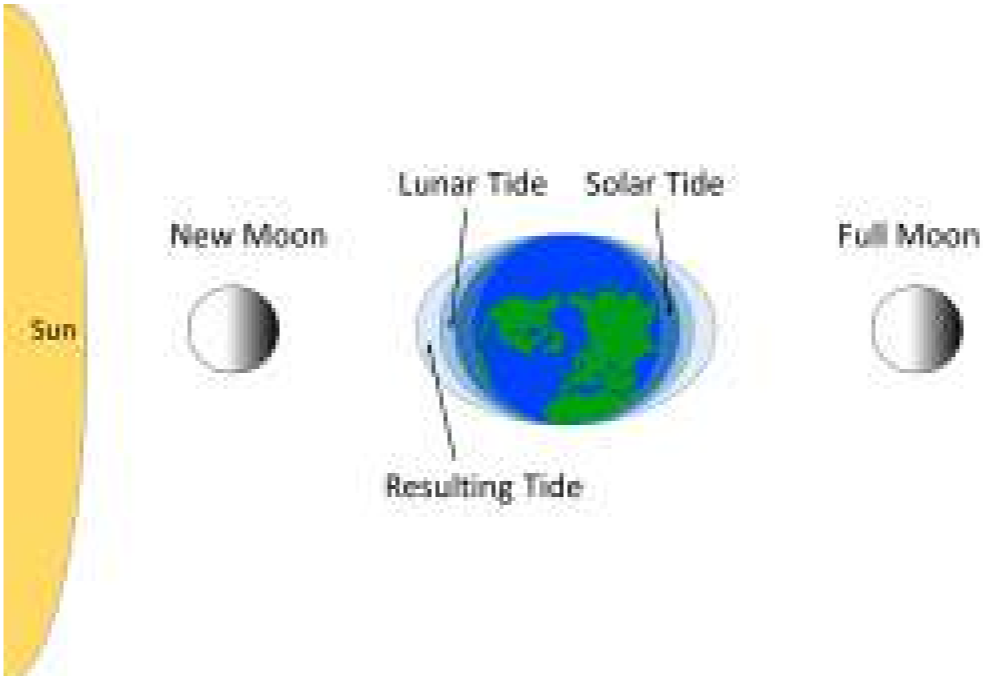


Figure 11.1.5 *Spring tides with high tidal ranges occur when the solar and lunar tides are added together during full and new moons when the Earth, sun and moon are aligned (PW).*

When the sun, Earth and moon are at 90^{o} to each other, the solar and lunar bulges are out of phase, and cancel each other out (destructive interference). Now the tidal range is small, with low high tides and high low tides (Figure 11.1.6). These are **neap tides**, and occur every two weeks, when the moon is in its 1/4 and 3/4 phases (Figure 11.1.7).

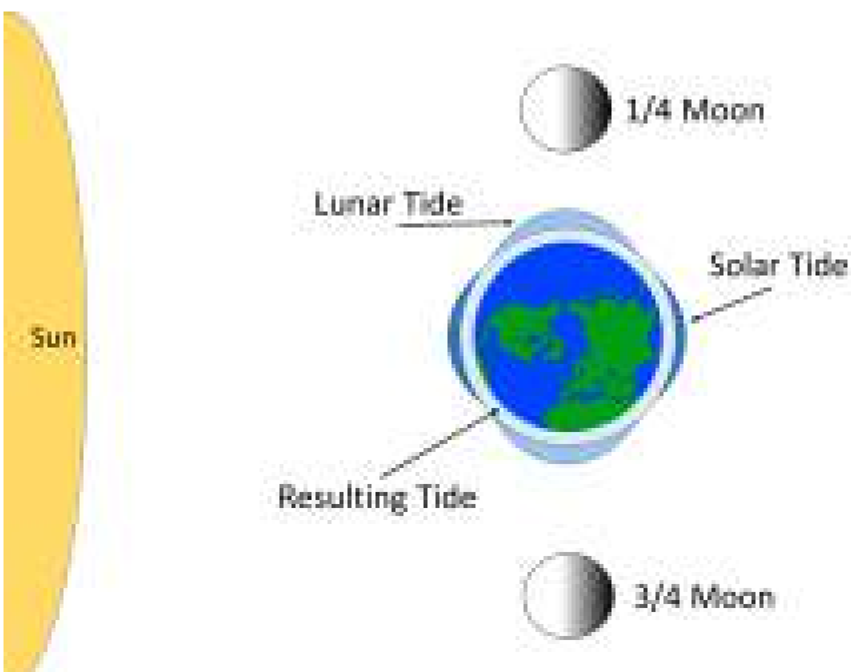


Figure 11.1.6 *Neap tides are created during 1/4 and 3/4 moons when the Earth, sun and moon are perpendicular to each other. The solar and lunar tides cancel each other out, resulting in a small tidal range (PW).*

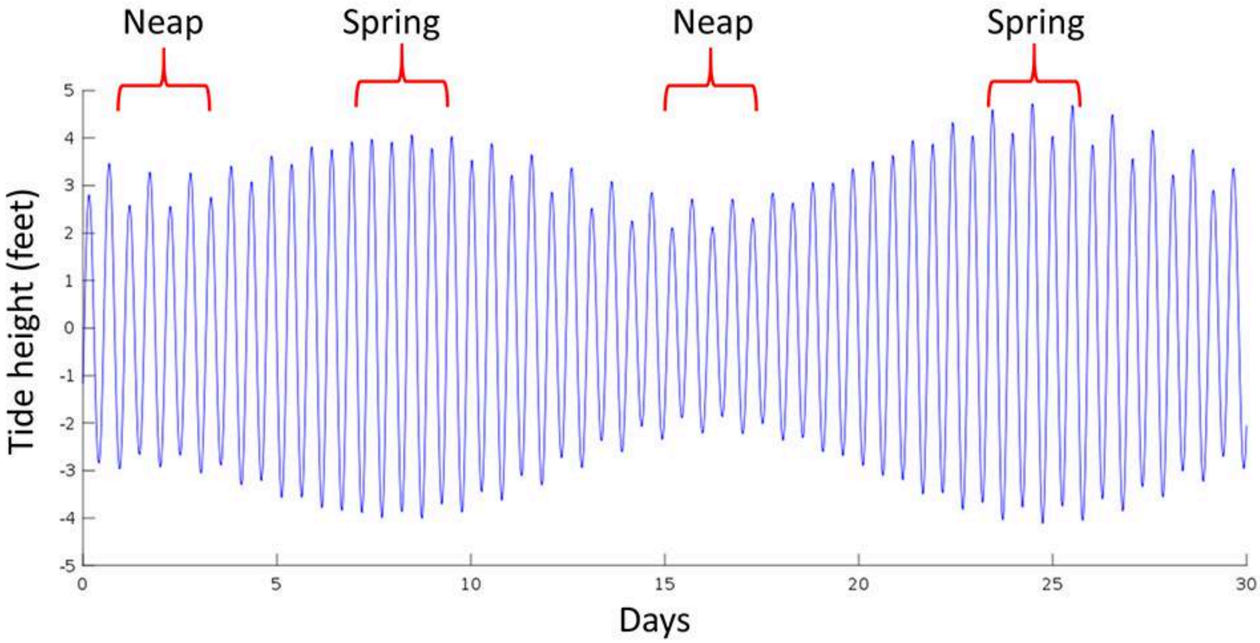


Figure 11.1.7 *30 days of tidal data from Bridgeport, CT, USA, showing spring and neap tidal ranges (Modified by PW from Cody Logan (clpo13), Public Domain via Wikimedia Commons).*

11.2 Dynamic Theory of Tides

The Equilibrium Theory of tides predicts that each day there will be two high and two low tides, each one occurring at the same time day after day, with each pair producing tides of similar heights. While this view provides a basic explanation for the primary forces that generate the tides, it does not take into account such variables as the effects of the continents, the depth of the water, and many other factors. In all, there are almost 400 variables that must be incorporated into predicting the tides! The **Dynamic Theory** of tides takes these other factors into account, and shows that the tides are much more complicated and variable from place to place than the Equilibrium Theory would suggest. For example, some areas receive only one high and one low tide per day (see section 11.3). Furthermore, the tidal range varies greatly across the globe; in the Mediterranean Sea, there can be a difference of only 10 cm between high and low tides, while the Bay of Fundy in Canada experiences a tidal range of up to 17m (56 ft) every day (Figure 11.2.1).

Figure 11.2.1 *Tidal range in the Bay of Fundy, Canada. Both photographs were taken on the same day in July 2003 (By Dylan Kereluk from White Rock, Canada (Flickr) [CC BY 2.0], via Wikimedia Commons).*

Examination of any tide chart will show that the tides don't occur at same time each day; in fact, each tidal peak occurs about 50 minutes later than it did in the previous day. This is due to the orbit of the moon around the Earth. Imagine a high tide that occurs at a particular location (X) at 1:00 pm (Figure 11.2.2). The high tide occurs as location X moves through the bulge of water facing the moon. It will take the Earth 24 hours to complete one revolution, to bring location X back to site of the water bulge that caused that high tide. However, during those 24 hours, the moon has also moved as it orbits the Earth, so the high tide bulge has moved beyond its original location. The Earth thus has to rotate an additional distance for location X to reach the bulge and experience that same high tide. Because it takes the moon about 28 days to orbit the Earth, the moon gets "ahead" of the Earth's rotation by about 50 minutes per day. Therefore, it takes location X 24 hours and 50 minutes to rotate through the same tidal bulge, and as a result, the tidal peaks occur about 50 minutes later each day. In our example, an afternoon high tide at 1:00 pm on one day would be followed by a high tide at about 1:50 pm the following day. This 24 hour and 50 minute cycle is referred to as a **tidal day**.

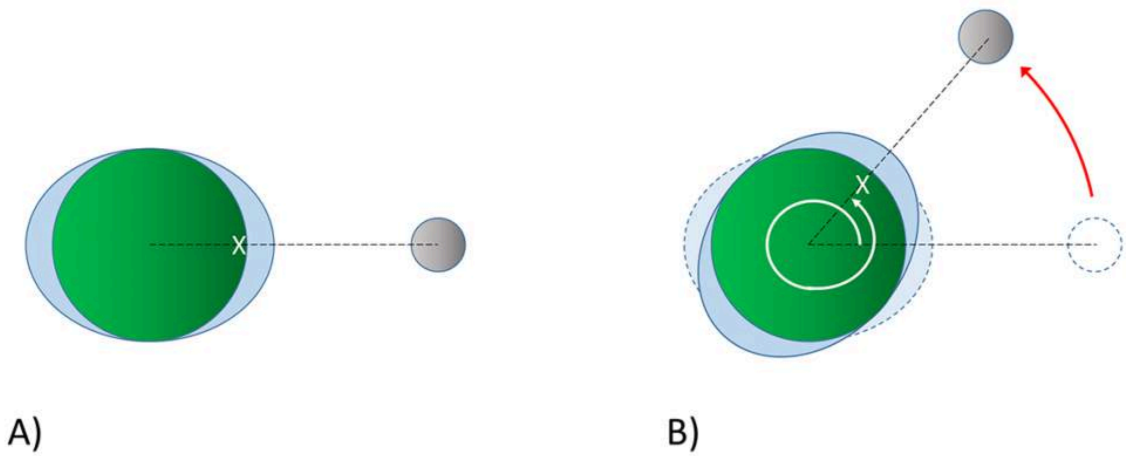


Figure 11.2.2 *In A) a high tide is occurring at point X on Earth's surface. 24 hours later, X has made a complete rotation and is back in its original position. However, the moon has moved during that time (B), so X must travel an additional distance (white arrow) to once again become aligned with the moon and experience a high tide. For this reason, corresponding tides occur approximately 60 minutes later each day (PW).*

The motion of the moon impacts the tidal cycles in other ways. As the moon orbits the Earth, its orbital plane is at an angle relative to the rotational plane of Earth. This angle, or **declination**, means than the moon fluctuates between an angle of 28.5^{o} north of the equator, to 28.5^{o} south of the equator roughly every two weeks (the cycle from maximum to minimum and back takes about 27 days). Figure 11.2.3 illustrates a case where the moon is at its maximum declination 28.5^{o} north of the equator, creating its corresponding tidal maxima. A point on the Earth at the latitude indicated by the red line would experience two high tides as it rotated through 24 hours, at points A and B. But the two high tides would not be of equal heights; the high tide at A would be higher than the high tide at B. This helps create a mixed semi-diurnal tide; two high tides of different heights per day (see section 11.3).

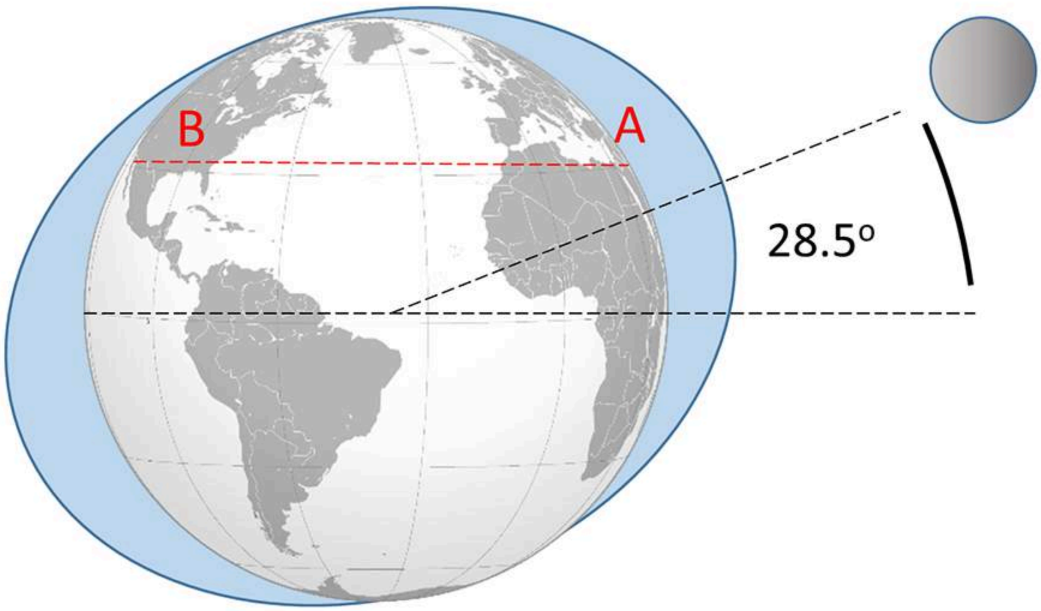


Figure 11.2.3 *The effect of the moon declination on tide heights. The moon oscillates between 28.5° north and 28.5° south of the equator every two weeks, leading to uneven tidal heights each day at a particular latitude (PW).*

Finally, the continents and the bottom topography of the oceans have an impact on the tides that are experienced in an area. Because the tides are essentially waves with extremely long wavelengths extending halfway across the Earth, they behave as shallow water waves, and they are influenced and refracted by the bottom contours, leading to regional tidal variations. When the tidal crests encounter land, they are are reflected, and the wave moves back out to sea, theoretically until it encounters another continent on the opposite side of the ocean basin. The crest is once again reflected, and the water oscillates back and forth as a standing wave across the ocean basin. However, because of the scale over which these tidal waves move, we must take into account the influence of the Coriolis Effect. As the tidal crest is reflected back across the ocean basin, its path is deflected by the Coriolis force; to the right in the Northern Hemisphere, and to the left in the Southern Hemisphere (see section 9.1). Using the Northern Hemisphere as an example, imagine a tidal crest that has reached land on the western side of an ocean basin. It would have a tendency to be reflected and move across the basin towards the east. But the Coriolis force deflects the movement to the right, causing the crest to instead head south. When the crest hits land in the south, it would now tend to reflect towards the north, but once again the Coriolis deflection to the right kicks in, and the wave instead moves to the east. From the east the reflected wave is deflected to the north, and so on. The result of all of this is that instead of a simple standing wave moving back and forth across the ocean, the tidal crest follows a circular pattern around the ocean basin, counterclockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere. This is analogous to shaking a pan full of water in a circular manner, and watching the water follow a similar circular path as it sloshes around inside. This large scale circular rotation pattern of tides is called **amphidromic circulation** (Figure 11.2.4). The rotation occurs around a central **amphidromic point** or **node**, that shows little tidal variation, while the largest tidal ranges occur on the edges of the circulation pattern. In Figure 11.2.4 the amphidromic points are indicated by the dark blue areas where the white lines converge, like spokes from a bicycle wheel, and the dark red and brown areas show the regions of maximum tidal heights. The tidal maxima will rotate around the amphidromic points, taking about 12 hours for a complete rotation, leading to two high and two low tides per day in many places. If a tidal maximum is occurring along one of the white lines in Figure 11.2.4 at a certain time in the Northern Hemisphere, one hour later that high tide will have moved to the white line to the left (counterclockwise), and so on until it completes a rotation. In the Southern Hemisphere, the tide will move to the line to the right for clockwise rotation.

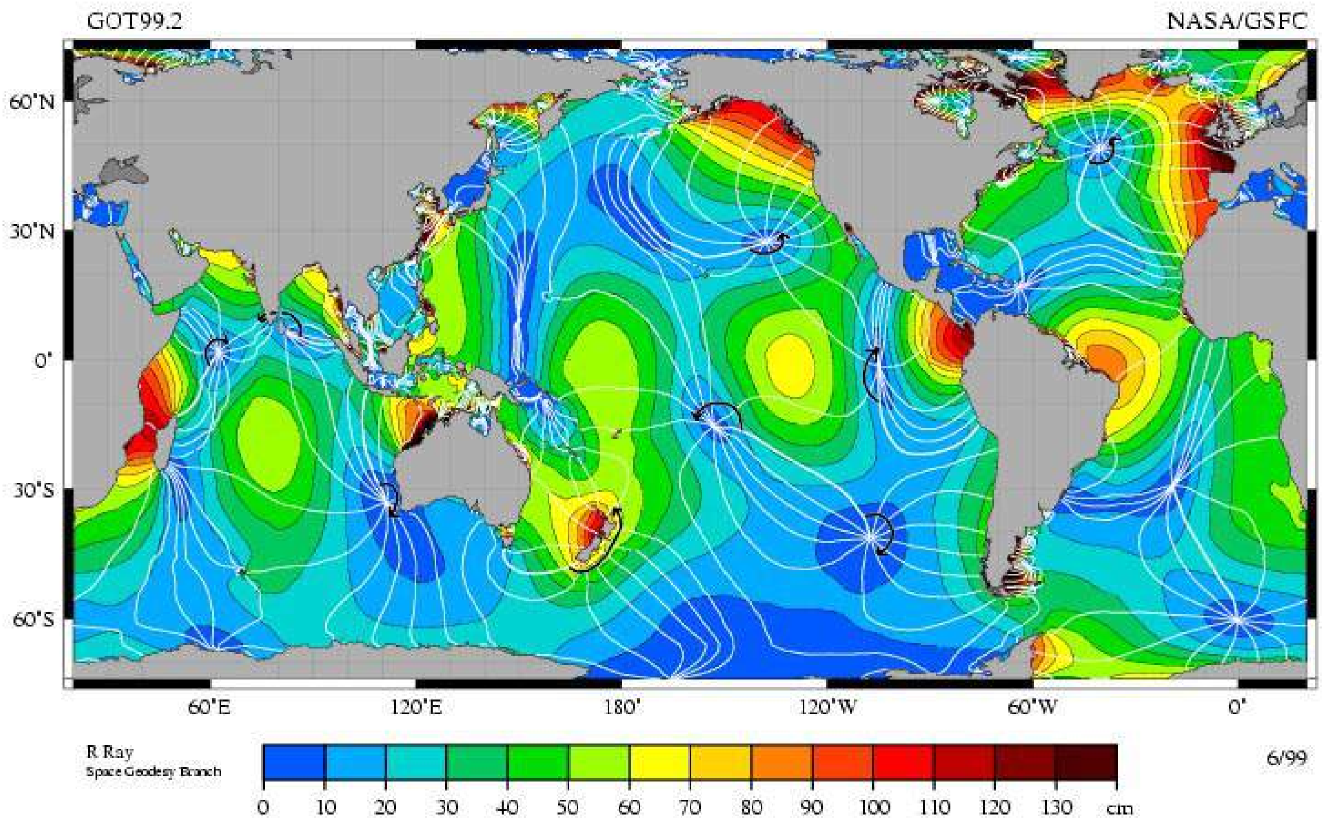


Figure 11.2.4 *Amphidromic circulation. Amphidromic points are represented where the white lines converge in areas of minimal tidal range. Tidal crests rotate around the amphidromic points, clockwise in the Northern Hemisphere and counterclockwise in the Southern Hemisphere. See text for details (NASA, Public Domain via Wikimedia Commons).*

The result of all of these variables is that the tides will not always occur twice each day, at the same time and with equal heights as the Equilibrium Theory of tides may suggest. Instead, each region of the oceans has a unique set of factors that contribute to the types of tides it will experience. The major types of tides are discussed in the next section.

11.3 Tide Classification

With so many variables playing a role in the production of tides, it is understandable that not every place on Earth will experience exactly the same tidal conditions. There are three primary classifications for tides, depending on the number and relative heights of tidal cycles per day.

A **diurnal tide** consists of only one high tide and one low tide per day (Figure 11.3.1). “Diurnal” refers to a daily occurrence, so a situation where there is only one complete tidal cycle per day is considered a diurnal tide. Diurnal tides are common in the Gulf of Mexico, along the west coast of Alaska, and in parts of Southeast Asia.

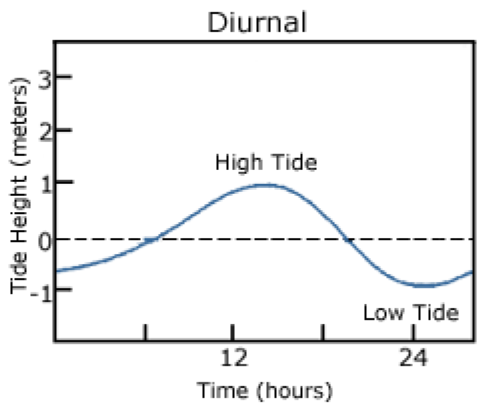


Figure 11.3.1 *A diurnal tide, with one high and one low tide per day (By NOAA [Public domain], via Wikimedia Commons).*

A **semidiurnal tide** exhibits two high and two low tides each day, with both highs and both lows of toughly equal height (Figure 11.3.2). “Semidiurnal” means “half of a day”; one tidal cycle takes half of a day, therefore there are two complete cycles per day. Semidiurnal tides are common along the east coasts of North America and Australia, the west coast of Africa, and most of Europe.

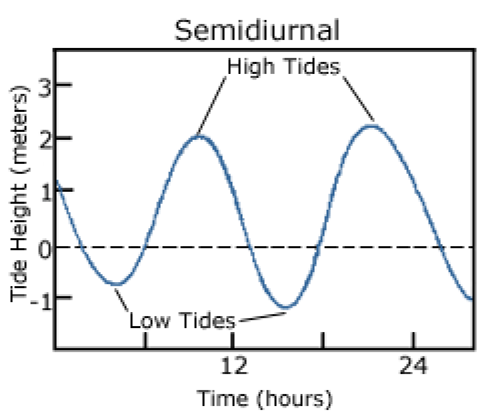


Figure 11.3.2 *A semi-diurnal tide, with two high and two low tides per day, each of roughly equal heights (By NOAA [Public domain], via Wikimedia Commons).*

Mixed semidiurnal tides (or **mixed tides**), have two high tides and two low tides per day, but the heights of each tide differs; the two high tides are of different heights, as are the two low tides (Figure 11.3.3). The differences in height may be the result of amphidromic circulation, the angle of the moon, or any of the other variables discussed in section 11.2. Mixed semidiurnal tides are found along the Pacific coast of North America.

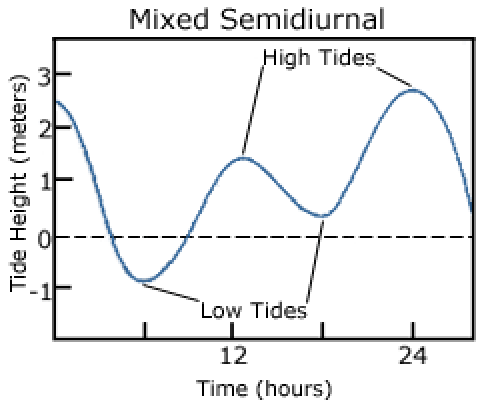


Figure 11.3.3 *A mixed semi-diurnal tide, with two high and two low tides per day, each with a different height (By NOAA [Public domain], via Wikimedia Commons).*

Figure 11.3.4 shows the distribution of the various tide types throughout the world.

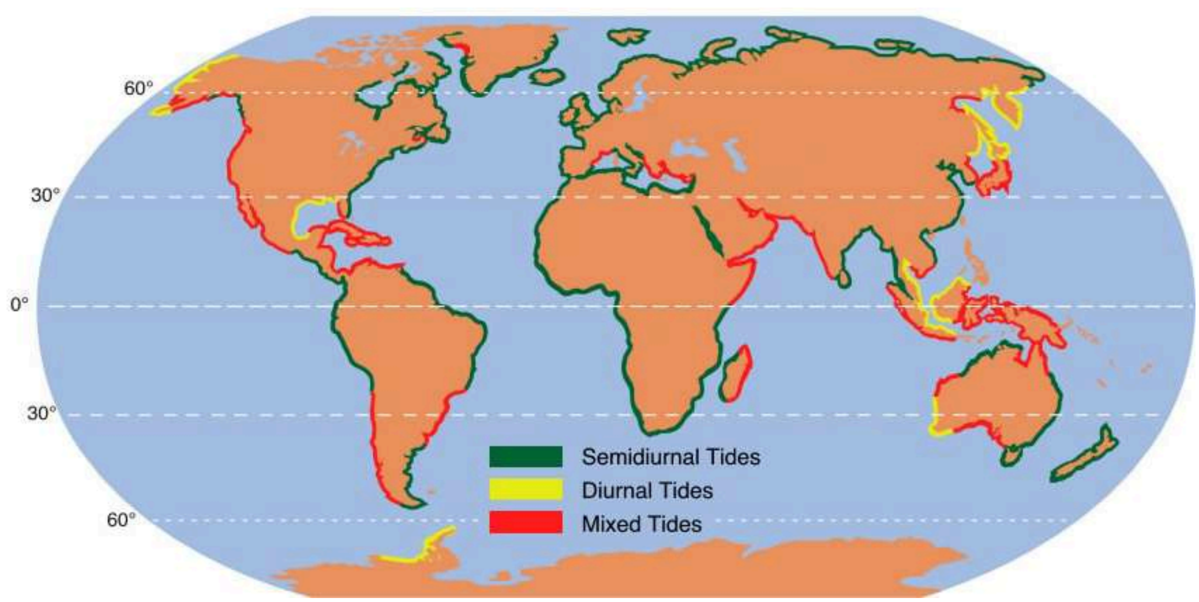


Figure 11.3.4 *Global distribution of the different types of tides (By KVDP (Own work) [Public domain], via Wikimedia Commons).*

Tidal Currents

The movement of water with the rising and falling tide creates tidal currents. As the tide rises, water flows into an area, creating a **flood current**. As the tide falls and water flows out an **ebb current** is created. **Slack water**, or **slack tides** occur during the transition between incoming high and outgoing low tides, when there is no net water movement.

The strength of a tidal current depends on the volume of water that enters and exits with each tidal cycle (the **tidal volume** or **tidal prism**), and the area through which the water flows. A large tidal volume moving through a large area may create only a weak tidal current, as the volume is spread over a wide area. On the other hand, a narrow area may produce a strong tidal current even if the tidal volume is small, as all of the water is forced through a small area. It follows that the strongest tidal currents will result from a large tidal range moving through a narrow area.

Tidal bores occur where rivers meet the ocean. If the incoming tidal current is stronger than the river outflow, the tidal bore appears as a wave, or moving wall of water that moves up the river as the tide comes in (Figure 11.3.5).

Figure 11.3.5 *A tidal bore near Silverdale in the United Kingdom (Arnold Price [CC BY-SA 2.0], via Wikimedia Commons).*

In many cases these tidal bores may move through a river or inlet for many kilometers, and if they are large enough they can form continually breaking waves that surfers can ride much farther and longer than a traditional ocean wave, such as the Severn Bore in England, shown in the video below.

A YouTube element has been excluded from this version of the text. You can view it online here: https://rwu.pressbooks.pub/webboceanography/?p=347

Additional links for more information

- For an even more dramatic tidal bore, watch this video of the "Silver Dragon" on China's Qiantang River

CHAPTER 12: OCEAN SEDIMENTS

Chapter 12: Ocean Sediments

Learning Objectives

After reading this chapter you should:

- know how sediments are classified based on physical characteristics (size, sorting etc.)
- identify the four main sources of marine sediments
- differentiate between organisms that produce different biogenous sediments
- understand the factors that determine the distribution of sediment types in the ocean
- know the different ways sediment samples can be obtained
- understand how biogenous sediments can be used to reconstruct past climate change

Let's be honest; for the majority of people with an interest in the oceans and oceanography it is not the allure of the sediments that first grabs their attention. At first glance the muddy seafloor may not seem that interesting, but the sediments play a vital role in marine ecosystems and our understanding of ocean and geological processes. The sediments provide habitat for a multitude of marine organisms, and they contain information about past climates, plate tectonics, ocean circulation patterns, and the timing of major extinctions, just to name a few. In this chapter we will examine the major types of sediments, and their distribution on the ocean floor.

12.1 Classifying Sediments

The term "sediment" refers to the tiny particles of rocks and other materials that sink to the ocean floor and eventually settle and accumulate on the bottom. All regions of the seafloor contain some form of sediment, although there are many different types of sediments from a variety of sources, and the amount of accumulated sediment can vary greatly from place to place. Globally, ocean sediments average about 1 km thick, but they can exceed 15 km thick in areas of high accumulation (Figure 12.1.1). These areas include regions near the mouths of rivers where there is high sediment discharge, and passive margins near the continents where the seafloor has had millions of years for sediment to accumulate. On the other hand, sediments are sparse along divergent plate boundaries where new oceanic crust is being formed, as the crust is too new for significant accumulation (see section 4.5), and in the central oceans that are far away from any significant sediment sources.

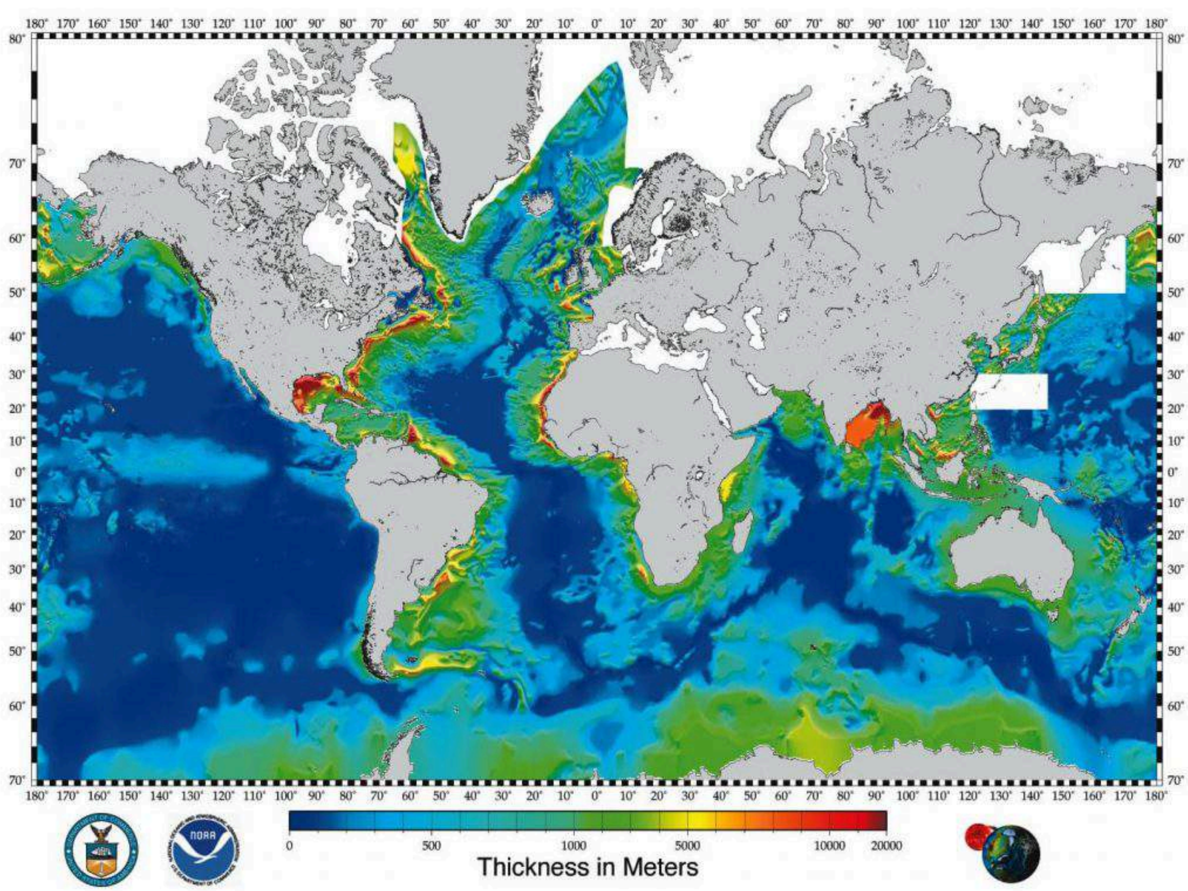


Figure 12.1.1 *Total sediment thickness of the world ocean (By Divins, D.L., NGDC Total Sediment Thickness of the World's Oceans & Marginal Seas, http://www.ngdc.noaa.gov/mgg/sedthick/sedthick.html (http://www.ngdc.noaa.gov/mgg/image/sedthick9.jpg) [Public domain], via Wikimedia Commons).*

As time passes over millions of years, these sediments can become **lithified** or turned into sedimentary rock. It has been estimated that over half of the exposed rock on the continents is sedimentary rock originally deposited in ancient oceans and uplifted by plate tectonics. Many tall mountains, including Mt. Everest, are composed of rock formations that contain fossils of marine creatures. These rocks were originally formed as ocean sediments which were then lithified and pushed upwards during the process of mountain formation.

Sediment Classification

There are a number of ways that we can classify ocean sediments, and some of the most common distinctions are based on the sediment texture, the sediment composition, and the sediment's origin.

Texture

Sediment texture can be examined through several variables. The first is **grain size**. Sediments are classified by particle size, ranging from the finest clays (diameter <0.004 mm) to the largest boulders (> 256 mm)(Figure 12.1.2). Among other things, grain size represents the conditions under which the sediment was deposited. High energy conditions, such as strong currents or waves, usually results in the deposition of only the larger particles as the finer ones will be carried away. Lower energy conditions will allow the smaller particles to settle out and form finer sediments.

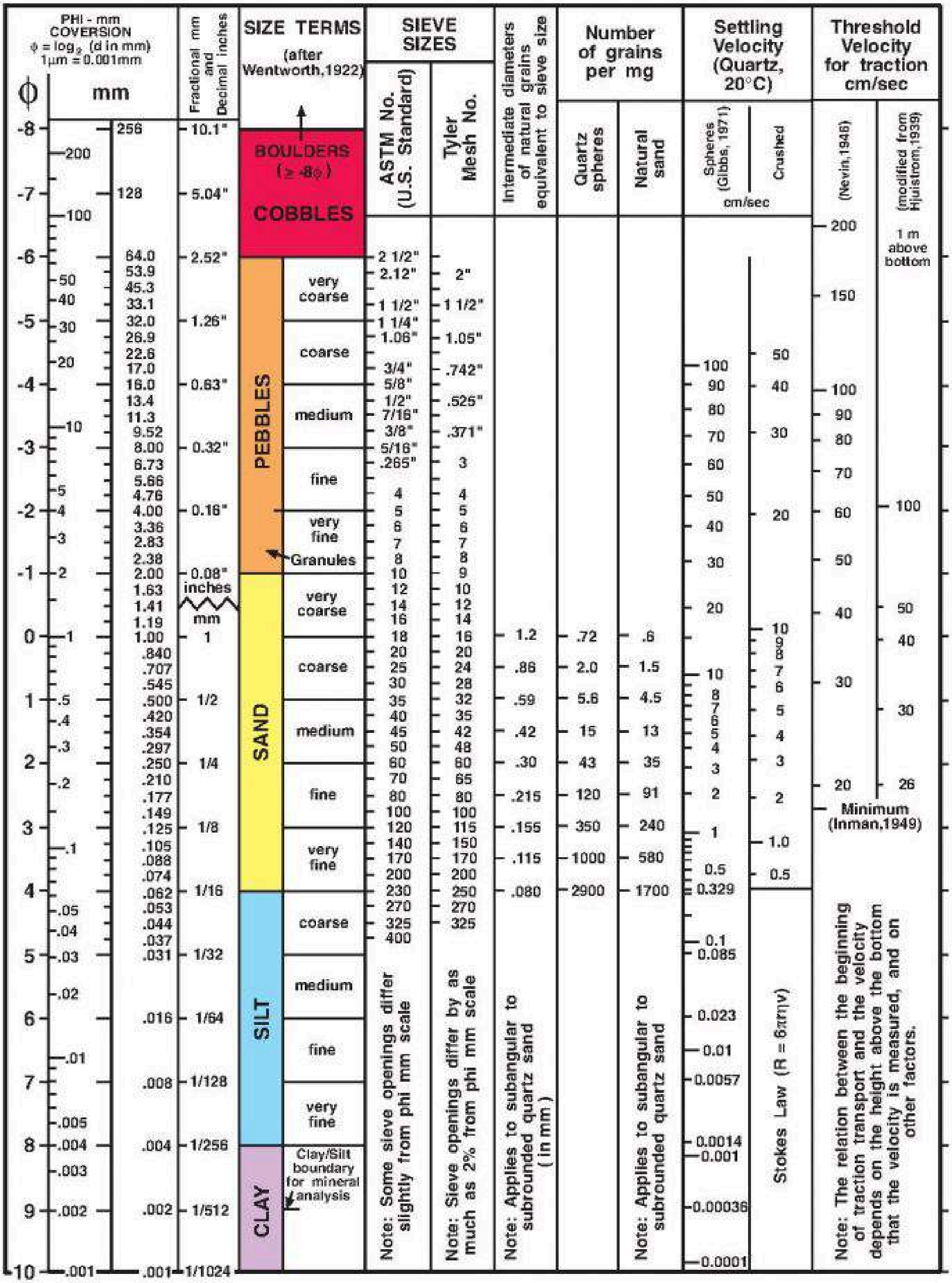


Figure 12.1.2 *Wentworth grain size chart for classifying sediments (By Jeffress Williams, Matthew A. Arsenault, Brian J. Buczkowski, Jane A. Reid, James G. Flocks, Mark A. Kulp, Shea Penland, and Chris J. Jenkins, USGS [Public domain], via*

Wikimedia Commons).

Sorting is another way to categorize sediments. Sorting refers to how uniform the particles are in terms of size (Figure 12.1.3). If all of the particles are of a similar size, such as in beach sand, the sediment is well-sorted. If the particles are of very different sizes, the sediment is poorly sorted, such as in glacial deposits.

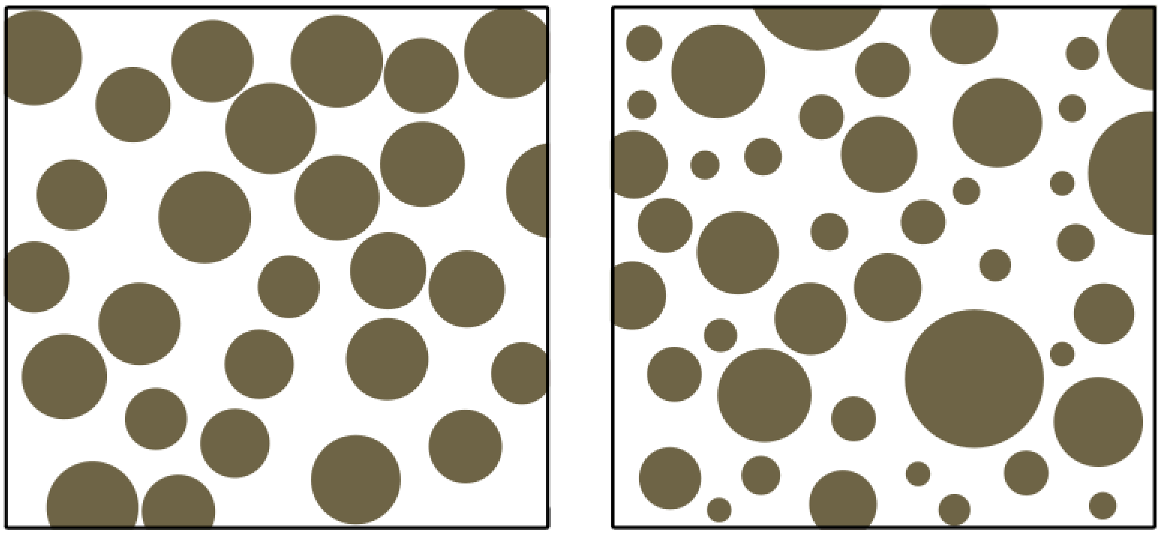


Figure 12.1.3 *Well-sorted sediments (left) have particles that are all of a similar size. Poorly sorted sediments (right) consist of particles of a wide range of sizes (Woudloper (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

Finally, sediment texture can be described based on its **maturity**, or how long the particles have been transported by water or other vectors. Maturity can be reflected by the degree of rounding of the particles, the amount of sorting, and the composition of the sediment. In the case of **rounding**, the more mature the sediment, the rounder the particles, as a result of the particles being abraded over time. A high degree of sorting indicates maturity, because over time the smaller particles will be washed away, and a given amount of energy will move particles of a similar size over the same distance. Lastly, the older and more mature a sediment the higher the quartz content, at least in sediments derived from rock particles. Quartz is a common mineral in terrestrial rocks, and it is very hard and resistant to abrasion. Over time, particles made from other materials are worn away, leaving only quartz behind. Beach sand is a very mature sediment; it is composed primarily of quartz, and the particles are rounded and of similar size (well-sorted).

Sources of sediments

Sediments are also classified based on their source of origin. There are four main categories for the origin of marine sediments:

- **Lithogenous** sediments are derived from preexisting rock. They are also called **terrigenous** sediments since most of it comes from the land masses and makes its way into the ocean.
- **Biogenous** sediments are composed of the remains of marine organisms.
- **Hydrogenous** sediments are formed when materials that are dissolved in water precipitate out and form solid particles.

- **Cosmogenous** sediments are derived from extraterrestrial sources.

The next few sections will address each of these sediment types in more detail.

12.2 Lithogenous Sediments

Lithogenous or terrigenous sediment is primarily composed of small fragments of preexisting rocks that have made their way into the ocean. These sediments can contain the entire range of particle sizes, from microscopic clays to large boulders, and they are found almost everywhere on the ocean floor. Lithogenous sediments are created on land through the process of weathering, where rocks and minerals are broken down into smaller particles through the action of wind, rain, water flow, temperature- or ice-induced cracking, and other erosive processes. These small eroded particles are then transported to the oceans through a variety of mechanisms:

- **Streams and rivers**: Various forms of runoff deposit large amounts of sediment into the oceans, mostly in the form of finer-grained particles (Figure 12.2.1). About 90% of the lithogenous sediment in the oceans is though to have come from river discharge, particularly from Asia. Most of this sediment, especially the larger particles, will be deposited and remain fairly close to the coastline, however, smaller clay particles may remain suspended in the water column for long periods of time and may be transported great distances from the source.

Figure 12.2.1 *River discharge in the Yukon Delta, Alaska. The pale color demonstrates the large amounts of sediment released into the ocean via the rivers (By Jesse Allen and Robert Simmon (NASA Earth Observatory) [Public domain], via Wikimedia Commons).*

- **Wind**: Wind-borne (**aeolian**) transport can take small particles of sand and dust and move them thousands of kilometers from the source. These small particles can fall into the ocean when the wind dies down, or can serve as the nuclei around which raindrops or snowflakes form. Aeolian transport is particularly important near desert areas (Figure 12.2.2).

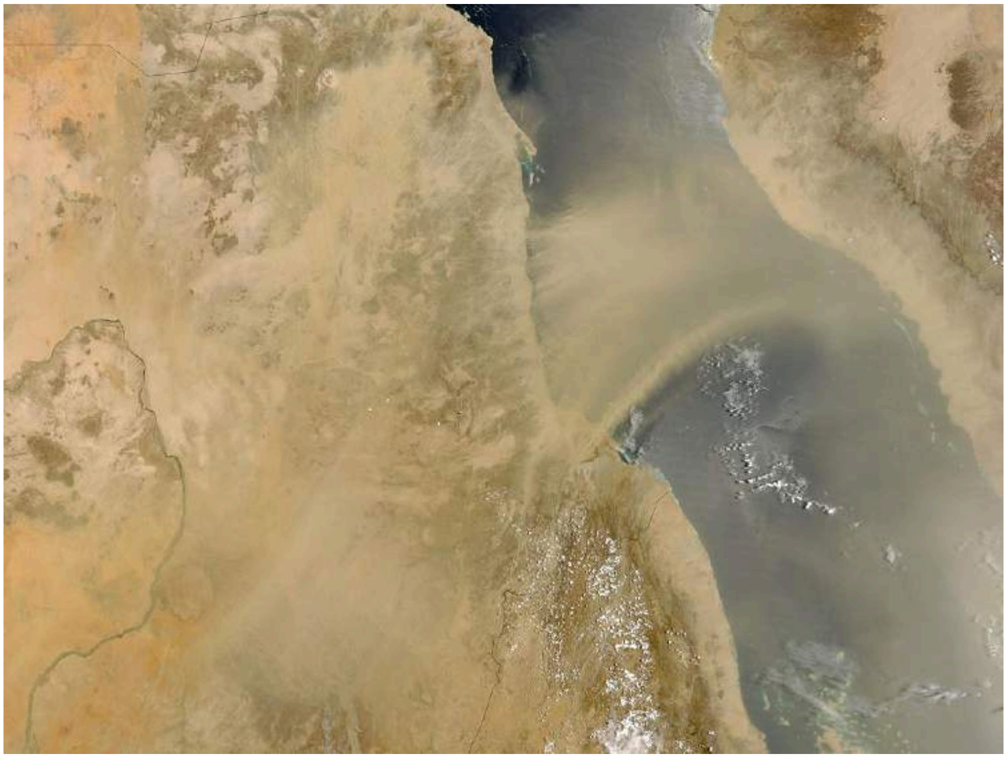


Figure 12.2.2 *A plume of wind-borne particles from Sudan (left) blow over the Red Sea (By NASA (http://visibleearth.nasa.gov/view_rec.php?id=5645) [Public domain], via Wikimedia Commons).*

- **Glaciers and ice rafting**: As glaciers grind their way over land, they pick up lots of soil and rock particles, including very large boulders, that get carried by the ice. When the glacier meets the ocean and begins to break apart or melt, these particles get deposited (Figure 12.2.3). Most of the deposition will happen close to where the glacier meets the water, but a small amount of material is also transported longer distances by rafting, where larger pieces of ice drift far from the glacier before releasing their sediment.

__Figure 12.2.3__ When glaciers reach the sea they can break apart, depositing their sediments into the ocean, including very large pieces of rock (By Ianaré Sévi (Own work) [CC BY-SA 3.0], via Wikimedia Commons).

- **Gravity**: Landslides, mudslides, avalanches, and other gravity-driven events can deposit large amounts of material into the ocean when they happen close to shore.
- **Waves**: Wave action along a coastline will erode rocks and will pull loose particles from beaches and shorelines into the water.
- **Volcanoes**: Volcanic eruptions emit vast amounts of ash and other debris into the atmosphere, where it can then be transported by wind to eventually get deposited in the oceans (Figure 12.2.4).

__Figure 12.2.4__ Eruption of the Mayon Volcano, Philippines, in 1984. Much of the material spewed from a volcanic eruption may eventually make its way into the oceans (By C.G. Newhall [Public domain], via Wikimedia Commons).

- **Gastroliths**: An interesting, although relatively minor avenue for the transport of lithogenous sediments to the ocean is in the form of gastroliths. Gastrolith means "stomach stones" and comes from the fact that many animals, including seabirds, pinnipeds, and some crocodiles will deliberately swallow stones in one area and regurgitate them in another. Often these stones swallowed on land will be regurgitated at sea. Why swallow stones? Possible explanations include using the stones to help grind up food in the stomach, to act as ballast to aid in buoyancy regulation, or to fill the stomach and reduce feelings of hunger during fasting periods on shore.

Most of these processes deposit lithogenous sediment fairly close to shore. The sediment particles can then be transported farther away by waves and currents, where they may eventually escape the continental shelf and reach the deep ocean floor.

Composition

Lithogenous sediments usually reflect the composition of whatever materials they were derived from, so they are dominated by the major minerals that make up most terrestrial rock. This includes quartz, feldspar, clay minerals, iron oxides, and terrestrial organic matter. Quartz (silicon dioxide, the main component of glass) is one of the most common minerals found in nearly all rocks, and it is very resistant to abrasion (see section 12.1), so it is a dominant component of lithogenous sediments, including sand.

12.3 Biogenous Sediments

Biogenous sediments come from the remains of living organisms that settle out as sediment when the organisms die. It is the "hard parts" of the organisms that contribute to the sediments; things like shells, teeth or skeletal elements, as these parts are usually mineralized and are more resistant to decomposition than the fleshy "soft parts" that rapidly deteriorate after death.

Macroscopic sediments contain large remains, such as skeletons, teeth, or shells of larger organisms. This type of sediment is fairly rare over most of the ocean, as large organisms don't die in enough of a concentrated abundance to allow these remains to accumulate. One exception is around coral reefs; here there is a great abundance of organisms that leave behind their remains, in particular the fragments of the stony skeletons of corals that make up a large percentage of tropical sand.

Microscopic sediment consists of the hard parts of microscopic organisms, particularly their shells, or **tests**. Although very small, these organisms are highly abundant and as they die by the billions every day their tests sink to the bottom to create biogenous sediments. Sediments composed of microscopic tests are far more abundant than sediments from macroscopic particles, and because of their small size they create fine-grained, mushy sediment layers. If the sediment layer consists of at least 30% microscopic biogenous material, it is classified as a biogenous **ooze**. The remainder of the sediment is often made up of clay.

The primary sources of microscopic biogenous sediments are unicellular algaes and protozoans (single-celled amoeba-like creatures) that secrete tests of either calcium carbonate ($CaCO_3$) or silica (SiO_2). Silica tests come from two main groups, the **diatoms** (algae) and the **radiolarians** (protozoans) (Figure 12.3.1).

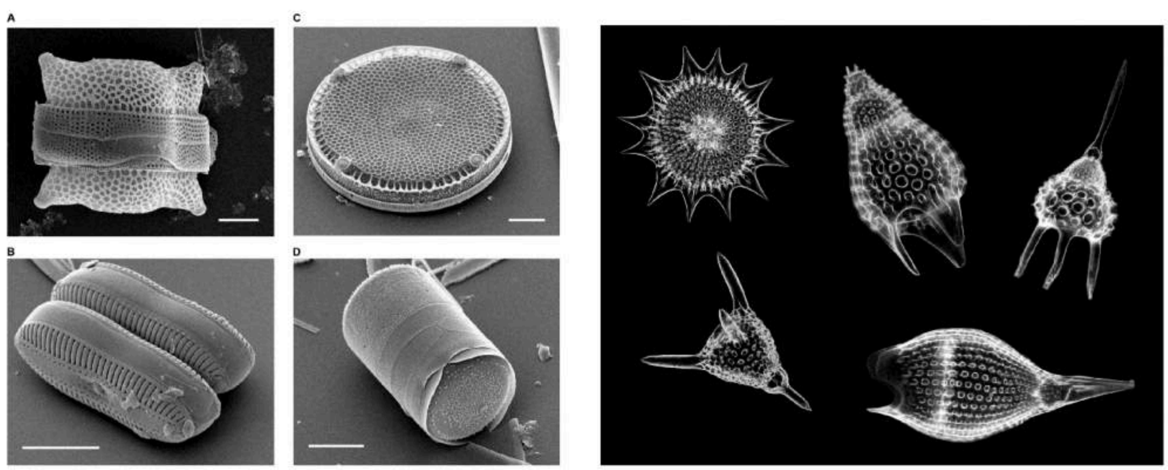


Figure 12.3.1 *Various diatom (left) and radiolarian (right) tests (Diatom images courtesy of Mary Ann Tiffany, San Diego State University [CC BY 2.5], via Wikimedia Commons; radiolarian images by Andreas Drews, https://pxhere.com/en/photo/239774, [CC by 2.0]).*

Diatoms are important members of the phytoplankton, the small, drifting algal photosynthesizers. A diatom consists of a single algal cell surrounded by an elaborate silica shell that it secretes for itself. Diatoms come in a range of shapes, from elongated, pennate forms, to round, or centric shapes that often have two halves, like a Petri dish (Figure 12.3.1 *left*). In areas where diatoms are abundant, the underlying sediment is rich in silica diatom tests, and is called **diatomaceous earth** (see box below).

What use are diatoms?

Diatoms are a vital piece of the global ecosystem for their role in oceanic primary production and the creation of much of the oxygen that organisms breathe. But diatoms are also important for many industrial and agricultural applications. Because of the very fine grain size, and the lattice-like structure of the diatom tests, diatomaceous earth has been used as a filtering agent in things like swimming pool filters and beer brewing. The microscopic tests have been added as an abrasive to toothpaste, facial cleansers and household cleaning agents. Alfred Nobel used diatomaceous earth to stabilize nitroglycerine in the production of dynamite. Diatomaceous earth also displays insecticide properties by stimulating dehydration in insects. It is marketed for this purpose in agriculture, as well as for household use to combat ants, cockroaches, and bedbugs. "Food grade" diatomaceous earth has also entered the market, with proponents touting a range of health benefits arising from its consumption. That's a pretty impressive range of uses from a microscopic algae!

Radiolarians are planktonic protozoans (making them part of the zooplankton), that like diatoms, secrete a silica test. The test surrounds the cell and can include an array of small openings through which the radiolarian can extend an amoeba-like "arm" or pseudopod (Figure 12.3.1 *right*). Radiolarian tests often display a number of rays protruding from their shells which aid in buoyancy. Oozes that are dominated by diatom or radiolarian tests are called **siliceous oozes**.

Like the siliceous sediments, the calcium carbonate, or calcareous sediments are also produced from the tests of microscopic algae and protozoans; in this case the **coccolithophores** and **foraminiferans**. Coccolithophores are single-celled planktonic algae about 100 times smaller than diatoms. Their tests are composed of a number of interlocking $CaCO_3$ plates (coccoliths) that form a sphere surrounding the cell (Figure 12.3.2 *left*). When coccolithophores die the individual plates sink out and form an ooze. Over time, the coccolithophore ooze lithifies to becomes chalk. The famous White Cliffs of Dover in England are composed of coccolithophore-rich ooze that turned into chalk deposits (Figure 12.3.2 *right*).

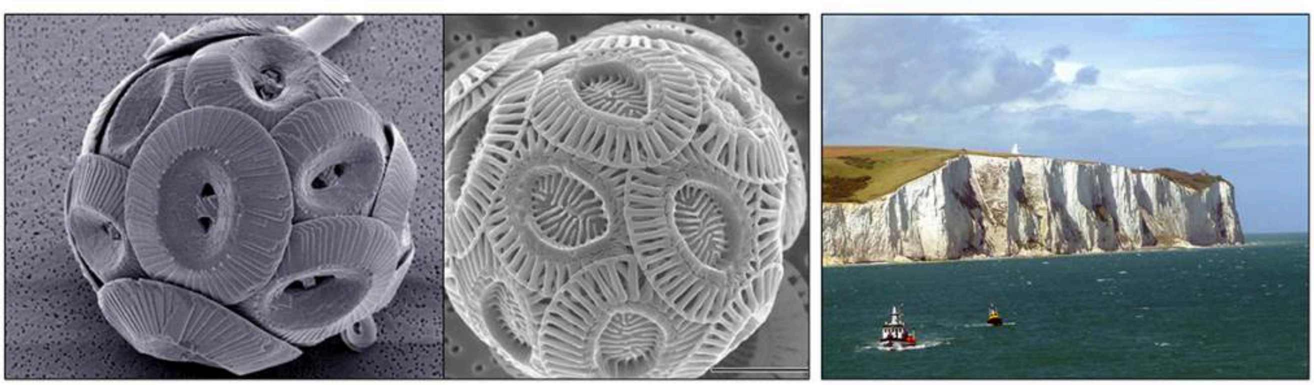


Figure 12.3.2 *(**Left**) coccolithophore tests (left; By Richard Lampitt, Jeremy Young, The Natural History Museum, London (http://planktonnet.awi.de/); center; by Alison R. Taylor (University of North Carolina Wilmington Microscopy Facility) (PLoS Biology, June 2011, Cover ([1])) [Both images CC BY 2.5], via Wikimedia Commons). (**Right**); the White Cliffs of Dover (Immanuel Giel (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

Foraminiferans (also referred to as “forams”) are protozoans whose tests are often chambered, similar to the shells of snails. As the organism grows, is secretes new, larger chambers in which to reside. Most foraminiferans are benthic, living on or in the sediment, but there are some planktonic species living higher in the water column. When coccolithophores and foraminiferans die, they form **calcareous oozes**.

Figure 12.3.3 *Foraminifera tests collected from a beach in Myanmar (By Psammophile [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 3.0], via Wikimedia Commons).*

Older calcareous sediment layers contain the remains of another type of organism, the **discoasters**; single-celled algae related to the coccolithophores that also produced calcium carbonate tests. Discoaster tests were star-shaped, and reached sizes of 5-40 μm across (Figure 13.3.4). Discoasters went extinct approximately 2 million years ago, but their tests remain in deep, tropical sediments that predate their extinction.

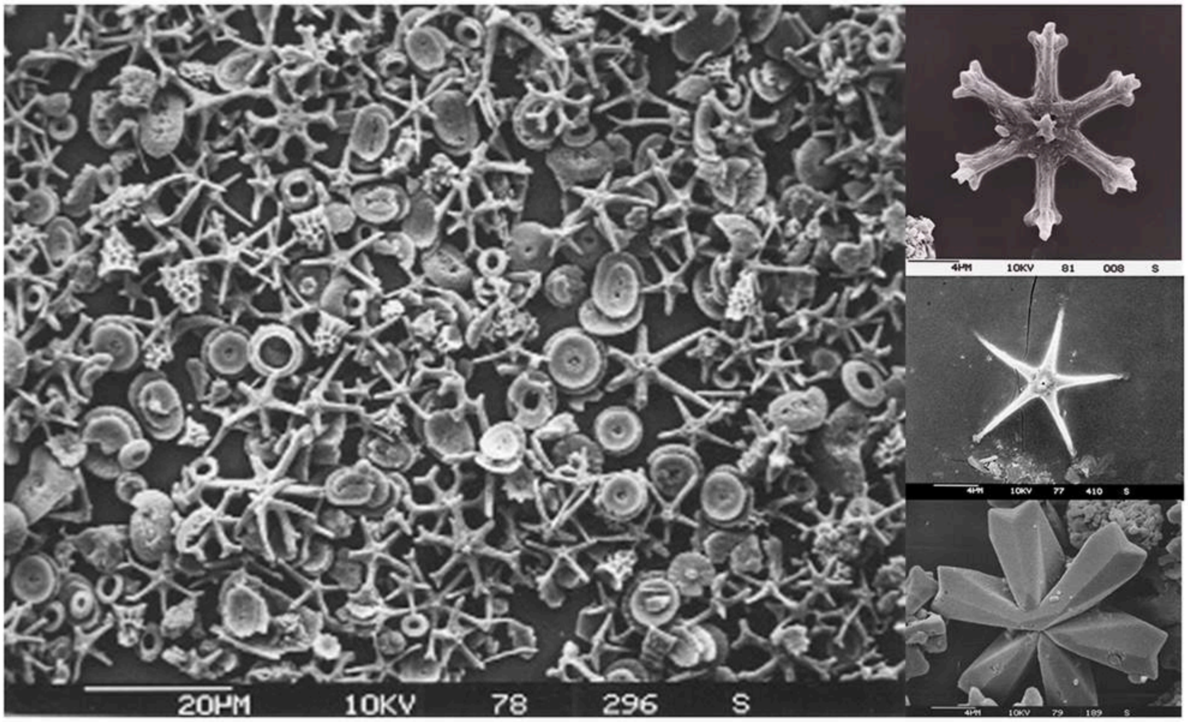


Figure 12.3.4 *Discoaster tests. Left: discoster tests with assorted coccoliths. Top right Discoaster surculus; center right: Discoaster pentaradiatus; bottom right: Discoaster surculus (All images by Hannes Grobe (Own work) [CC BY 3.0], via Wikimedia Commons).*

Because of their small size, these tests sink very slowly; a single microscopic test may take about 10-50 years to sink to the bottom! Given that slow descent, a current of only 1 cm/sec could carry the test as much as 15,000 km away from its point of origin before it reaches the bottom. Yet despite this, we find that the sediments in a particular location are well-matched to the types of organisms and degree of productivity that occurs in the water overhead. This means that the sediment particles must be sinking to the bottom at a much faster rate, so that they accumulate below their point of origin before the currents can disperse them. What is the mechanism for this increased sinking rate? Apparently most of the tests do not sink as individual particles; about 99% of them are first consumed by some other organism, and are then aggregated and expelled as large fecal pellets, which sink much more quickly and reach the ocean floor in only 10-15 days. This does not give the particles as much time to disperse, and the sediment below will reflect the production occurring near the surface. The increased rate of sinking through this mechanism is called the “fecal express.”

Reconstructing past climate through sediment analysis

As outlined in the opening to this chapter, examining marine sediments allows us to learn much about oceanographic and atmospheric processes, both past and present. Biogenous sediments are no exception, and they can allow us to reconstruct past climate history from oxygen isotope ratios.

Oxygen atoms exist in three forms, or isotopes, in ocean water: O^{16}, O^{17} and O^{18} (the number refers to the atomic masses of the isotopes). O^{16} is the most common form, followed by O^{18} (O^{17} is rare). O^{16} is lighter than O^{18}, so it evaporates more easily, leading to water vapor that has a higher proportion of O^{16}. During periods of cooler climate, water vapor condenses into rain and snow, which forms glacial ice that has a high proportion of O^{16}. The remaining seawater therefore has a relatively higher proportion of O^{18}. Marine organisms who incorporate dissolved oxygen into their shells as calcium carbonate will therefore have shells with a higher proportion of O^{18} isotope. In other words, the ratio of O^{16}:O^{18} in shells will be low during periods of colder climate.

When the climate warms, glacial ice melts, releasing O^{16} from the ice and returning it to the oceans, increasing the O^{16}:O^{18} ratio in the water. Now, when organisms incorporate oxygen into their shells, the shells will contain a higher O^{16}:O^{18} ratio. Scientists can therefore examine biogenous sediments, calculate the O^{16}:O^{18} ratios for samples of known ages, and from those ratios, infer the climate conditions under which those shells were formed. The same types of measurements can also be taken from ice cores; a decrease of 1 ppm O^{18} between ice samples represents a decrease in temperature of 1.5^{o} C.

12.4 Hydrogenous Sediments

*Methane hydrate section modified from "Physical Geology" by Steven Earle**

As we saw in section 5.3 seawater contains many different dissolved substances. Occasionally chemical reactions occur that cause these substances to precipitate out as solid particles, which then accumulate as **hydrogenous sediment**. These reactions are usually triggered by a change in conditions, such as a change in temperature, pressure, or pH, which reduces the amount of a substance that can remain in a dissolved state. There is not a lot of hydrogenous sediment in the ocean compared to lithogenous or biogenous sediments, but there are some interesting forms.

Hydrothermal vents were discussed in section 4.11. Recall that in these systems, seawater percolates into the seafloor, where it becomes superheated by magma before being expelled by the vent. This superheated water contains many dissolved substances, and when it encounters the cold seawater after leaving the vent, these particles precipitate out, mostly as metal sulfides. These particles make up the "smoke" that flows from a vent, and may eventually settle on the bottom as hydrogenous sediment (Figure 12.4.1).

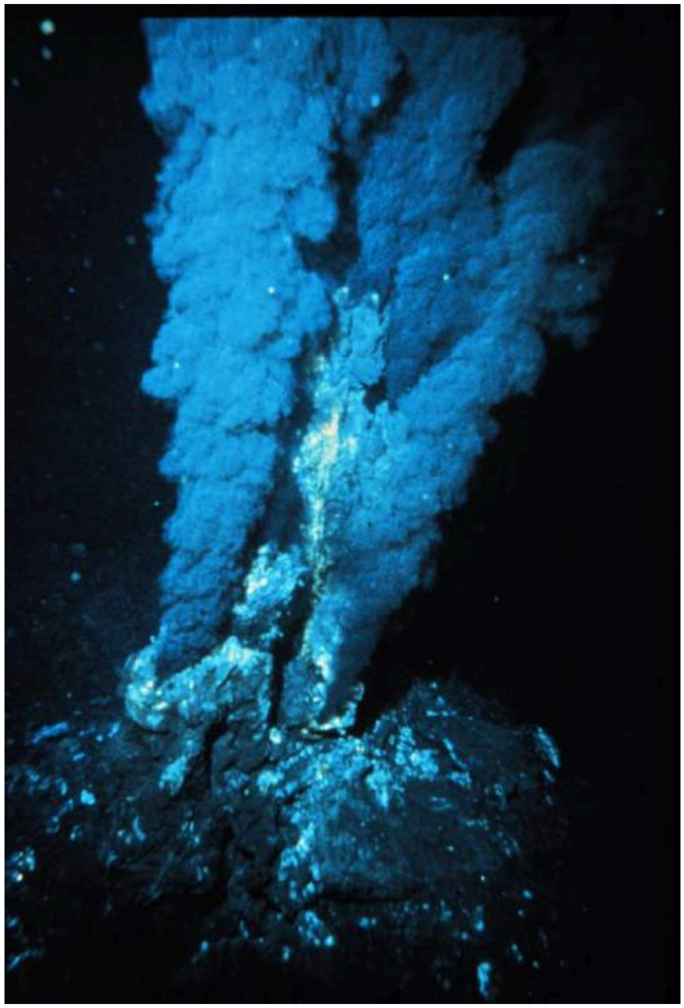


Figure 12.4.1 *A "black smoker" hydrothermal vent. The "smoke" consists of dissolved particles that precipitate into solids when exposed to colder water (NOAA, http://www.photolib.noaa.gov/htmls/nur04506.htm).*

Manganese nodules are rounded lumps of manganese and other metals that form on the seafloor, generally ranging between 3-10 cm in diameter, although they may sometimes reach up to 30 cm (Figure 12.4.2). The nodules form in a manner similar to pearls; there is a central object around which concentric layers are slowly deposited, causing the nodule to grow over time. The composition of the nodules can vary somewhat depending on their location and the conditions of their formation, but they are usually dominated by manganese- and iron oxides. They may also contain smaller amounts of other metals such as copper, nickel and cobalt. The precipitation of manganese nodules is one of the slowest geological processes known; they grow on the order of a few millimeters per million years. For that reason, they only form in areas where there are low rates of lithogenous or biogenous sediment accumulation, because any other sediment deposition would quickly cover the nodules and prevent further nodule growth. Therefore, manganese nodules are usually limited to areas in the central ocean, far from significant lithogenous or biogenous inputs, where they can sometimes accumulate in large numbers on the seafloor (Figure 12.4.2 *right*). Because the nodules contain a number of commercially valuable metals, there has been significant interest in mining the nodules over the last several decades, although most of the efforts have thus far remained at the exploratory stage. A number of factors have prevented large-scale extraction of nodules, including the high costs of deep sea mining operations, political issues over mining rights, and environmental concerns surrounding the extraction of these non-renewable resources.

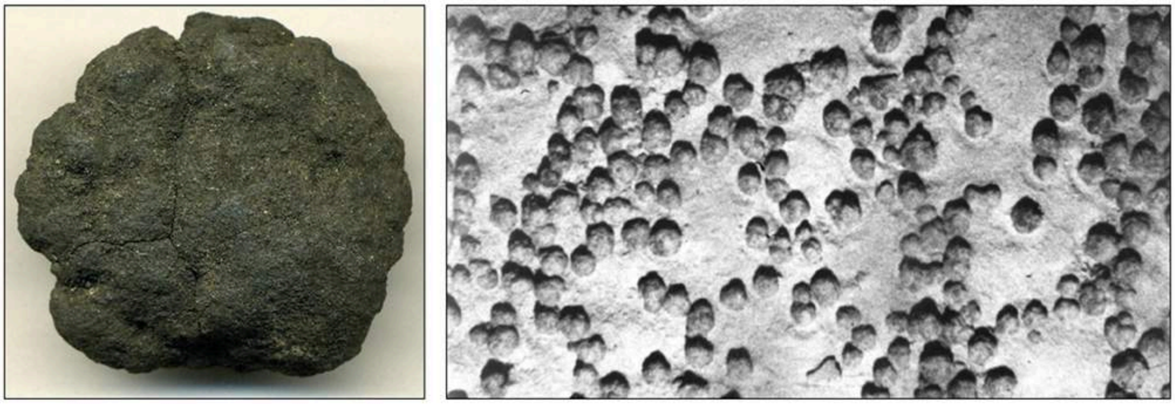


Figure 12.4.2 *(Left) manganese nodule from the subtropical eastern Pacific Ocean. The nodule is 4.1 cm in diameter (By James St. John, https://www.flickr.com/photos/jsjgeology/15139986302/in/photostream/ [CC-BY 2.0]). (Right) field of manganese nodules on the seafloor (By United States Geological Survey [Public domain], via Wikimedia Commons).*

Evaporites are hydrogenous sediments that form when seawater evaporates, leaving the dissolved materials to precipitate into solids, particularly halite (salt, NaCl). In fact, the evaporation of seawater is the oldest form of salt production for human use, and is still carried out today. Large deposits of halite evaporites exist in a number of places, including under the Mediterranean Sea. Beginning around 6 million years ago, tectonic processes closed off the Mediterranean Sea from the Atlantic, and the warm climate evaporated so much water that the Mediterranean was almost completely dried out, leaving large deposits of salt in its place (an event known as the Messinian Salinity Crisis). Eventually the Mediterranean re-flooded about 5.3 million years ago, and the halite deposits were covered by other sediments, but they still remain beneath the seafloor.

Figure 12.4.3 *Salt farmers harvesting salt left behind from the evaporation of seawater, Pak Thale, Ban Laem, Phetchaburi, Thailand (By JJ Harrison (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

Oolites are small, rounded grains formed from concentric layers of precipitation of material around a suspended particle. They are usually composed of calcium carbonate, but they may also from from phosphates and other materials. Accumulation of oolites results in oolitic sand, which is found in its greatest abundance in the Bahamas (Figure 12.4.4).

Figure 12.4.4 *Oolites from a beach on Joulter's Cay, The Bahamas (By Wilson44691 (Own work) [Public domain], via Wikimedia Commons).*

Methane hydrates are another type of hydrogenous deposit with a potential industrial application. All terrestrial erosion products include a small proportion of organic matter derived mostly from terrestrial plants. Tiny fragments of this material plus other organic matter from marine plants and animals accumulate in terrigenous sediments, especially within a few hundred kilometers of shore. As the sediments pile up, the deeper parts start to warm up (from geothermal heat), and bacteria get to work breaking down the contained organic matter. Because this is happening in the absence of oxygen (a.k.a. anaerobic conditions), the by-product of this metabolism is the gas methane (CH_4). Methane released by the bacteria slowly bubbles upward through the sediment toward the seafloor. At water depths of 500 m to 1,000 m, and at the low temperatures typical of the seafloor (close to 4°C), water and methane combine to create a substance known as methane hydrate. Within a few meters to hundreds of meters of the seafloor, the temperature is low enough for methane hydrate to be stable and hydrates accumulate within the sediment (Figure 12.4.5 *left*). Methane hydrate is flammable because when it is heated, the methane is released as a gas (Figure 12.4.5 *right*). The methane within seafloor sediments represents an enormous reservoir of fossil fuel energy. Although energy corporations and governments are anxious to develop ways to produce and sell this methane, anyone that understands the climate-change implications of its extraction and use can see that this would be folly.

***Figure 12.4.5** (Left): Methane hydrate within muddy sea-floor sediment from an area offshore from Oregon (By Wusel007 (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 3.0], via Wikimedia Commons). (Right): Methane hydrate on fire (USGS, http://www.usgs.gov/blogs/features/files/2012/01/New-Image.jpg).*

12.5 Cosmogenous Sediments

Cosmogenous sediment is derived from extraterrestrial sources, and comes in two primary forms; microscopic spherules and larger meteor debris. Spherules are composed mostly of silica or iron and nickel, and are thought to be ejected as meteors burn up after entering the atmosphere. Meteor debris comes from collisions of meteorites with Earth. These high impact collisions eject particles into the atmosphere that eventually settle back down to Earth and contribute to the sediments. Like spherules, meteor debris is mostly silica or iron and nickel. One interesting form of debris from these collisions are tektites, which are small droplets of glass. They are likely composed of terrestrial silica that was ejected and melted during a meteorite impact, which then solidified as it cooled upon returning to the surface.

Figure 12.5.1 *Tektite-like glass found near the Volkov River in western Russia (By James St. John [CC BY 2.0], via Wikimedia Commons).*

Cosmogenous sediment is fairly rare in the ocean and it does not usually accumulate in large deposits. However, it is constantly being added to through space dust that continuously rains down on Earth. About 90% of incoming cosmogenous debris is vaporized as it enters the atmosphere, but it is estimated that 5 to 300 tons of space dust land on the Earth's surface each day!

12.6 Sediment Distribution

Now that we have an understanding of the types of sediments found in the ocean, we can turn our attention to the processes that cause different types of sediments to dominate in different locations. Sediment accumulation will depend on the the amount of material coming from the source, the distance from the source, the amount of time that sediment has had to accumulate, how well the sediments are preserved, and the amounts of other types of sediments that are also being added to the system.

Rates of sediment accumulation are relatively slow throughout most of the ocean, in many cases taking thousands of years for any significant deposits to form. Lithogenous sediment accumulates the fastest, on the order of 1 m or more per thousand years for coarser particles. However, sedimentation rates near the mouths of large rivers with high discharge can be orders of magnitude higher. Biogenous oozes accumulate at a rate of about 1 cm per thousand years, while small clay particles are deposited in the deep ocean at around 1 mm per thousand years. As described in section 12.4, manganese nodules have an incredibly slow rate of accumulation, gaining 0.001 mm per thousand years.

Marine sediments are thickest near the continental margins (refer to figure 12.1.1) where they can be over 10 km thick. This is because the crust near passive continental margins is often very old, allowing for a long period of accumulation, and because there is a large amount of terrigenous sediment input coming from the continents. Near mid-ocean ridge systems where new oceanic crust is being formed, sediments are thinner, as they have had less time to accumulate on the younger crust. As you move away from the ridge spreading center the sediments get progressively thicker (see section 4.5), increasing by approximately 100-200 m of sediment for every 1000 km distance from the ridge axis. With a seafloor spreading rate of about 20-40 km/million years, this represents a sediment accumulation rate of approximately 100-200 m every 25-50 million years.

Figure 12.6.1 shows the distribution of the major types of sediment on the ocean floor. Cosmogenous sediments could potentially end up in any part of the ocean, but they accumulate in such small abundances that they are overwhelmed by other sediment types and thus are not dominant in any location. Similarly, hydrogenous sediments can have high concentrations in specific locations, but these regions are very small on a global scale. So we will mostly ignore cosmogenous and hydrogenous sediments in the discussion of global sediment patterns.

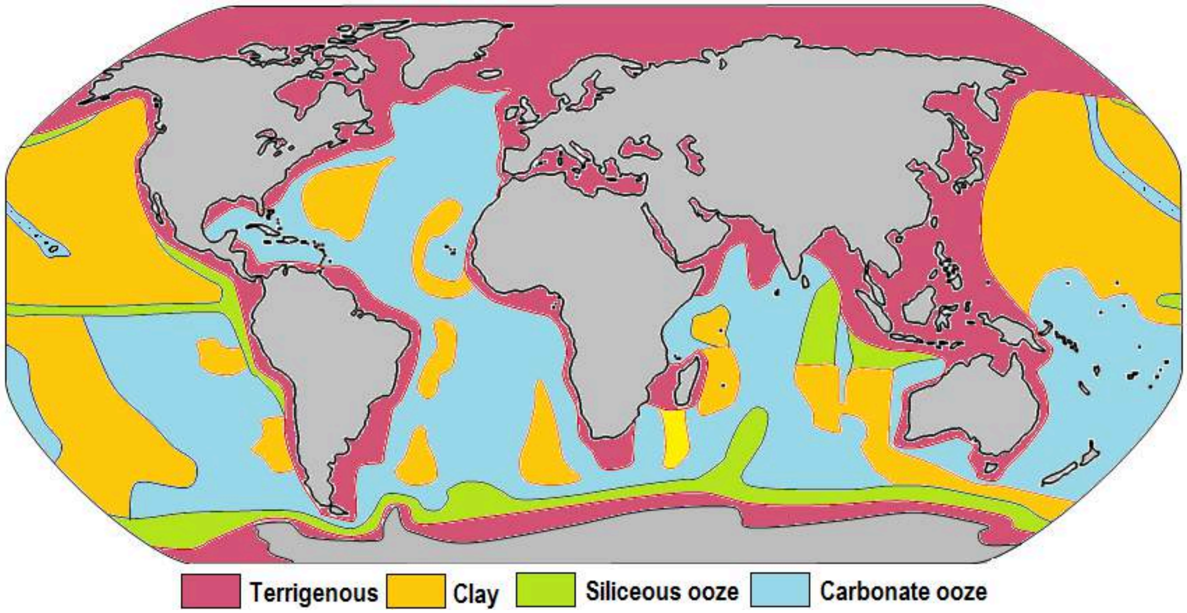


Figure 12.6.1 *The distribution of sediment types on the seafloor. Within each colored area, the type of material shown is what dominates, although other materials are also likely to be present (Steven Earle, "Physical Geology").*

Coarse lithogenous/terrigenous sediments are dominant near the continental margins as runoff, river discharge, and other processes deposit vast amounts of these materials on the continental shelf (section 12.2). Much of this sediment remains on or near the shelf, while turbidity currents can transport material down the continental slope to the deep ocean floor. Lithogenous sediment is also common at the poles where thick ice cover can limit primary production, and glacial breakup deposits sediments along the ice edge. Coarse lithogenous sediments are less common in the central ocean, as these areas are too far from the sources for these sediments to accumulate. Very small clay particles are the exception, and as described below, they can accumulate in areas that other lithogenous sediment will not reach.

The distribution of biogenous sediments depends on their rates of production, dissolution, and dilution by other sediments. We learned in section 7.4 that coastal areas display very high primary production, so we might expect to see abundant biogenous deposits in these regions. However, recall that sediment must be >30% biogenous to be considered a biogenous ooze, and even in productive coastal areas there is so much lithogenous input that it swamps the biogenous materials, and that 30% threshold is not reached. So coastal areas remain dominated by lithogenous sediment, and biogenous sediments will be more abundant in pelagic environments where there is little lithogenous input.

In order for biogenous sediments to accumulate their rate of production must be greater than the rate at which the tests dissolve. Silica is undersaturated throughout the ocean and will dissolve in seawater, but it dissolves more readily in warmer water and lower pressures; in other words, it dissolves faster near the surface than in deep water. Silica sediments will therefore only accumulate in cooler regions of high productivity where they accumulate faster than they dissolve. This includes upwelling regions near the equator and at high latitudes where there are abundant nutrients and cooler water. Oozes formed near the equatorial regions are usually dominated by radiolarians, while diatoms are more common in the polar oozes. Once the silica tests have settled on the bottom and are covered by subsequent layers, they are no longer subject to dissolution and the sediment will accumulate. Approximately 15% of the seafloor is covered by siliceous oozes.

Biogenous calcium carbonate sediments also require production to exceed dissolution for sediments to accumulate, but the processes involved are a little different than for silica. Calcium carbonate dissolves more readily in more acidic water. Cold seawater contains more dissolved CO_2 and is slightly more acidic than warmer water (section 5.5). Therefore calcium carbonate tests are more likely to dissolve in colder, deeper, polar water than in warmer, tropical, surface water. At the poles the water is uniformly cold, so calcium carbonate readily dissolves at all depths, and carbonate sediments do not accumulate. In temperate and tropical regions calcium carbonate dissolves more readily as it sinks into deeper water. The depth at which calcium carbonate dissolves as fast as it accumulates is called the **calcium carbonate compensation depth**, or **calcite compensation depth**, or simply the **CCD**. The **lysocline** represents the depths where the rate of calcium carbonate dissolution increases dramatically (similar to the thermocline and halocline). At depths shallower than the CCD carbonate accumulation will exceed the rate of dissolution, and carbonate sediments will be deposited. In areas deeper than the CCD, the rate of dissolution will exceed production, and no carbonate sediments can accumulate (Figure 12.6.2). The CCD is usually found at depths of 4 – 4.5 km, although it is much shallower at the poles where the surface water is cold. Thus calcareous oozes will mostly be found in tropical or temperate waters less than about 4 km deep, such as along the mid-ocean ridge systems and atop seamounts and plateaus. The CCD is deeper in the Atlantic than in the Pacific since the Pacific contains more CO_2, making the water more acidic and calcium carbonate more soluble. This, along with the fact that the Pacific is deeper, means that the Atlantic contains more calcareous sediment than the Pacific. All told, about 48% of the seafloor is dominated by calcareous oozes.

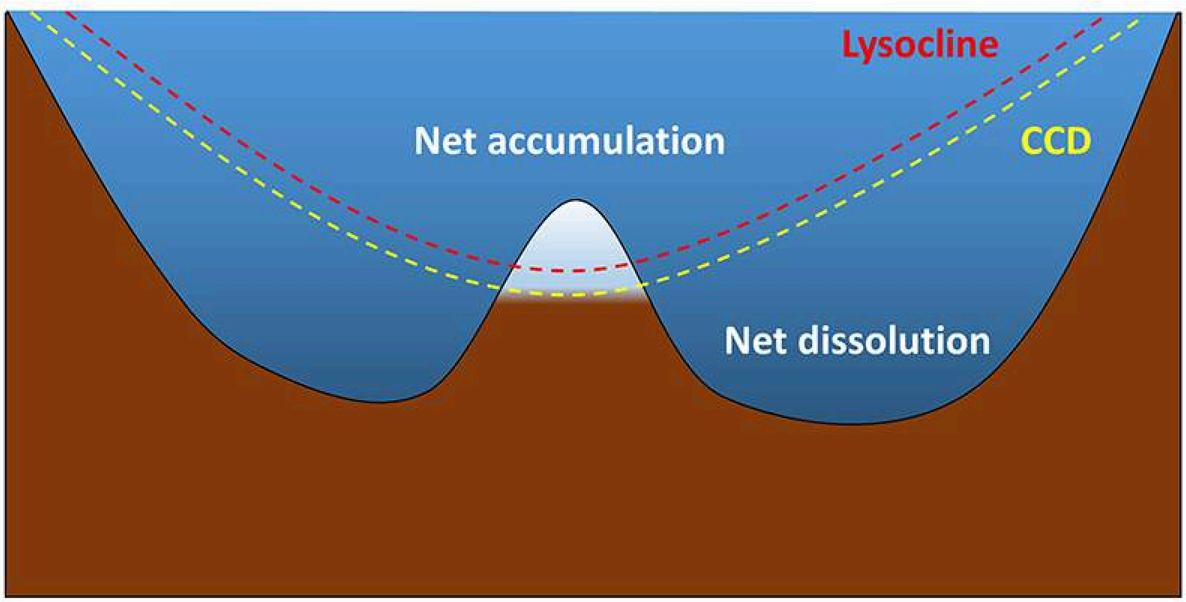


Figure 12.6.2 *Calcareous sediment can only accumulate in depths shallower than the calcium carbonate compensation depth (CCD). Below the CCD, calcareous sediments dissolve and will not accumulate. The lysocline represents the depths where the rate of dissolution increases dramatically (PW).*

Much of the rest of the deep ocean floor (about 38%) is dominated by abyssal clays. This is not so much a result of an abundance of clay formation, but rather the lack of any other types of sediment input. The clay particles are mostly of terrestrial origin, but because they are so small they are easily dispersed by wind and currents, and can reach areas inaccessible to other sediment types. Clays dominate in the central North Pacific, for example. This area is too far from land for coarse lithogenous sediment to reach, it is not productive enough for biogenous tests to accumulate, and it is too deep for calcareous materials to reach the bottom before dissolving. Because clay particles accumulate so slowly, the clay-dominated deep ocean floor is often home to hydrogenous sediments like manganese nodules. If any other type of sediment was produced here it would accumulate much more quickly and would bury the nodules before they had a chance to grow.

CHAPTER 13: COASTAL OCEANOGRAPHY

Chapter 13: Coastal Oceanography

Learning Objectives

After reading this chapter you should:

- recognize the various zones of a beach
- understand how the relationship between swash and backwash determines the composition of a beach
- understand the concept of longshore transport
- know the different erosional and depositional structures created by longshore transport
- understand the issues associated with different forms of hard stabilization: groins, jetties etc.
- know what an estuary is
- know the four types of geological estuaries, and how they form
- know the four types of estuaries based on salinity and mixing patterns

For most people, their image of the coast is the place where the land meets the sea, most likely in the form of a beach. But it is more than just the narrow strip along the water line; technically the term “coast” refers to the range of land over which the ocean has an effect on climate, foliage, and other environmental processes. This range may extend for tens of kilometers inland from the water’s edge. Furthermore, what we recognize as the coast today, may not have been a coastal area in the past, as sea level has varied from about 6 m above to 125 m below its current height over the past two million years.

This chapter begins with the features of coastal regions, the processes that shape the coastline, and how humans try to control these processes. Following that, we will examine the different types of estuaries that are found in coastal areas.

13.1 Beaches

For most people, when they think of coastal areas they picture a beach, and the beach that they imagine is probably a typical sandy beach composed of quartz sand grains (section 12.2). But beaches are comprised of whatever types of sediments are dominant in the local area. For example, parts of Hawaii and Iceland are famous for their black sand beaches, made up of eroded basalt and other volcanic materials. The beautiful tropical white sand beaches we see in travel ads are largely composed of the crushed calcium carbonate remains of coral skeletons (much of which has been chewed up and excreted by a fish before we happily run our toes through it!) Other beaches may lack sand altogether and instead be dominated by small shells, or larger rocks or pebbles (Figure 13.1.1).

Figure 13.1.1 *Various beach substrates. Clockwise from top left: Punaluu Black Sand Beach, Hawaii, USA (Diego Delso [CC BY-SA 3.0], via Wikimedia Commons); Shell Beach, Shark Bay, Western Australia (Brian W. Schaller (Own work) [CC BY-NC-SA 3.0], via Wikimedia Commons); white coral sand beach in the Maldives (http://www.elitedivingagency.com/, [CC BY-SA 3.0] via Wikimedia Commons); rocky beach at Killbear Provincial Park in Ontario, Canada (John Vetterli (originally posted to Flickr as Beach) [CC BY-SA 2.0], via Wikimedia Commons).*

The shoreline is divided up into multiple zones (Figure 13.1.2). The **backshore** is the region of the beach above the high tide line, which is only submerged under unusually high wave conditions, such as during storms. The

foreshore lies between the high tide and low tide lines; it is submerged during high tide and is exposed during low tide. The **nearshore** extends from the low tide line to the depth where wave action is no longer influenced by the bottom, i.e. to where the depth exceeds the wave base (section 10.1). Finally, the **offshore** zone represents the depths beyond the nearshore region.

Along the beach itself, the area above the high tide line is called the **berm**, which is usually dry and relatively flat. The berm often ends with a berm crest or berm **scarp**, which is a steeper wall carved out by wave action that leads down to the foreshore. The foreshore has a number of other names, including the **beach face**, the **intertidal** or **littoral zone**, and if the area is fairly flat, the **low tide terrace**. Just off shore from the beach there are often **longshore bars** and longshore troughs running parallel to the beach. The longshore bars are accumulations of sand that are deposited by wave action and longshore currents (section 13.2). The decrease in depth above longshore bars is what often causes waves to start to break well before reaching the beach (section 10.3).

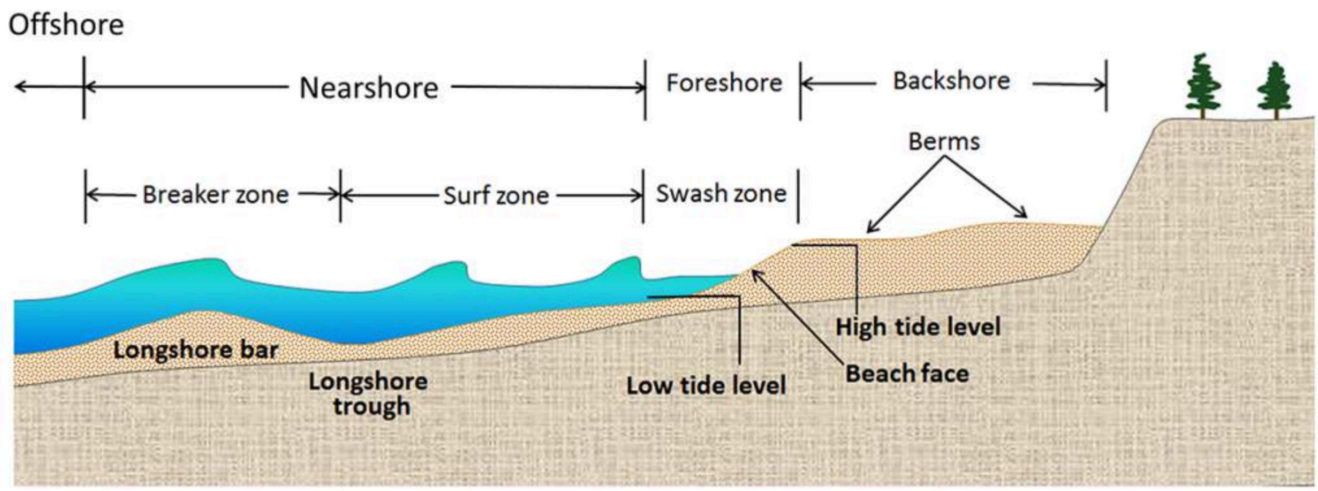


Figure 13.1.2 *The zones of a typical beach (Modified by PW from Steven Earle, "Physical Geology").*

The sand or other particles that make up the beach are distributed by wave action. The water that moves over a beach through incoming waves is called **swash**, and it also contains suspended sand grains that can get deposited on the beach. Some of the swash percolates into the sand while the rest of the water washes back out as **backwash** as the wave recedes. Backwash removes sand from the beach and returns it to the ocean. Sand will therefore be deposited or eroded depending on which process is dominant. If wave action is light, a lot of incoming water gets absorbed by the sand, so swash dominates. Under heavier waves the beach becomes saturated with water, so less can be absorbed, and backwash is dominant. This leads to seasonal cycles in beach structure; waves are heavier during the winter as a result of stormier conditions at sea, so backwash dominates and sand is removed from the beach and deposited offshore in longshore bars. In the summer the waves are gentler, swash dominates, and the sand is transported from the longshore bar and deposited on the shore to create a wider, sandy beach (Figure 13.1.3).

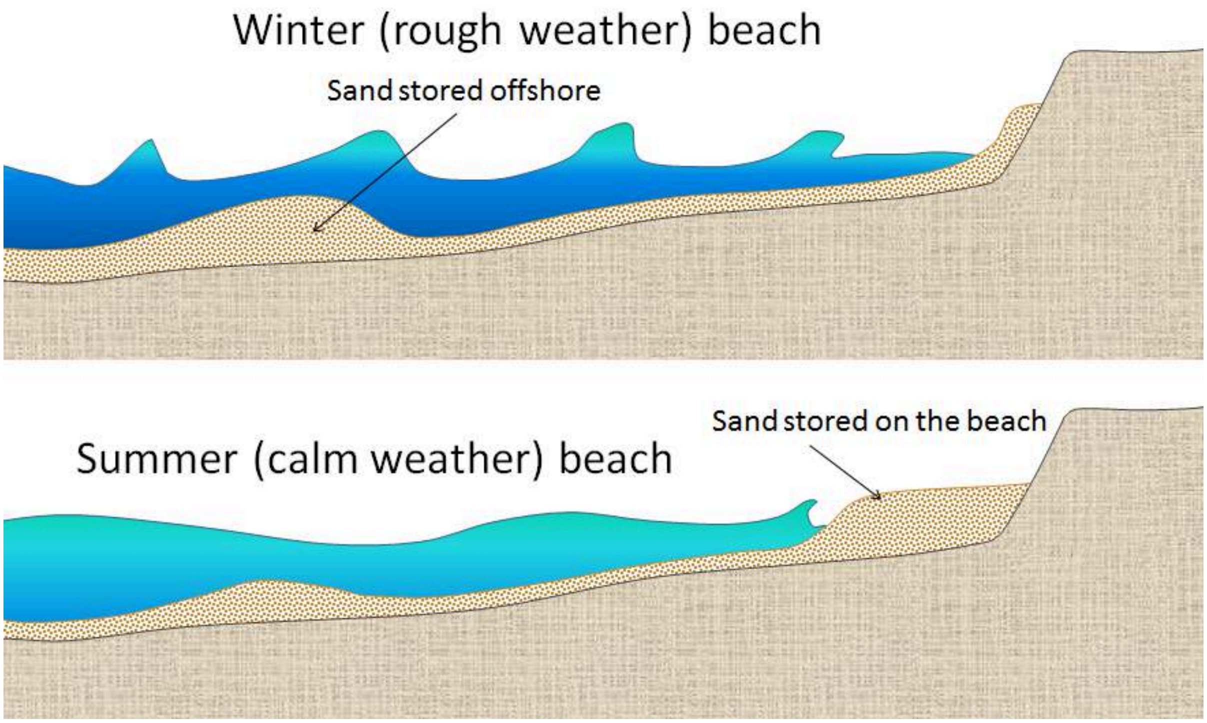


Figure 13.1.3 *The differences between summer and winter on beaches in areas where the winter conditions are rougher and waves have a shorter wavelength but higher energy. In winter, sand from the beach is stored offshore (Steven Earle, “Physical Geology”).*

13.2 Longshore Transport

*Modified from "Physical Geology" by Steven Earle**

We learned in section 10.3 that refraction causes waves to approach parallel to shore. However, most waves still reach the shore at a small angle, and as each one arrives, it pushes water along the shore, creating what is known as a **longshore current** within the surf zone (the areas where waves are breaking) (Figure 13.2.1). Longshore currents can move up to 4 km/hr, strong enough to carry people with them, as everyone knows who has been swimming in the ocean only to look up and see that they have been carried far down the beach from their towel!

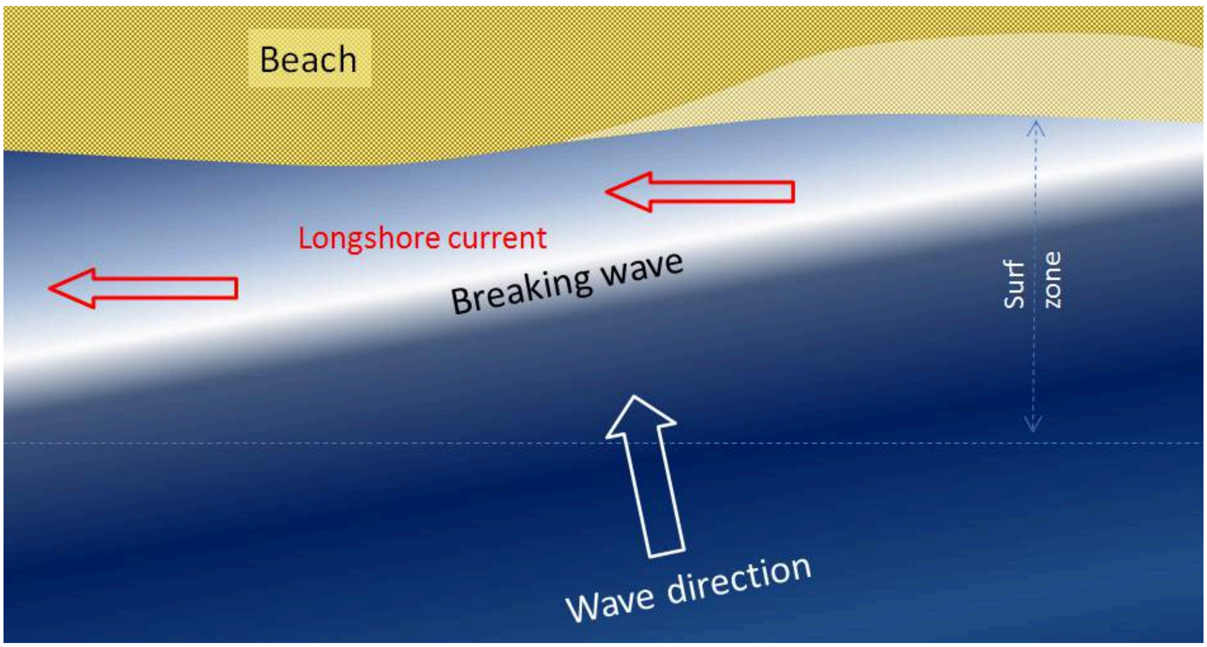


Figure 13.2.1 *Longshore currents are caused by waves approaching shore at a small angle, moving water parallel to the shore (Steven Earle, "Physical Geology").*

Another important effect of waves reaching the shore at an angle is that when they wash up onto the beach, they do so at an angle, but when that same wave water washes back down the beach, it moves straight down the slope of the beach (Figure 13.2.2). The upward-moving water, known as the swash, pushes sediment particles along the beach, while the downward-moving water, the backwash, brings them straight back. With every wave that washes up and then down the beach, particles of sediment are moved along the beach in a zigzag pattern.

The combined effects of sediment transport within the surf zone by the longshore current and sediment movement along the beach by swash and backwash is known as **longshore transport**, or **littoral drift**. Longshore transport moves a tremendous amount of sediment along coasts (both oceans and large lakes) around the world, and it is responsible for creating a variety of depositional features that we will discuss in section 13.4. The net movement of sediment due to longshore transport is to the south along both coasts of the continental United States, because the storms and high winds that originally create the swell tend to occur at higher latitudes and move to the south.

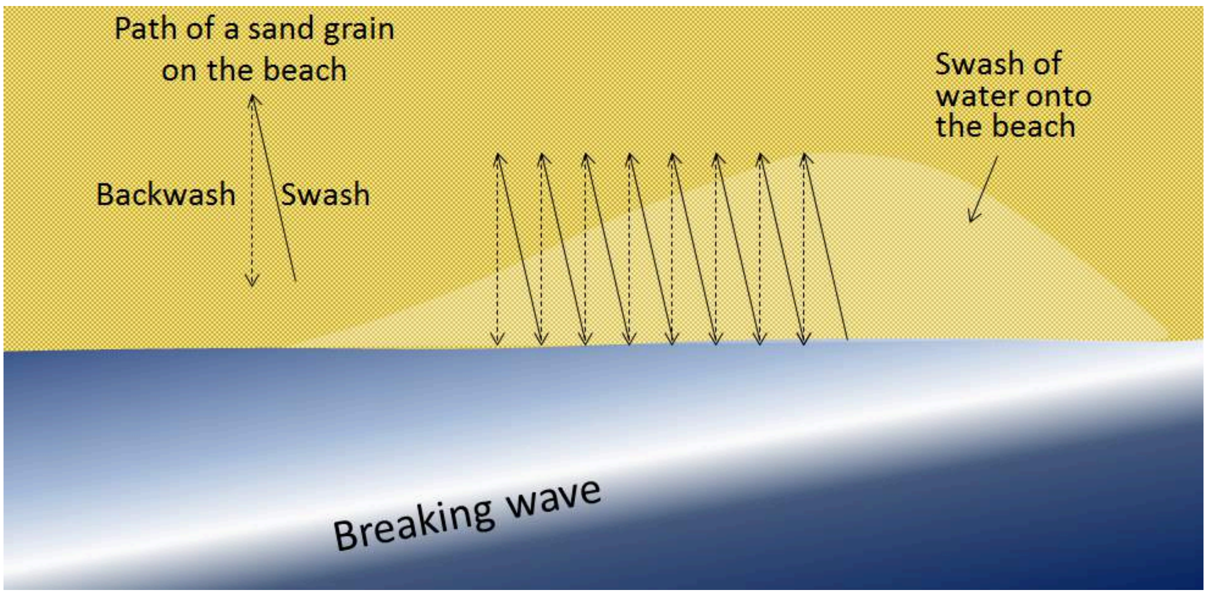


***Figure 13.2.2** The zigzag pattern of sediment movement along a beach creating longshore transport. In this figure the longshore transport moves particles to the left (Steven Earle, "Physical Geology").*

A **rip current** (often incorrectly called a “rip tide”; they are not really related to tides) is another type of current that develops in the nearshore area, and has the effect of returning water that has been pushed up to the shore by incoming waves or accumulated through longshore currents, particularly converging longshore currents. Rip currents often occur where there is a channel between sandbars that makes it easier for the retreating water to escape. As shown in Figure 13.2.3, rip currents flow straight out from the shore, and because the water is directed through a narrow space, the current can be very strong. The currents lose strength quickly just outside of the surf zone, but they can be dangerous to swimmers who get caught in them and are pulled away from shore. Swimmers caught in a rip current should not try to swim directly back to shore, as it is difficult to fight the current and the swimmer can quickly tire. Instead, swim parallel to the beach for a short distance until you are outside of the rip current, and then you can easily swim to shore.

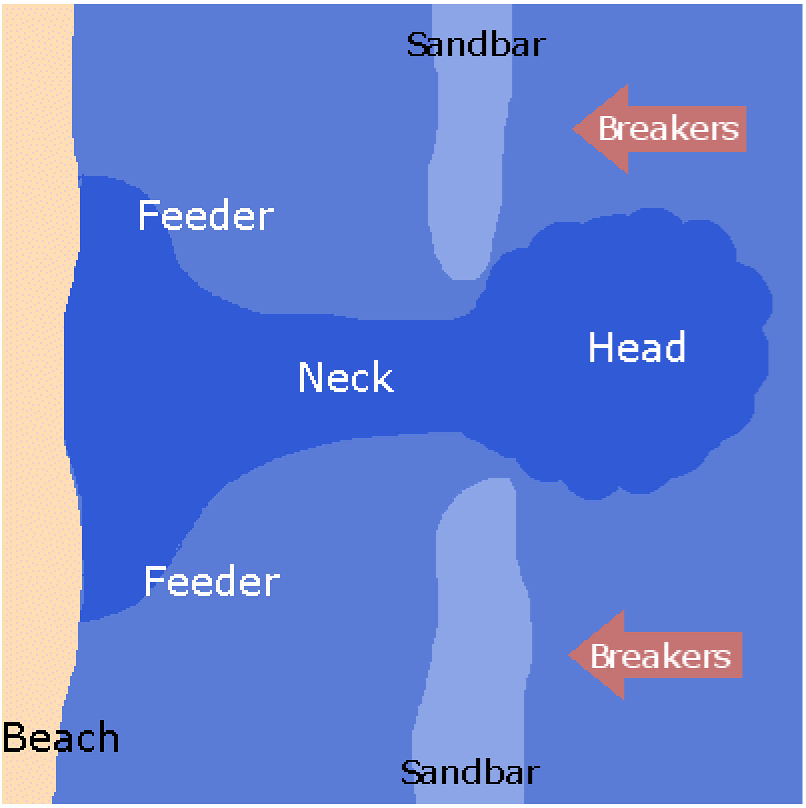


Figure 13.2.3 *Creation of a rip current from wave action and longshore transport. Water accumulates on the beach, and then rushes out to sea through a narrow channel, creating a strong current (National Weather Service, Wilmington, NC (NOAA) [Public domain], via Wikimedia Commons).*

Rip currents are visible in Figure 13.2.4, a beach at Tunquen in Chile near Valparaiso. As is evident from the photo, the rips correspond with embayments in the beach profile. Three of them are indicated with arrows, but it appears that there may be several others farther along the beach.

Figure 13.2.4 *Rip currents along a beach in Chile, indicated by the arrows. Longshore currents converging in a curved beach have nowhere to go but straight back out to sea, creating a rip current (Steven Earle, "Physical Geology").*

13.3 Landforms of Coastal Erosion

*Modified from "Physical Geology" by Steven Earle**

Large waves crashing onto a shore bring a tremendous amount of energy that has a significant eroding effect, and several unique erosion features commonly form on rocky shores with strong waves.

When waves approach an irregular shore, they are slowed down to varying degrees, depending on differences in the water depth, and as they slow, they are bent or refracted (section 10.3). In Figure 13.3.1, wave energy is represented by the blue arrows. That energy is evenly spaced out in the deep water, but because of refraction, the energy of the waves is being focused on the **headlands**. On irregular coasts, the headlands receive much more wave energy than the intervening bays, and thus they are more strongly eroded. The result of this is **coastal straightening**, where an irregular coast will eventually become straightened, although that process may take millions of years.

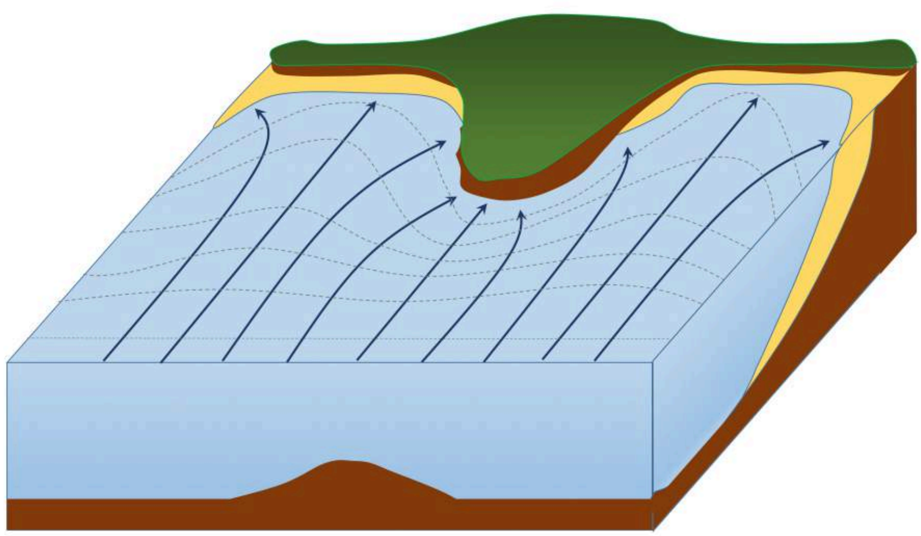


Figure 13.3.1 *The approach of waves (blue lines) towards a coastal headland. The blue arrows represent wave energy; most of that energy is focused on the headlands, causing greatest erosion in this area (PW).*

Wave erosion is greatest in the surf zone, where the wave base is impinging strongly on the seafloor and where the waves are breaking. The result is that the substrate in the surf zone is typically eroded to a flat surface known as a **wave-cut platform** (or **wave-cut terrace**) (Figure 13.3.2). A wave-cut platform extends across the intertidal zone.

Figure 13.3.2 *A wave-cut platform in bedded sedimentary rock on Gabriola Island, B.C. The wave-eroded surface is submerged at high tide (Steven Earle, "Physical Geology").*

Arches and **sea caves** form as a result of the erosion of relatively non-resistant rock. Wave action and strong longshore currents can carve a cave into a headland, and if the erosion extends all the way through, it becomes an arch. If a hole develops in the ceiling of a cave, a **blowhole** can be created, shooting water into the air when waves crash in the cave. An arch in the Barachois River area of western Newfoundland, Canada, is shown in Figure 13.3.3. This feature started out as a sea cave, and then, after being eroded from both sides, became an arch. During the winter of 2012-2013, the arch collapsed, leaving a small stack at the end of the point.

Figure 13.3.3 *Top: An arch in tilted sedimentary rock at the mouth of the Barachois River, Newfoundland, July 2012. Bottom: The same location in June 2013. The arch has collapsed and a small stack remains (Photo: Dr. David Murphy, used with permission in Steven Earle, "Physical Geology").*

The tower of rock left behind from a collapsed arch is called a **sea stack** (Figure 13.3.4). But sea stacks can also form during the formation of wave-cut platforms or other features, when relatively resistant rock that does not get completely eroded remains behind to form the stack.

Figure 13.3.4 *A sea stack, likely created from the collapse of a sea arch (Doug Lee [CC BY-SA 2.0], via Wikimedia Commons).*

Figure 13.3.5 summarizes the process of transformation of an irregular coast into a straightened coast with **sea cliffs** (wave-eroded escarpments) and the remnants of stacks, arches, and wave-cut platforms. The next stages of this process would be the continued landward erosion of the sea cliffs and the complete erosion of the stacks and wave-cut platforms in favor of a continuous and nearly straight sandy beach.

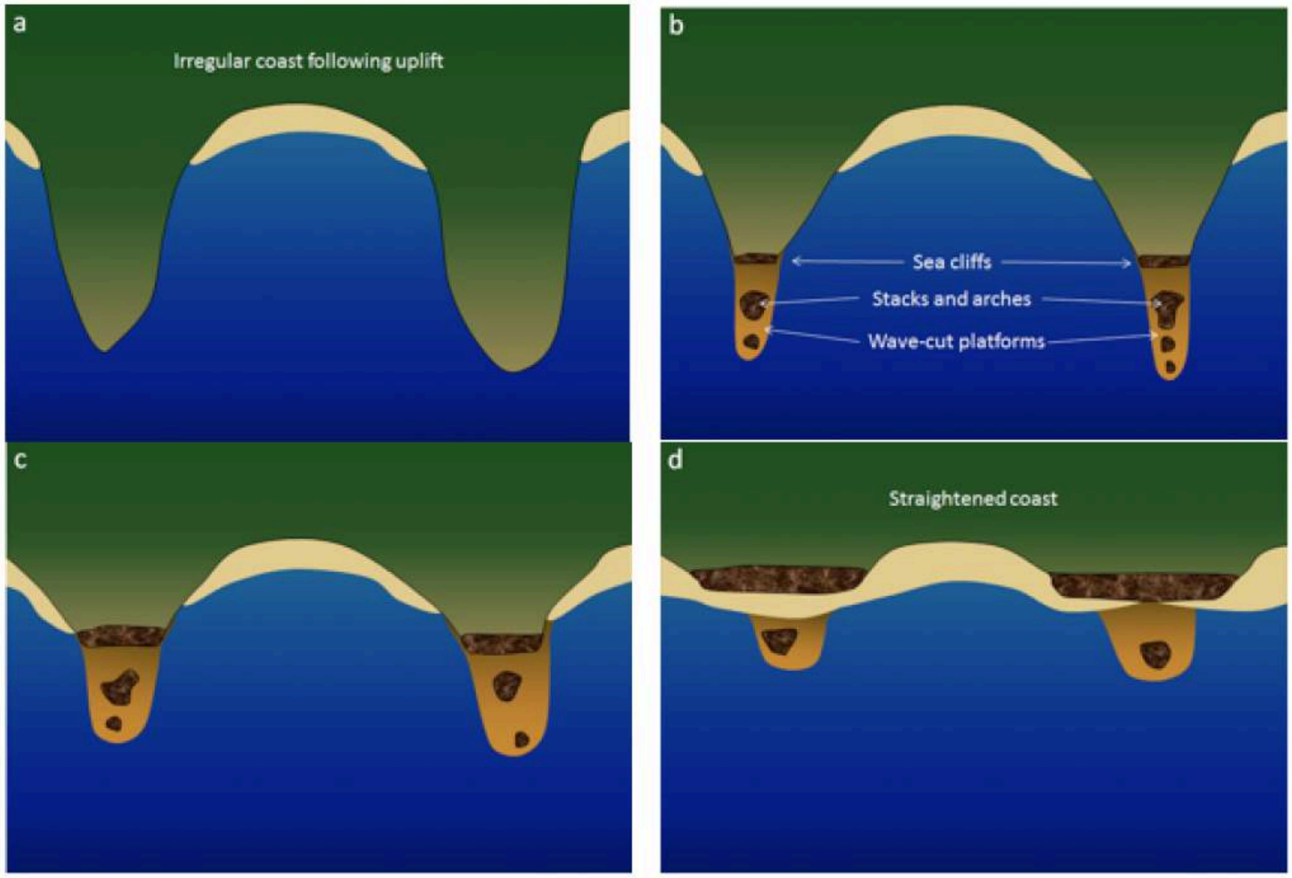


Figure 13.3.5 *Evolution of a straightened coast through the erosion to stacks and arches, sea cliffs, and wave-cut platforms (Steven Earle, "Physical Geology").*

13.4 Landforms of Coastal Deposition

*Modified from "Physical Geology" by Steven Earle**

Some coastal areas are dominated by erosion, an example being the Pacific coast of North America, while others are dominated by deposition, examples being the Atlantic and Caribbean coasts of the United States. But on almost all coasts, both deposition and erosion are happening to varying degrees most of the time, although in different places. On deposition-dominant coasts, the coastal sediments are still being eroded from some areas and deposited in others.

On coasts that are dominated by depositional processes, most of the sediment being deposited typically comes from large rivers. Much of the sediment is immediately deposited at the mouth of the river, creating large fan-shaped **deltas**. An obvious example is where the Mississippi River flows into the Gulf of Mexico at New Orleans; another is the Yellow (Huang He) River in China (Figure 13.4.1).

Figure 13.4.1 *The Yellow River delta in China, created by one of the most sediment-laden rivers on Earth (NASA [Public domain], via Wikimedia Commons).*

The evolution of sandy depositional features on sea coasts is primarily influenced by waves and currents, especially longshore currents. As sediment is transported along a shore, either it is deposited on beaches, or it creates other depositional features. A **spit**, for example is an elongated sandy deposit that extends out into open water in the direction of a longshore current (Figure 13.4.2).

__Figure 13.4.2__ Farewell Spit, on the northern shore of New Zealand's South Island (By NASA/GSFC/METI/ERSDAC/JAROS, and U.S./Japan ASTER Science Team (NASA's Earth Observatory) [Public domain], via Wikimedia Commons).

A spit that extends across a bay to the extent of closing, or almost closing it off, is known as a **baymouth bar** (Figure 13.4.3). Most bays have streams flowing into them, and since this water has to get out, it is rare that a baymouth bar will completely close the entrance to a bay.

__Figure 13.4.3__ Left: Illustration of a baymouth bar and tombolo (Steven Earle, "Physical Geology"). Right: A baymouth bar at the mouth of the Klamath River in northern California (Linda Tanner, https://www.flickr.com/photos/goingslo/5827465324, Creative Commons CC-BY 2.0).

Tombolos are common where islands are abundant, and they typically form where there is a wave shadow behind a nearshore island (Figure 13.4.4). This becomes an area with reduced energy, and so the longshore current slows and sediments accumulate. Eventually enough sediments accumulate to connect the island to the mainland with a tombolo (Figure 13.4.5).

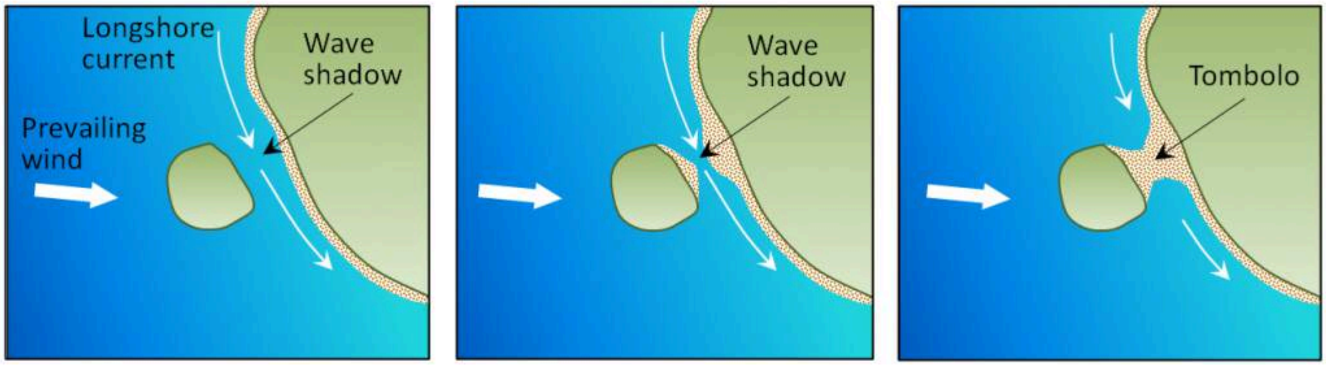


Figure13.4.4 *Formation of a tombolo. In the wind shadow of an island, there is little wave action, so the sediment moved by longshore transport gets deposited, eventually linking the island to the mainland (Steven Earle, "Physical Geology").*

Figure 13.4.5 *A stack (with a wave-cut platform) connected to the mainland by a tombolo, Leboeuf Bay, Gabriola Island, British Columbia (Steven Earle, "Physical Geology").*

In areas where coastal sediments are abundant and coastal relief is low (because there has been little or no recent coastal uplift), it is common for **barrier islands** to form (Figure 13.4.6). Barrier islands are elongated islands composed of sand that form offshore from the mainland, potentially reaching several kilometers wide and hundreds of kilometers long. They are common along the U.S. Gulf Coast from Texas to Florida, and along the U.S. Atlantic Coast from Florida to Massachusetts. The islands often form as the result of sediment moving offshore through river discharge, while wave action works to push the sediment back towards the shore. The resulting sediment buildup is then stretched into long barrier islands by longshore transport.

Figure 13.4.6 *Assateague Island on the Maryland coast, U.S. This barrier island is about 60 km long and only 1 km to 2 km wide. The open Atlantic Ocean is to the right and the lagoon is to the left. This part of Assateague Island has recently been eroded by a tropical storm, which pushed massive amounts of sand into the lagoon. (http://soundwaves.usgs.gov/2014/04/images/DelmarvaAssateague_aerial_ViewCV.jpg).*

Mature barrier islands contain a number of ecological zones. Beginning on the ocean side of the island there is a beach, consisting of the zones we discussed in section 13.1. Behind the beach lie dunes that are built up by sand transported by wind. The ocean side of the dunes are home to grasses and other plants which help stabilize the sand from erosion, and also help slow down the wind to allow sand to settle and accumulate. Beyond the dunes lies a more heavily vegetated barrier flat, covered by larger shrubs and trees that are tolerant to the high winds and salty conditions. As the land slopes down on the side of the island facing the mainland, the low-lying areas transition into a salt marsh or mud flat habitat, which is protected from wave action, but is influenced by tidal changes. The mud flats are colonized by grasses, which slow down the movement of water and lead to increased sediment deposition, building up the land in the marsh. Different species of grasses eventually dominate the different elevations of the salt marsh, depending on their tolerance for submersion in seawater. These salt marshes are very important habitats for many invertebrates, birds, and juvenile fish. Between the island and the mainland lies a lagoon, which usually contains brackish water from the mixing of fresh water runoff from the land and the seawater within a somewhat enclosed space. Barrier islands, although attractive locations for beach houses, are not permanent structures, and people should be wary of building on them. Over time, the erosion on the seaward side of the island, and the expansion of the marsh on the landward side, causes the island to slowly move towards the mainland, eventually closing off the lagoon. Maintaining dune grasses is one way to slow this movement, and as we will see in the next section, people have developed a number of other strategies to try to curtail the natural erosion of beaches.

13.5 Human Interference with Shorelines

The continued erosion and deposition of coastal sediments is a natural process, with features forming and disappearing as sea level and other conditions change. However, we have also come to enjoy and rely on many of these beaches and other coastal features for commerce, recreation, and living space. So from our perspective, we often see the transient nature of the coast as a threat to our activities, and as a result we have developed a number of ways to try and influence the erosion process, usually through **hard stabilization**, or the building of structures to stop the flow of sand. While some of these efforts have been successful, many others have actually exacerbated the problem, as we will see below.

Groins (or groynes) are barriers built perpendicular to the shore (Figure 13.5.1 *left*). Groins are built to interrupt longshore transport and trap sand upstream of the groin, which they do very well. But downstream of the groin the source of replacement sand is cut off, while sand continues to be removed, so erosion can become even more pronounced on that side of the groin. To prevent that erosion, another groin must be built downstream of the first one, which then creates its own erosion problems, leading to another groin, and so on. Eventually a beach may become covered in a series of groins, called a **groin field**, all trying to stabilize the natural flow of sand (Figure 13.5.1 *right*).

Figure 13.5.1 *Left: A groin on Bolinas Beach, California. Longshore transport would be moving from right to left, accumulating sand on the right of the groin (Beatrice Murch, https://www.flickr.com/photos/blmurch/952523971, CC BY-SA 2.0). Right: A groin field. Longshore transport would move from the bottom of the picture towards the top (https://pxhere.com/en/photo/1048378, Public Domain, CC0).*

Jetties are like longer groins, often built to protect the mouths of harbors to prevent them from filling with sand. Because they are longer they can trap more sand than groins, and they also can contribute to increased erosion on the downstream side (Figure 13.5.2). If too much sand accumulates upstream of the jetty it can spread past the jetty and into the mouth of the harbor, in which case the jetty may need to be extended.

Figure 13.5.2 The jetties protecting Santa Cruz harbor, California. Sand has accumulated on the left (north) side of the harbor, and the beach is eroded on the right (south) side (Google Maps, Map Data: Google).

Breakwaters are walls that are usually built parallel to the shore. Their purpose is not necessarily to interfere with sediment transport, but instead to protect the areas behind them from heavy wave action, so they are often deployed at the mouth of a harbor, or to protect the boats in a mooring field. But breakwaters do have an unintended impact on sediment distribution. Longshore transport continues to move sand along the beach, but once it gets behind the breakwater the lack of wave action interrupts the flow, and the sand settles and accumulates. The beach grows behind a breakwater, until eventually they may become connected (Figure 13.5.3). With the longshore transport interrupted, increased erosion can occur downstream of the breakwater.

Figure 13.5.3 A series of breakwaters along the coast in Skagen, Denmark. In this case, sand has accumulated to the point that they now connect the breakwaters to land (Google Maps; Imagery: DigitalGlobe, Aerodata,International Surveys, TerraMetrics, Data SIO, NOAA, US Navy, NGA, GEBCO, Map Data: Google).

Santa Monica Pier

The beach around the Santa Monica Pier in California provides a good example of the effects of breakwaters on a sandy shore. A breakwater was constructed in the early 1930s to protect the pier and the boats that moored near it. Following the construction, the once-straight beach became much wider behind the breakwater as sand accumulated in the absence of strong wave action (*below left*). Now that the breakwater is no longer in place, the bulge in the shoreline is gone, and the beach is much straighter once again (*below right*).

Left; aerial image of Santa Monica Pier and breakwater from 1936 (Courtesy of Santa Monica Public Library Image Archives/ Spence Air Photos). Right; Santa Monica Pier in 2011 (© JCS, CC BY-SA 3.0, via Wikimedia Commons).

Seawalls are constructed at the top of the surf zone, where the waves crash against the shore. The walls are designed as a barrier between the waves and the shore, to prevent the land from being eroded (Figure 13.5.4). They are often utilized in beachfront property to prevent the ground under a home from being undermined by the waves. However, as with the other forms of hard stabilization we have discussed, seawalls are not without their own environmental consequences. The sudden release of wave energy on a seawall can create turbulence, which undermines the sediment at the base of the wall and causes it to erode. Furthermore, on a softer, natural coastline some of the wave energy is absorbed or dissipated, but with a hard seawall most of the wave energy is reflected, leading to stronger longshore currents and faster erosion. In many places where seawalls have been built the beaches are getting steeper, and erosion rates have increased, with the potential for seawalls to collapse along with whatever they are supporting. Because of this, some coastal communities are phasing out seawall construction to try to return to more natural beach fronts.

Figure 13.5.4 *A sea wall at Horsey Gap, UK (Evelyn Simak [CC BY-SA 2.0], via Wikimedia Commons).*

13.6 Estuaries

Estuaries are partially enclosed bodies of water where the salt water is diluted by fresh water input from land, creating **brackish** water with a salinity somewhere between fresh water and normal seawater. Estuaries include many bays, inlets, and sounds, and are often subject to large temperature and salinity variations due to their enclosed nature and smaller size compared to the open ocean.

Estuaries can be classified geologically into four basic categories based on their method of origin. In all cases they are a result of rising sea level over the last 18,000 years, beginning with the end of the last ice age; a period that has seen a rise of about 130 m. The rise in sea level has flooded coastal areas that were previously above water, and prevented the estuaries from being filled in by all of the sediments that have been emptied into them.

The first type is a **coastal plain estuary**, or **drowned river valley**. These estuaries are formed as sea level rises and floods an existing river valley, mixing salt and fresh water to create the brackish conditions where the river meets the sea. These types of estuaries are common along the east coast of the United States, including major bodies such as the Chesapeake Bay, Delaware Bay, and Narragansett Bay (Figure 13.6.1). Coastal plain estuaries are usually shallow, and since there is a lot of sediment input from the rivers, there are often a number of depositional features associated with them such as spits and barrier islands.

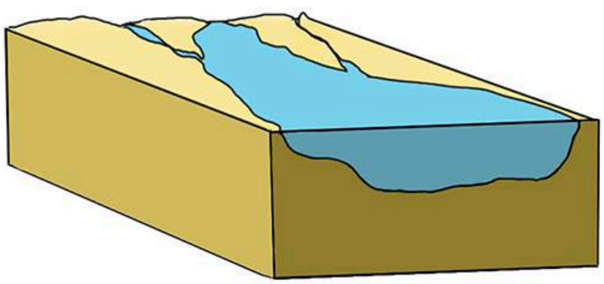


__Figure 13.6.1__ A coastal plain estuary. Sea level has risen and flooded what was once a river valley. The satellite image shows Chesapeake Bay and Delaware Bay, two coastal plain estuaries (left: JR, right: NASA, Public Domain via Wikimedia Commons).

The presence of sand bars, spits, and barrier islands can lead to **bar-built estuaries**, where a barrier is created between the mainland and the ocean. The water that remains inside the sand bar is cut off from complete mixing with the ocean, and receives freshwater input from the mainland, creating estuarine conditions (Figure 13.6.2).

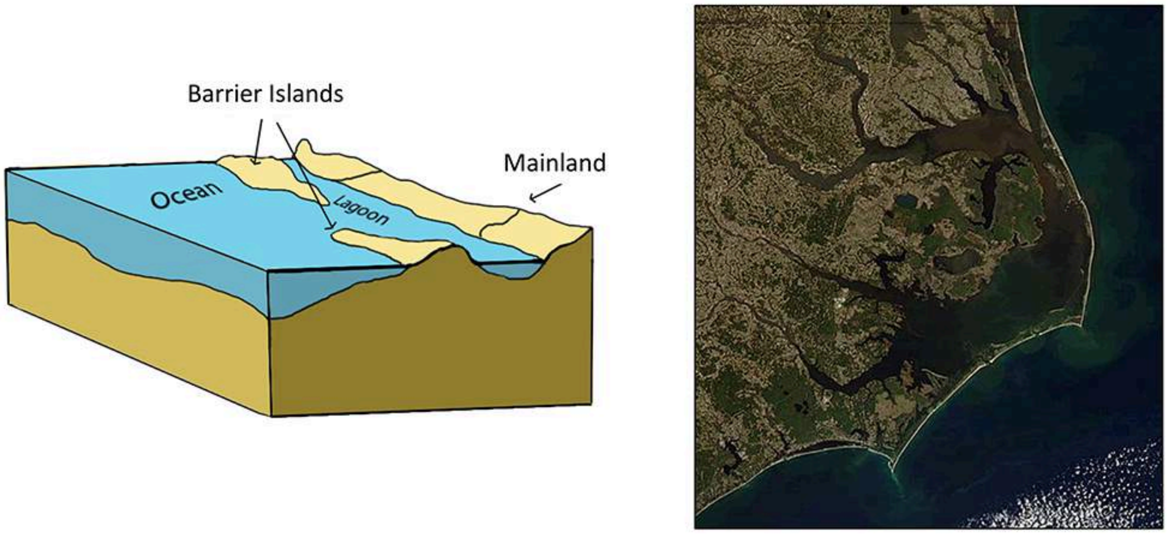


Figure 13.6.2 A bar-built estuary. Sand bars and barrier islands have partially isolated a lagoon from the rest of the ocean. Freshwater input into the lagoon from the mainland creates brackish conditions in the estuary. At right is a satellite image of Pamlico Sound, North Carolina, a bar-built estuary surrounded by spits and barrier islands (left: JR, right: NASA, Public Domain via Wikimedia Commons).

Fjords are estuaries formed in deep, U-shaped basins that were carved out by advancing glaciers. When the glaciers melted and retreated, sea level rose and filled these troughs, creating deep, steep-walled fjords (Figure 13.6.3). Fjords are common in Norway, Alaska, Canada, and New Zealand, where there are mountainous coastlines once covered by glaciers.

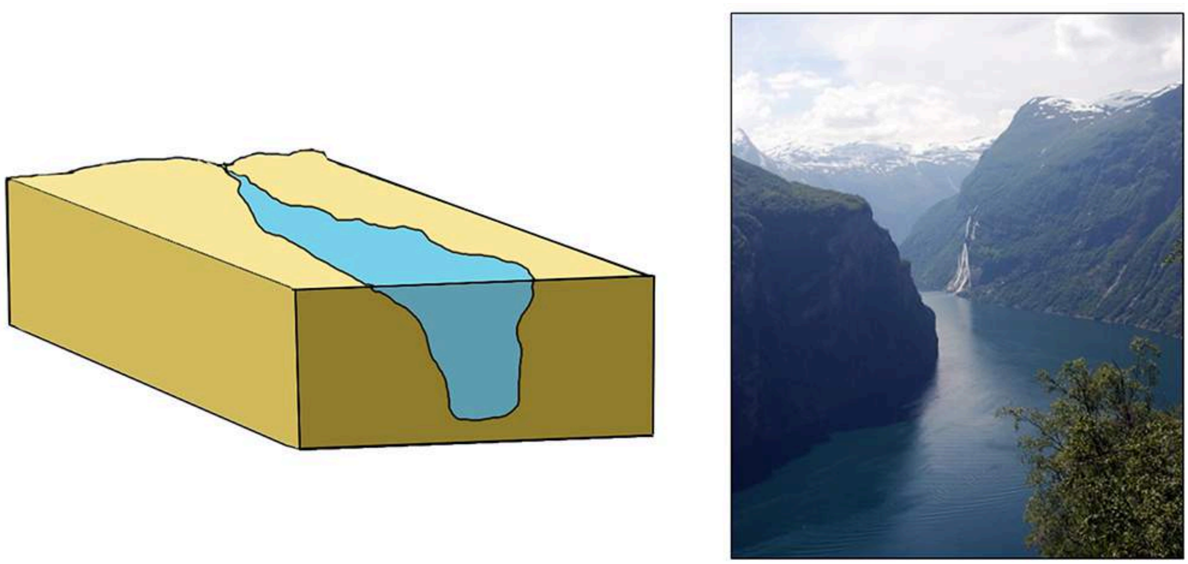


Figure 13.6.3 A fjord is a deep estuary that was carved out by glacial movements. At right is Geirangerfjord, Norway (left: JR, right: Fgmedia, [CC-BY-SA-3.0], via Wikimedia Commons).

Tectonic estuaries are the result of tectonic movements, where faulting causes some sections of the crust to subside, and those lower elevation sections then get flooded with seawater. San Francisco Bay is an example of a tectonic estuary (Figure 13.6.4).

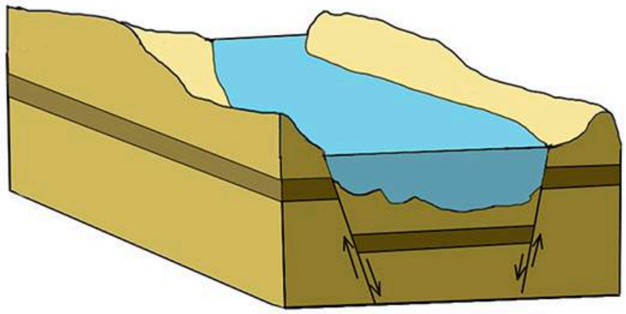


Figure 13.6.4 *A tectonic estuary, formed from the subsidence of crust along fault lines, and the subsequent filling by seawater. San Francisco Bay is a tectonic estuary, shown at right (left: JR, right: USGS, Public Domain, via Wikimedia Commons).*

Estuaries are also classified based on their salinity and mixing patterns. The amount of mixing of fresh and salt water in an estuary depends on the rate at which fresh water enters the head of the estuary from river input, and the amount of seawater that enters the estuary mouth as a result of tidal movements. The input of fresh water is reflected in the **flushing time** of the estuary. This refers to the time it would take for the in-flowing fresh water to completely replace all the fresh water currently in the estuary. Seawater input is measured by the **tidal volume**, or **tidal prism**, which is the average volume of sea water entering and leaving the estuary during each tidal cycle. In other words, it is the volume difference between high and low tides. The interaction between the flushing time, tidal volume, and the shape of the estuary will determine the extent and type of water mixing within the estuary.

In a **vertically mixed**, or **well-mixed** estuary there is complete mixing of fresh and salt water from the surface to the bottom. In a particular location the salinity is constant at all depths, but across the estuary the salinity is lowest at the head where the fresh water enters, and is highest at the mouth, where the seawater comes in. This type of salinity profile usually occurs in shallower estuaries, where the shallow depths allow complete mixing from the surface to the bottom.

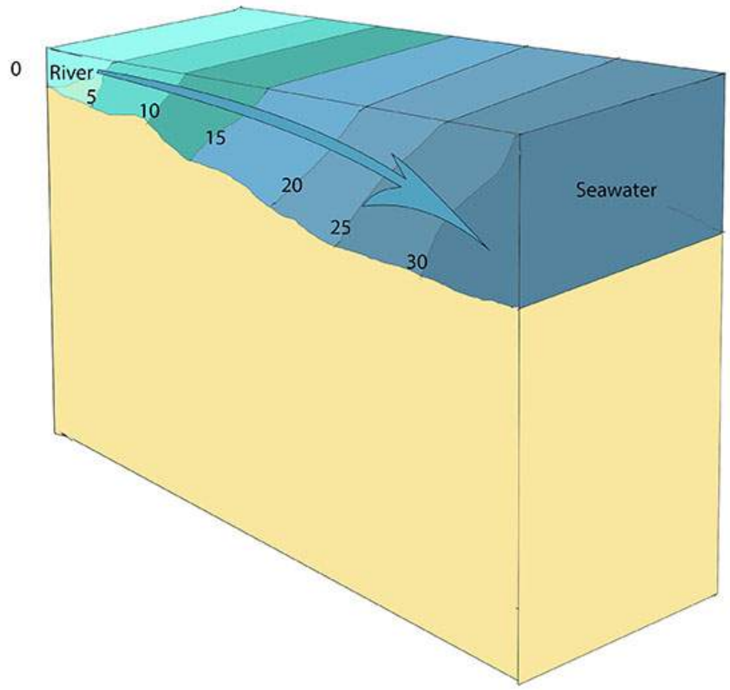


__Figure 13.6.5__ A well-mixed estuary. The shallow basin allows nearly complete mixing or fresh and seawater from top to bottom. Salinities are in ppt (JR).

Slightly stratified or **partially mixed** estuaries have similar salinity profiles to vertically mixed estuaries, where salinity increases from the head to the mouth, but there is also a slight increase in salinity with depth at any point. This usually occurs in deeper estuaries than those that are well-mixed, where waves and currents mix the surface water, but the mixing may not extend all the way to the bottom.

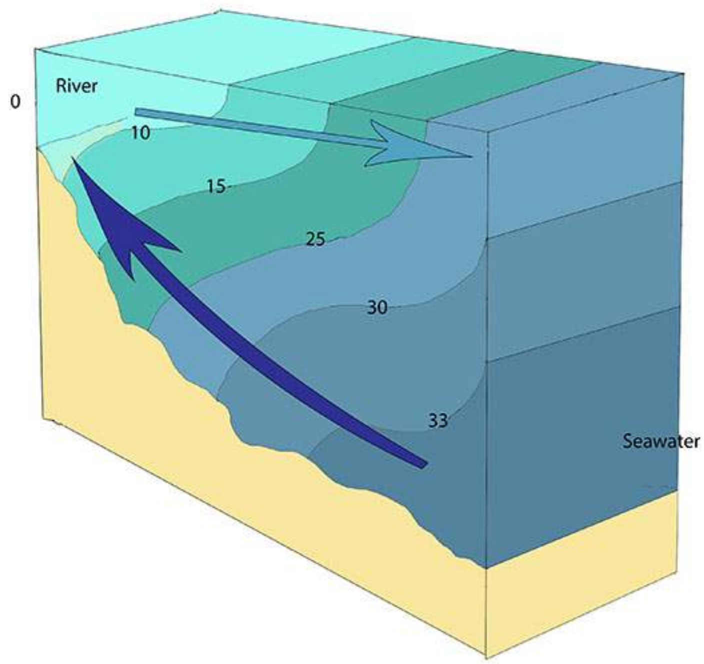


Figure 13.6.6 *A slightly-stratified estuary. Generally deeper than a well-mixed estuary, the inflow of seawater (dark blue arrow) and fresh water (light blue arrow) create an estuary where salinity increases with depth, and at the surface when moving from the head to the mouth of the estuary. Salinities are in ppt (JR).*

A **salt wedge** estuary occurs where the outflow of fresh water is strong enough to prevent the denser ocean water to enter through the surface, and where the estuary is deep enough that surface waves and turbulence have little mixing effect on the deeper water. Fresh water flows out along surface, salt water flows in at depth, creating a wedge shaped lens of seawater moving along the bottom. The surface water may remain mostly fresh throughout the estuary if there is no mixing, or it can become brackish depending on the level of mixing that occurs.

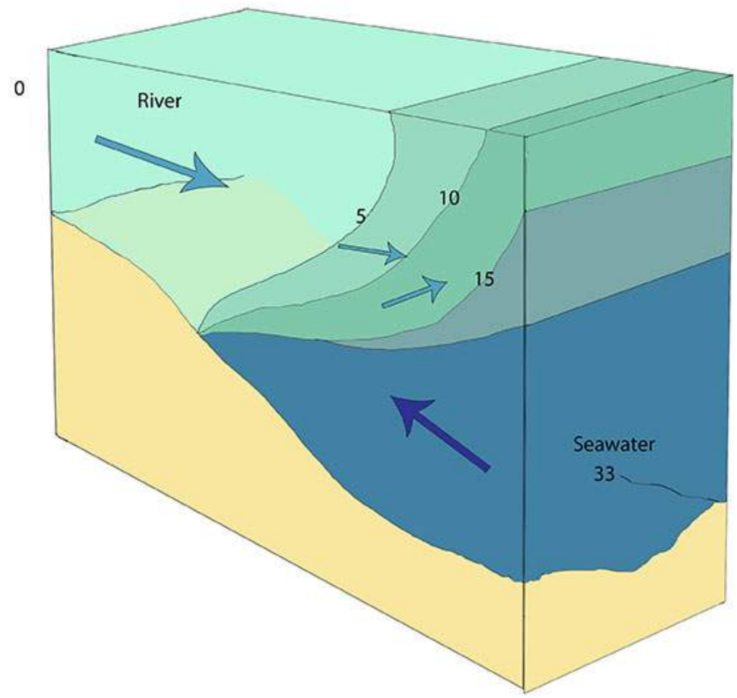


Figure 13.6.7 *A salt wedge estuary. Strong river outflow (light blue arrows) creates a layer of mostly fresh water that sits on top of a wedge of encroaching seawater along the bottom (dark blue arrow). Salinities are in ppt (JR).*

Highly stratified profiles are found in very deep estuaries, such as in fjords. Because of the depth, mixing of fresh and salt water only occurs near the surface, so in the upper layers salinity increases from the head to the mouth, but the deeper water is of standard ocean salinity.

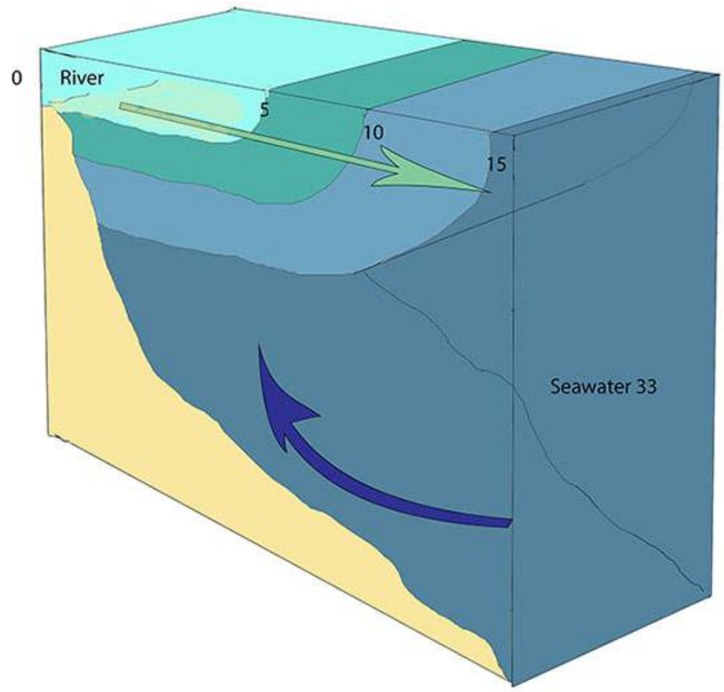


Figure 13.6.8 *A highly stratified estuary. Strong river outflow and a deep basin prevent mixing between surface and bottom water, creating an estuary that is vertically stratified. Salinities are represented in ppt (JR).*

Estuaries are very important commercially, as they are home to the majority of the world's metropolitan areas, they serve as ports for industrial activity, and a large percentage of the world's population lives near estuaries. Estuaries are also very important biologically, especially in their role as the breeding grounds for many species of fish, birds, and invertebrates.

13.7 Sea Level Change

*Modified from "Physical Geology" by Steven Earle**

Sea level change has been a feature on Earth for billions of years, and it has important implications for coastal processes, estuaries, and both erosional and depositional features. There are two main mechanisms of sea level change, eustatic and isostatic, as described below.

Eustatic sea level changes are global sea level changes related to changes in the volume of water in the ocean. These can be due to changes in the volume of glacial ice on land, thermal expansion of the water, or to changes in the shape of the seafloor caused by plate tectonic processes. For example, seafloor spreading widens an ocean basin, thus changing its volume and affecting sea level.

Over the past 20,000 years, there has been approximately 125 m of eustatic sea level rise due to glacial melting. Most of that took place between 15,000 and 7,500 years ago during the major melting phase of the North American and Eurasian Ice Sheets (Figure 13.7.1). At around 7,500 years ago, the rate of glacial melting and sea level rise decreased dramatically, and since that time, the average rate has been in the order of 0.7 mm/year.

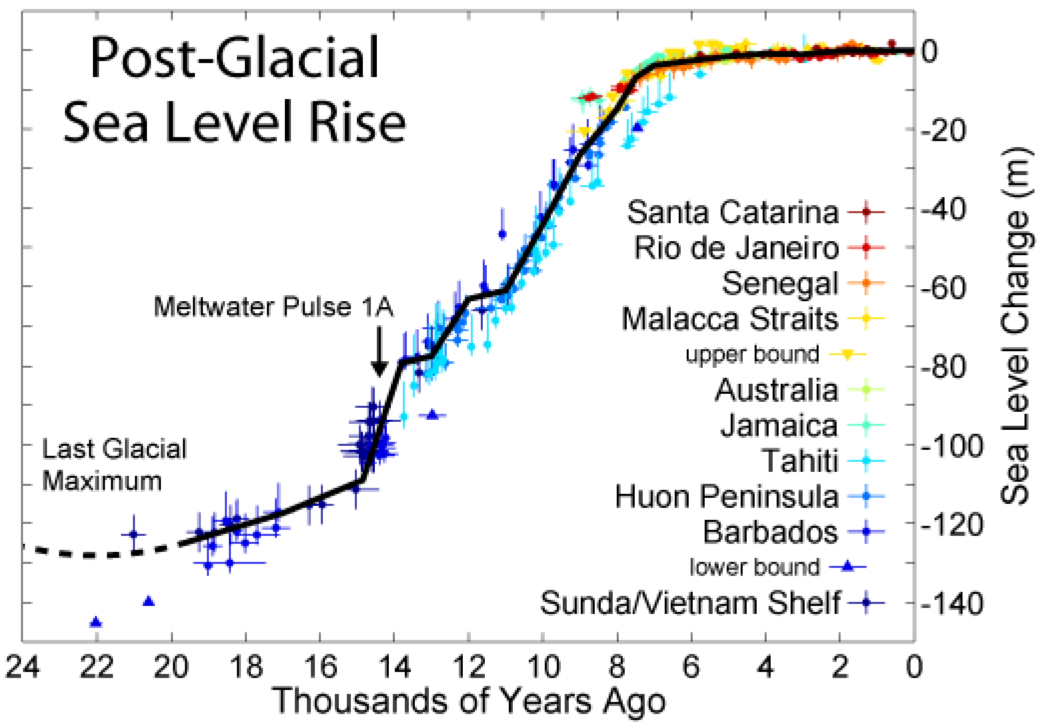


Figure 13.7.1 *Sea level rise resulting from the melting of glacial ice over the past 24,000 years (Robert A. Rhode, CC BY-SA 3.0, via Wikimedia Commons).*

Anthropogenic climate change led to accelerating sea level rise starting around 1870. Since that time, the average rate has been 1.1 mm/year, but it has been gradually increasing. Since 1992, the average rate has been 3.2 mm/year (Figure 13.7.2). Much of this is due to increased glacial melting as the global climate gets warmer (section 14.3), but a large part is due to **thermal expansion** of the water. As water warms, the molecules gain more kinetic energy and move faster and farther apart; the result is that the same amount of water now takes up more space. So even without the input of new water from melting ice, warming ocean temperatures will cause sea level to rise.

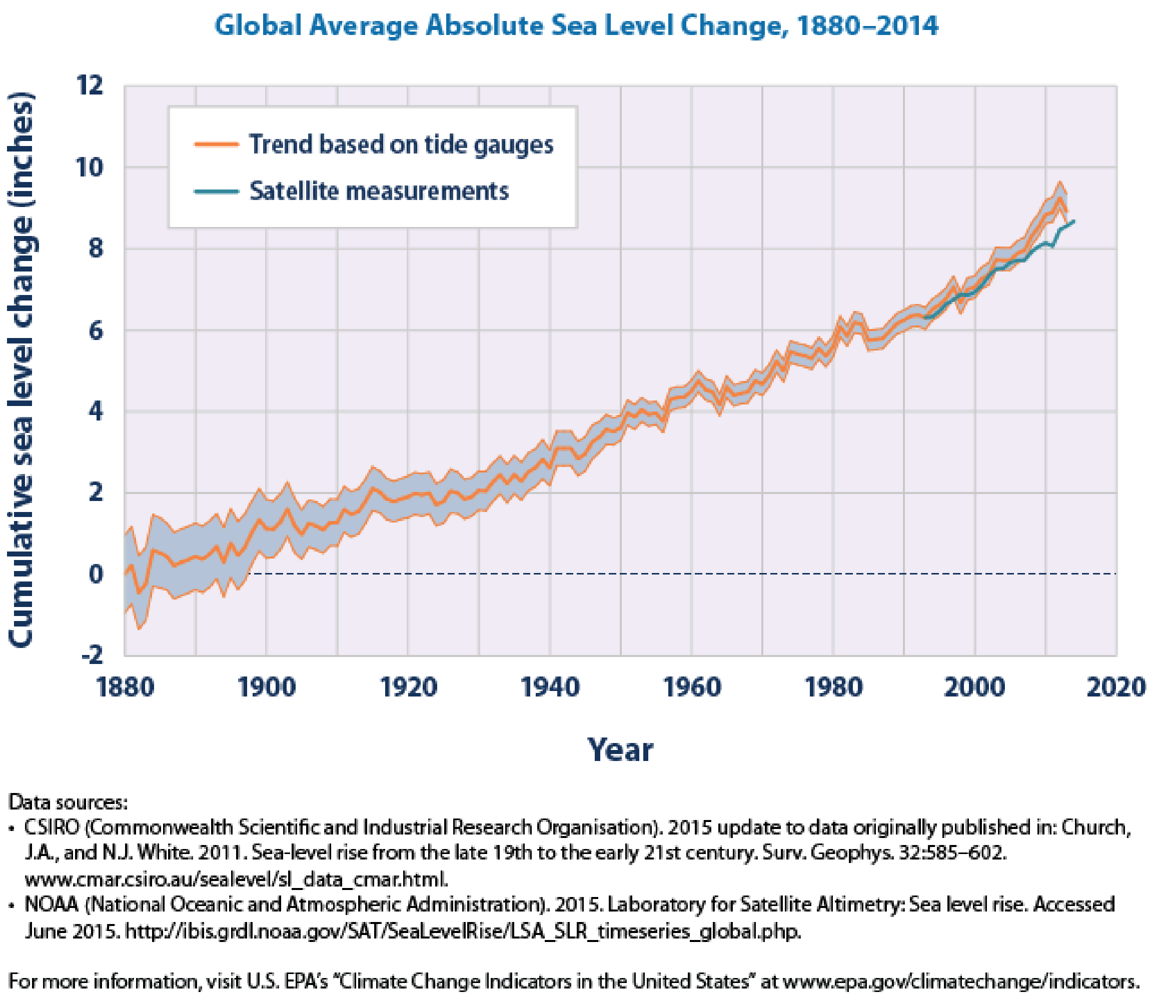


Figure 13.7.2 *Average absolute sea level change, which refers to the height of the ocean surface, regardless of whether nearby land is rising or falling. Satellite data are based solely on measured sea level, while the long-term tide gauge data include a small correction factor because the size and shape of the oceans are changing slowly over time (US EPA [Public domain], via Wikimedia Commons).*

Isostatic sea level changes are local changes caused by subsidence or uplift of the crust related either to changes in the amount of ice on the land, or to growth or erosion of mountains. Almost all of Canada and parts of the northern United States were covered in thick ice sheets at the peak of the last glaciation. Following the melting of this ice, there has been an isostatic rebound of continental crust in many areas. This ranges from several hundred meters of rebound in the central part of the Laurentide Ice Sheet (around Hudson Bay) to 100 m to 200 m in places such as Vancouver Island and the mainland coast of British Columbia. In other words, although global sea level was about 130 m lower during the last glaciation, the glaciated regions were depressed at least that much in most places, and more than that in places where the ice was thickest. Tectonic processes, such as the uplift of crust, can also cause localized changes in sea level.

CHAPTER 14: ICE

Chapter 14: Ice

Learning Objectives

After reading this chapter you should:

- understand the roles played by ice in global climate
- understand the difference between glacial ice and sea ice
- understand the steps in the formation of sea ice
- know the difference between pack and fast ice
- understand the process of iceberg formation
- know the ways that icebergs are classified
- understand how global ice cover is responding to climate change

Ice covers about 6% of the Earth's surface, and is responsible for some of the most beautiful and spectacular features in the oceans. Yet what is less appreciated is ice's influence on global weather and climate. For example:

- Ice cover affects the albedo (reflectivity) of the Earth's surface, influencing the amount of solar radiation that is absorbed.
- Ice insulates the ocean from heat exchange with the atmosphere, preventing large fluctuations in polar water temperatures. One meter of ice cover reduces heat exchange between Earth and the atmosphere by 100 times.
- Seasonal freezing and thawing of ice moderates ocean temperatures through the latent heat of freezing and fusion.
- In all of these ways, ice contributes to the differential heating of Earth's surface, which serves as the basis for global wind and climate patterns.

We have discussed these processes in previous chapters, so in this section we will examine the process of ice formation, the different types of ice that are found in the ocean, and the impacts of climate change on global ice cover.


(By NASA / James Yungel (Iceberg in sea ice) [Public domain], via Wikimedia Commons).

14.1 Types of Ice

The ice that is seen floating on the ocean's surface comes from one of two sources. **Glacial ice** is formed from the accumulation and compression of snow into glaciers, that then break apart and release ice to the ocean. Because glaciers can be several kilometers thick the icebergs that break off of them can be very large; so the tall icebergs at sea always come from glacial ice sheets. **Sea ice** refers to the ice formed from the freezing of sea water, and rarely exceeds a thickness of several meters (Figure 14.1.1). Sea ice covers about 7% of the ocean at any time, and makes up about 66% of the Earth's permanent ice cover by area, but only 0.1% of the ice in terms of volume. This is because sea ice is a vast but thin sheet of cover compared to the glacial ice caps that are more localized but may be several kilometers thick.

Figure 14.1.1 *An iceberg (glacial ice) embedded in a thinner layer of sea ice (NASA / James Yungel [Public domain], via Wikimedia Commons).*

Sea ice cover around Antarctica fluctuates between about 21 million km^2 in winter to around 1.3 million km^2 in the summer, with most Antarctic sea ice lasting only a year. Seasonal changes in ice cover are less pronounced in the Arctic, from about 14 million km^2 in winter to 6.5 million km^2 in summer. About half of the sea ice in the Arctic lasts more than a year to become multi-year ice. This difference arises because Antarctica is surrounded by water, so the ice expands into warmer water and eventually melts. The Arctic Ocean is enclosed by continents, so only about 10% of the ice escapes into the Atlantic between Greenland and Spitzbergen. The rest is trapped and becomes multi-year ice or perennial ice, averaging around 7 years old, and 3-5 m thick, compared to first year ice at 1-2 m thick.

Sea ice formation

Because of the salt content, seawater begins to freeze at about –1.8^o C, a lower temperature than for fresh water. Ice formation begins at the surface with the formation of small needle-like ice crystals called **frazil**, which accumulate and make the water appear slushy and cloudy; this stage is referred to as **grease ice** (Figure 14.1.2 A). In calmer water these small crystals can freeze together into a thin surface layer called **nilas**, which can reach a thickness of up to 10 cm (Figure 14.1.2 B).

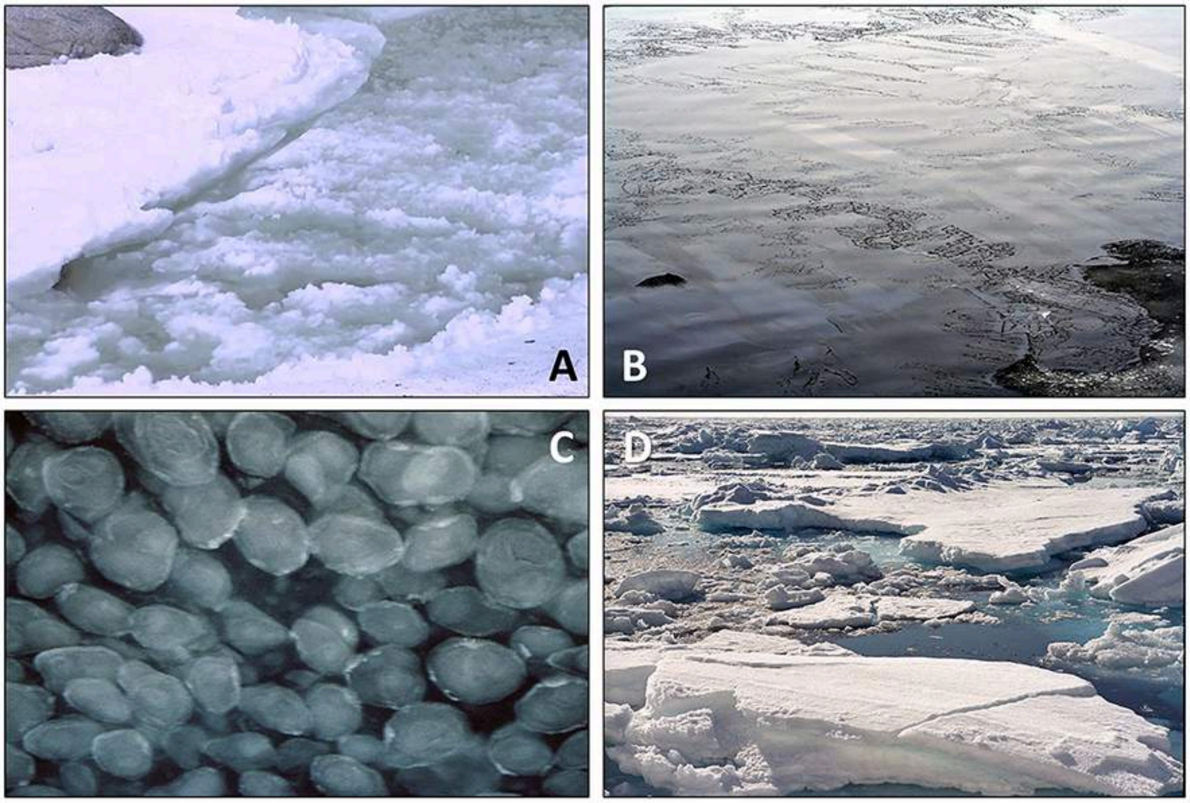


Figure 14.1.2 *Stages in sea ice formation. A) grease ice, B) nilas, C) pancake ice, D) ice floes. (A) National Park Service [Public domain], via Wikimedia Commons; B) Brocken Inaglory (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 4.0, via Wikimedia Commons; C) PW; D) Jerzy Strzelecki (Own work) [CC BY-SA 3.0], via Wikimedia Commons).*

Wave action can break up the nilas into small mats 1-2 m across, which then bump into each other and form rounded shapes with raised edges, called **pancake ice** (Figure 14.1.2 C). If temperatures remain cold the pancake ice freezes together into solid **ice floes**, a hard surface covering the ocean (Figure 14.1.2 D). Ice floes then freeze together into **ice fields**.

Once ice floes form, the water underneath becomes insulated and heat loss to the atmosphere declines, so the water no longer cools and no more ice formation occurs. As a result, young sea ice is usually relatively thin, not more than 3-4 m thick. The ice can get thicker through precipitation; there is not a lot of precipitation at the poles, but due to the low temperatures, whatever does occur tends to accumulate rather than melt. Over time the accumulated ice and snow can add to the overall thickness of the sea ice, but it will still never approach the thickness of glacial ice.

As sea ice crystals form, most of the salt is excluded, so sea ice contains much less salt than seawater, and can be melted for drinking if needed. But about 20% of the salt remains trapped in pockets of water between the ice crystals. As ice forms and salts are excluded into these pockets, the salinity of the remaining water increases and it can become too salty to freeze. These unfrozen pockets of briny water make sea ice a little softer and more slushy than fresh water ice, which is harder and more rigid. Eventually most of this brine leaks out, and the sea ice becomes more solid, but when it is “young ice” it can be more dangerous to walk on than fresh water ice of the same thickness. For example, 7-8 cm of fresh water ice is enough to support the weight of a single person, but you would need at least 15 cm of sea ice to do the same.

The very cold, dense briny water leaks out of the ice and sinks. The brine is “supercooled”; it is cooled below the normal freezing point of seawater, but remains liquid due to the high salt content. When this supercooled brine comes into contact with the surrounding water, it causes the water around it to freeze, creating hollow ice stalactites, or "brinicles" that can be several meters long. Brine continues to flow through the hollow brinicle, and the brinicle grows downwards (see below for an amazing time-lapse video of brinicle formation).

A YouTube element has been excluded from this version of the text. You can view it online here: https://rwu.pressbooks.pub/webboceanography/?p=437

When the large sheets of sea ice are formed, they exist in one of two formations. **Fast ice**, or **land-fast ice**, refers to the large, solid ice sheets that are attached to land. The **pack ice** consists of the smaller, free-floating pieces of sea ice. They may have formed independently, or may have broken off from the fast ice (Figure 14.1.3).

Figure 14.1.3 *Fast ice (left) and pack ice (right). (Left: Peterfitzgerald (Own work) [CC BY-SA 3.0], via Wikimedia Commons; Right: Markus Trienke, https://www.flickr.com/photos/mtrienke/34281559366/in/photostream/ [CC BY-SA 2.0]).*

The floating pack ice dampens waves and currents, buffering the sea surface from motion. So changes in the distribution of the pack ice can lead to changes in current patterns, and even alter ecosystem structure. But the pack ice is also subject to the currents flowing underneath, and the ice sheets are constantly on the move, breaking up or being pushed together. When pieces of ice converge, they often buckle and crack, or override one another as in convergent lithospheric plate boundaries (section 4.6). These collisions can create tall, jagged pressure ridges, that may extend for several kilometers, and that create hazards for polar explorers navigating the ice (Figure 14.1.4).

Figure 14.1.4 *Pressure ridges formed from colliding ice sheets (Michael Studinger [Public domain], via Wikimedia Commons).*

In the polar oceans, ice cover is not uniform. There are a number of areas where there is consistently open water, even though the areas surrounding them are covered by ice. These regions of persistent open water are called **polynyas** (Figure 14.1.5). Polynyas may be the result of currents or winds moving the ice, or areas of warmer water that prevent ice formation. In Figure 14.1.5, very strong winds blowing offshore from Antarctica's interior have created a polynya near the edge of the ice sheet.

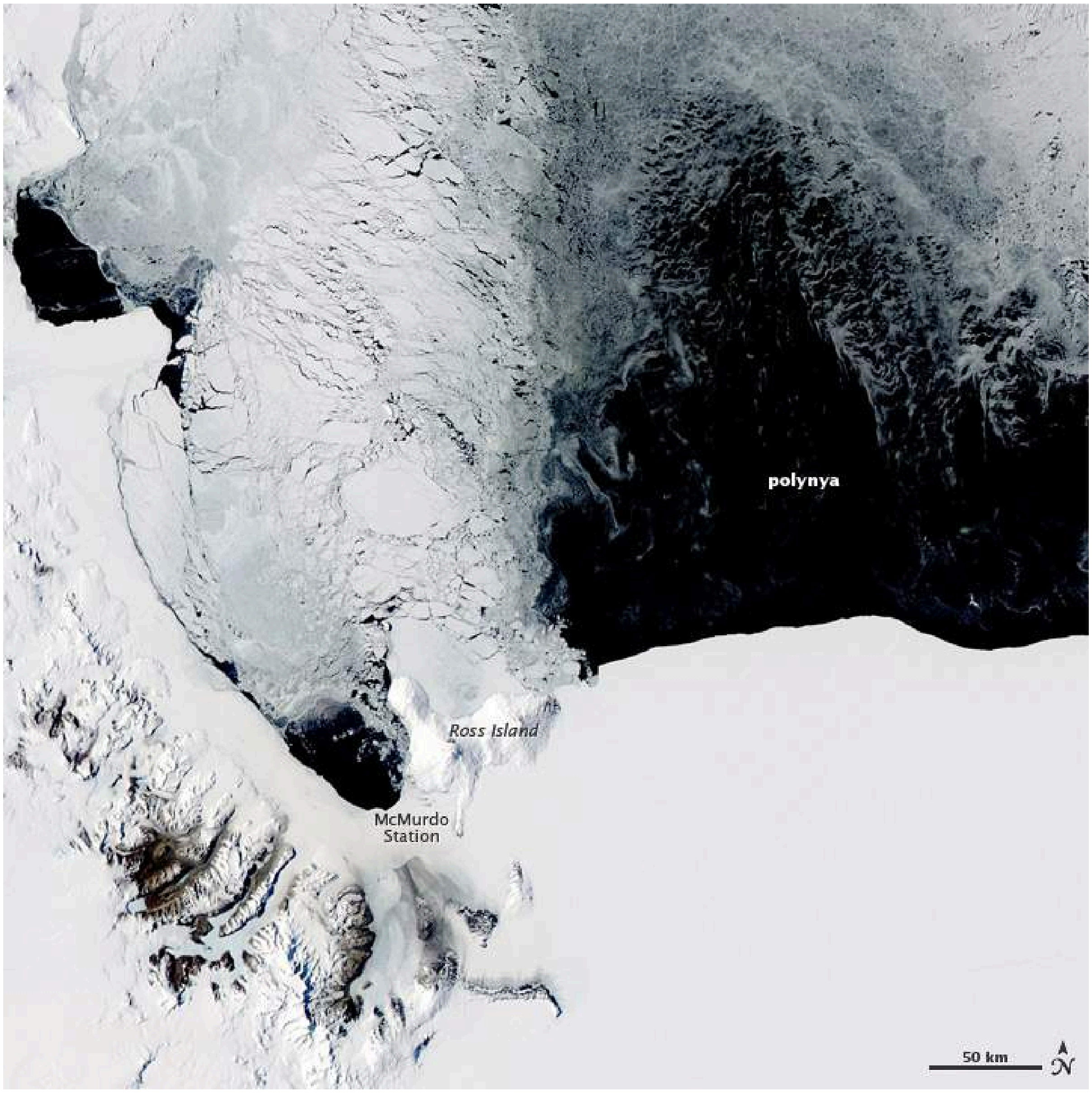


Figure 14.1.5 *A polynya near McMurdo Station in Antarctica (NASA's Earth Observatory [CC BY 2.0], via Wikimedia Commons).*

14.2 Icebergs

Many of the most spectacular ice formations in the ocean take the form of icebergs. Remember that the largest icebergs are not made of sea ice; they are floating pieces of glaciers that have broken off, or **calved** from the glacier tongue, and thus they are composed of fresh water ice.

Glacial ice forms from layers of snow that accumulate over time. The weight of the accumulated layers compresses the snow into a granular form known as **firn** that has a higher density and less air than regular snow. As the pressure from the continued snow accumulation increases, the firn is compressed into even denser glacial ice (Figure 14.2.1). Most icebergs appear white because the ice still contains a lot of air bubbles that scatter all of the wavelengths of white light. But icebergs composed of older ice, or highly compressed ice from deep within a glacier can appear a deep shade of blue (Figure 14.2.1). This ice contains much less air, and larger, denser ice crystals that absorb the longer (red) wavelengths of light, and transmit and scatter the shorter, blue wavelengths. The longer the path that light travels through the ice, the more the long wavelengths are absorbed, and the bluer the ice appears.

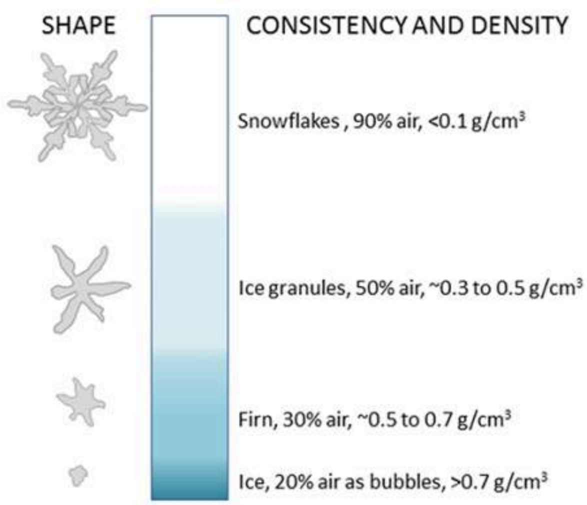


***Figure 14.2.1** Left: Changes in ice crystal structure and density as they are compacted from snow into glacial ice (Steven Earle, "Physical Geology"). Right: A blue iceberg (By Doug Knuth [CC BY-SA 2.0], via Wikimedia Commons).*

Icebergs come in a wide range of sizes, from a few meters across to hundreds of kilometers long. Table 14.2.1 shows the size categories used in the North Atlantic (in the Antarctic a different scale is used, as the icebergs tend to be larger there).

Table 14.2.1 Iceberg size classifications for the North Atlantic

Size Category	Height (m)	Length (m)
Growler	< 1	< 5
Bergy Bit	1-5	5-15
Small	5-15	15-60
Medium	16-45	61-120
Large	46-75	121-200
Very Large	> 75	> 200

Icebergs can also be classified based on their shapes (Figure 14.2.2). The primary distinction is between tabular and non-tabular icebergs. **Tabular icebergs** have steep sides and a flat top, and the length is greater than five times the height. **Non-tabular icebergs** include any icebergs that are not tabular, and these can be further subdivided into a few categories. **Domed** icebergs have a rounded top, **pinnacled** icebergs have tall spires, and **wedge** icebergs have a steep face next to a more gradually sloping side. **Drydock** icebergs have a water-covered channel running through it, potentially large enough for boats to pass through; hence the name. Finally, **blocky** icebergs have a flat top and steep sides, but their length to height ratio is not as great as it is for a tabular iceberg. Regardless of the shape of an iceberg, what we see above the surface only represents about 10% of the total mass, with the rest of the ice remaining submerged (see box below).

Figure 14.2.2 *Icebergs can be classified based on their shapes. Clockwise from top left: tabular, domed, pinnacled, blocky, drydock, and wedge icebergs. (Tabuar: Michael Clarke Stuff [CC BY-SA 2.0 via Wikimedia Commons]; domed: AWeith [CC BY-SA 4.0] via Wikimedia Commons; pinnacled: Povl Abrahamsen, https://www.flickr.com/photos/povl/13339227703 [CC-BY-NC 2.0]; wedge: https://pixabay.com/en/iceberg-st-john-s-newfoundland-2008688/ [CC0 Public domain]; drydock: Brocken Inaglory [CC-BY-SA-3.0] via Wikimedia Commons; blocky: Jerzy Strzelecki [CC-BY 3.0] via Wikimedia Commons).*

How much of an iceberg is visible above the surface?

Archimedes' Principle states that the upwards buoyant force of an object in water is equal to the weight of the water displaced by the object. We know that the density of an iceberg (ρ_i) is around 0.917 g/cm^3. The weight (w_i) of the iceberg is equal to its mass (m_i) x the acceleration due to gravity (g), which is 9.8 m/s^2. Since density = mass/volume (V), the weight of the iceberg is:

$$w_i = \rho_i V_i g$$

Since the iceberg is floating at equilibrium, the weight of the iceberg (w_i) is equal to the weight of the water it displaces (w_W). The weight of water displaced is therefore:

$$w_W = \rho_W V_W g$$

where the density of seawater (ρ_W) is about 1.024 g/cm^3. Now we have:

$$w_i = w_W$$

$$\rho_i V_i g = \rho_W V_W g$$

so $V_W = (\rho_i/\rho_W)V_i = (0.917/1.024)V_i = 0.89\ V_i$.

In other words, the volume of water displaced (V_W) is equal to about 89% of the volume of the iceberg (V_i). This means that 89% of the iceberg is submerged, leaving around 11% of the ice exposed above the surface.

Icebergs in the Arctic tend to be smaller and more non-tabular than Antarctic icebergs, because most of them formed as irregular chunks of glaciers that calved off into the ocean, primarily around Greenland. It is estimated that the glaciers around Greenland and the Canadian Arctic calve off 300 billion cubic meters of icebergs each year. Because Antarctica has much larger ice sheets, the icebergs produced in the Southern Hemisphere are usually larger than those in the Arctic, and they tend to be more tabular in shape, as entire sections of the ice sheet break off at once. The largest iceberg ever recorded, the B-15 iceberg, broke off of the Antarctic ice sheet in 2000. It measured 295 km long and 37 km wide, giving it an area of about 11,000 square kilometers; about the size of Connecticut. Over the next few years B-15 drifted along the Antarctic coast before being broken up into smaller pieces (Figure 14.2.3).

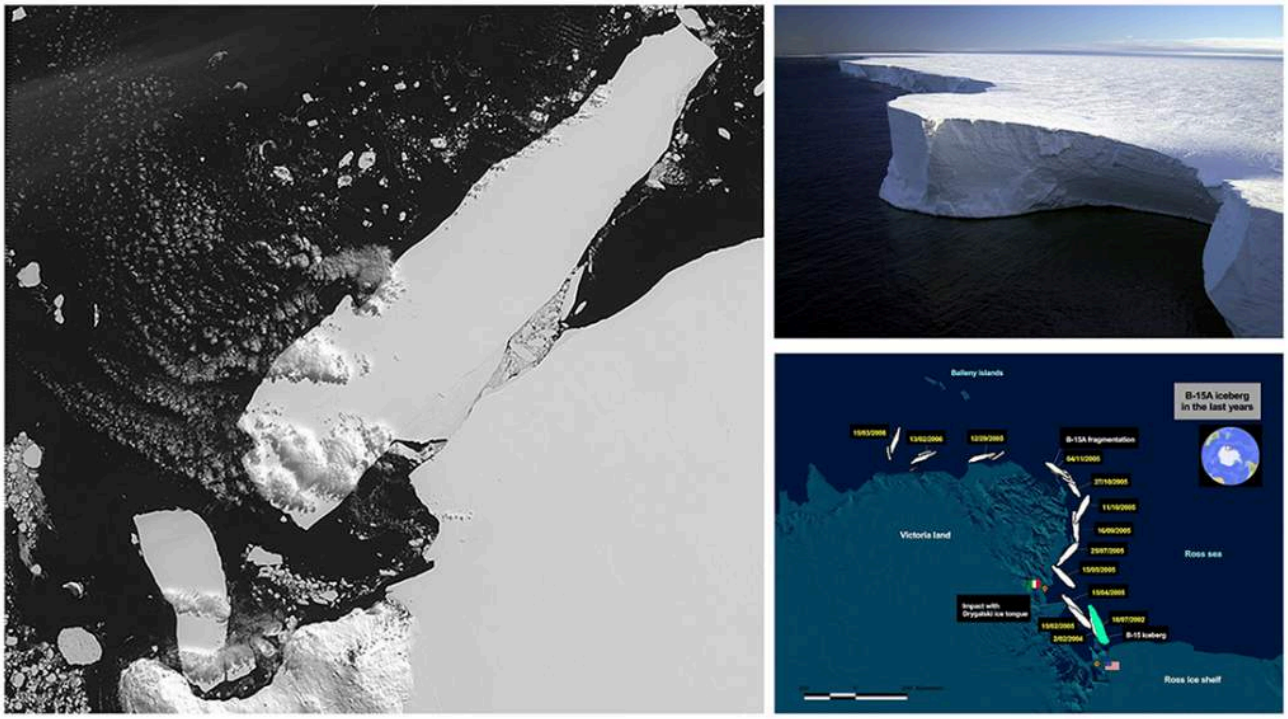


Figure 14.2.3 *Left: Iceberg B-15 soon after breaking off of the Antarctic ice sheet in 2000 (NASA/GSFC/Robert Bindschadler [Public domain], via Wikimedia Commons). Top right: Aerial view of a section of B-15A, a piece of the B-15 iceberg in January 2002 (NSF/Josh Landis, [Public domain], via Wikimedia Commons). Bottom right: Map showing the path of the B-15A iceberg over several years (Luca Pietranera [Public domain], via Wikimedia Commons).*

All of these icebergs floating through the polar oceans are potentially hazardous to ship traffic, as we all know from the tragic story of the Titanic. In response to that disaster, the North Atlantic maritime nations established the International Ice Observation and Ice Patrol Service in 1914, to monitor icebergs threatening the shipping lanes. Interestingly, the original charge of the patrol also included searching for and destroying derelict ship hulls that were still floating around, as these were also a potential hazard. While the patrols were successful at destroying these abandoned ship hulls, they achieved far less success at attempting to destroy the icebergs. Eventually they gave up on iceberg destruction and focused their efforts on monitoring iceberg movements. Today, the service still operates as the International Ice Patrol, operated by the U.S. Coast Guard. Twice each day the patrol issues alerts with the position and potential tracks of existing icebergs, and the extent of the ice cover, reporting on the roughly 600 icebergs per year that intrude below 48^{o} N latitude, the northern limit of the major shipping lines. No such iceberg patrol exists around Antarctica, as there is much less shipping traffic at those latitudes.

14.3 Ice and Climate Change

In section 8.5 we examined the causes of climate change, and the potential impacts on Earth's ecosystems. One obvious consequence of increasing global temperatures is a decline in ice cover as the warmer climate melts the ice caps. So are we seeing a decrease in ice cover, and if so, what are some consequences of this decline?

To address these questions we must once again begin by separating sea ice and glacial ice, as each type of ice is responding to climate change in its own way, with different consequences. For example, melting polar ice is often discussed in terms of its effect on raising sea levels (section 13.7). Sea level is not really impacted by the melting of sea ice, as the formation and melting of sea ice does not change the volume of water already in the ocean. The melting of glacial ice, on the other hand, takes water that was locked up in glaciers on land and returns it to the ocean. This does increase the volume of water in the oceans, and results in rising sea levels. So we will first examine the trends in sea ice cover, and then look at the situation with glacial ice.

While fluctuations in sea ice cover do not dramatically impact sea level, they can be used as indicators of global warming. Satellite data show that since 1979 the average Arctic sea ice extent declined by about 3% per decade, with the rate of decline increasing in recent years (Figure 14.3.1). The year 2017 set a record for the lowest maximum Arctic sea ice extent in the 38-year history of the satellite data, the third year in a row that the record was broken. In addition, the thickness of the Arctic sea ice has also declined, from a mean thickness of 3.64 meters in 1980 to 1.89 meters in 2008. So Arctic sea ice cover is clearly in decline as global temperatures increase.

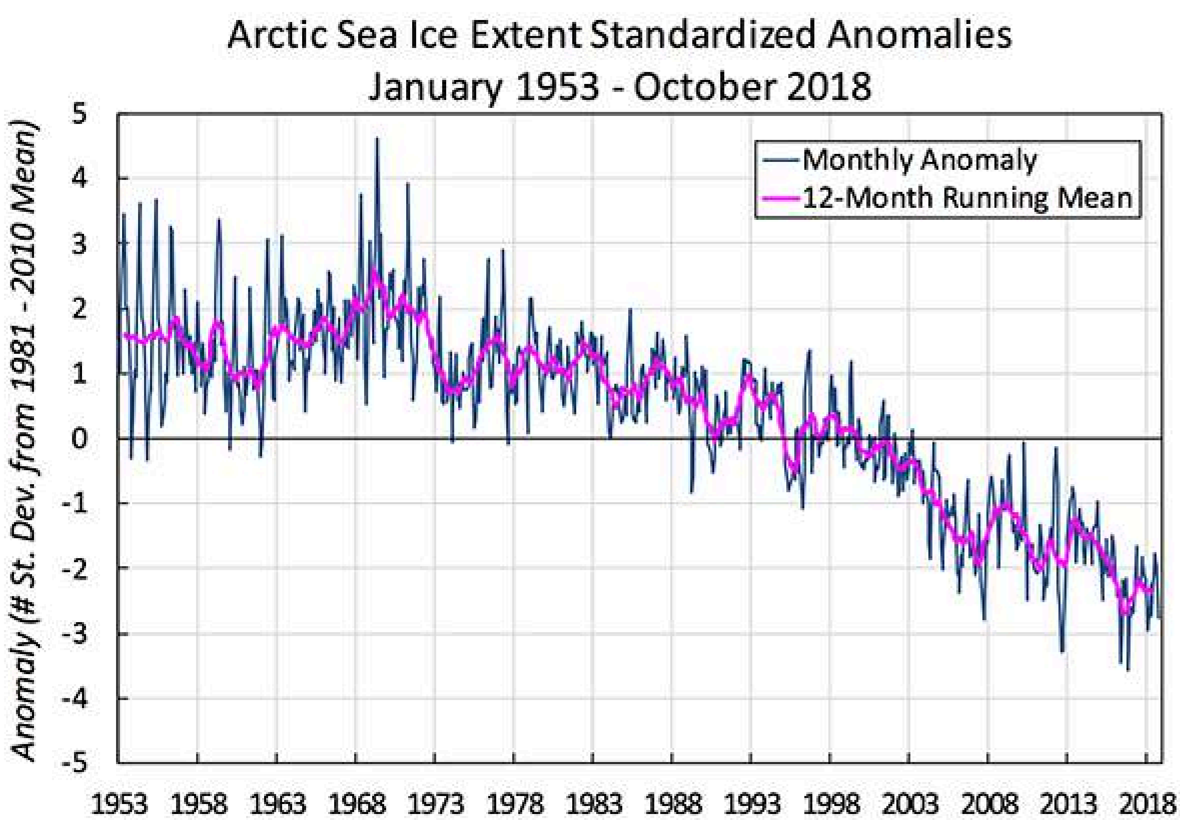


Figure 14.3.1 *Mean sea ice anomalies from 1953-2018. Positive values indicate ice cover greater than the mean from 1981-2010, negative values show ice cover below that mean (Image by Walt Meier and Julienne Stroeve, National Snow and Ice Data Center).*

The situation with Antarctic sea ice is less clear. Antarctic sea ice cover is increasing slightly, at a rate of about 1.1% per decade, although the upward trend in the Antarctic is only about one-third of the magnitude of the downward trend seen in the Arctic (Figure 14.3.2). There are several hypotheses for why this increase is occurring, including winds, currents, and La Niña conditions (section 9.6) in the Pacific Ocean. For example, the strong Antarctic Circumpolar Current prevents warmer water from reaching the ice edge, and the winds in the region keep the water very cold, promoting ice formation. One other possible explanation is that as portions of the Antarctic get warmer there is more precipitation and increased glacial ice melt. This leads to a layer of low-density, fresher water which remains at the surface and freezes more easily than saltier water. In the Arctic, there has always been significant freshwater input from rivers, so that low-density layer is already present. Whatever the reasons for the increasing sea ice in the Antarctic, the increase is still less than the decline in the Arctic, and the overall trend for global sea ice is a loss of about 35,000 km^2 per year since 1979.

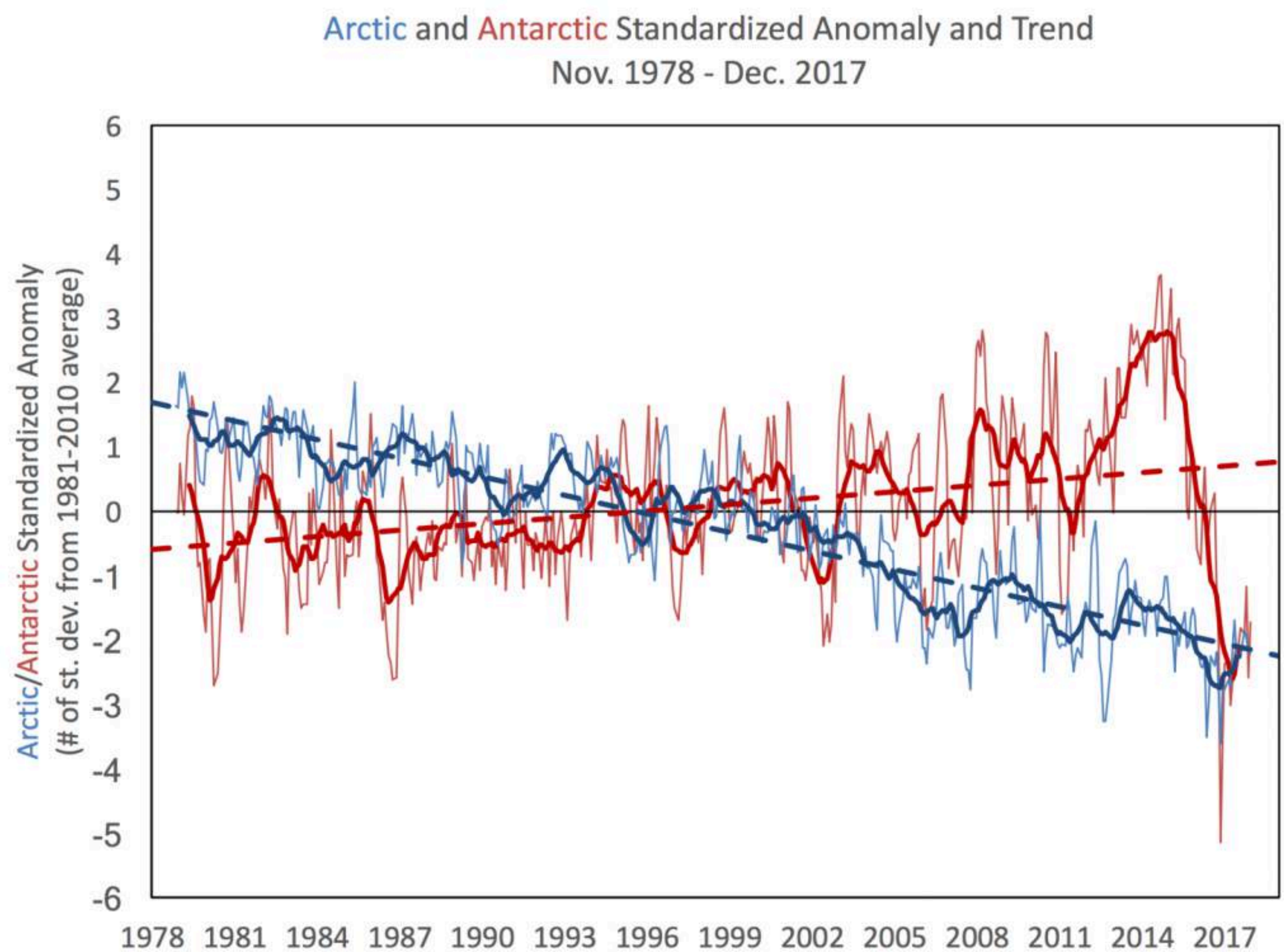


Figure 14.3.2 *Arctic and Antarctic Sea Ice Extent Anomalies, 1979-2017: Arctic sea ice extent underwent a strong decline from 1979 to 2017, but Antarctic sea ice underwent a slight increase, although some regions of the Antarctic experienced strong declining trends in sea ice extent (National Snow and Ice Data Center).*

In terms of glacial ice, once again the Arctic is showing signs of significant ice loss, particularly in Greenland; from 1979 to 2006, summer melt on the Greenland ice sheet increased by 30%. Ice melt during 2016 was the tenth highest on record (Figure 14.3.3). Once again, the trends are less defined in Antarctica. Warming and ice loss is occurring in the West Antarctic ice sheet and along the Antarctic Peninsula, which has warmed

2.5 °C since 1950. However, there is no clear warming trend in East Antarctica, and some of these areas are accumulating ice. Overall, the Antarctic ice is currently relatively stable, particularly when compared to Greenland. If the polar ice sheets were to melt, they would have a dramatic impact on sea level. It is estimated that melting of the Greenland ice would raise global sea level by about 7 m, while complete melting of the Antarctic ice sheet would result in a rise of almost 60 m!

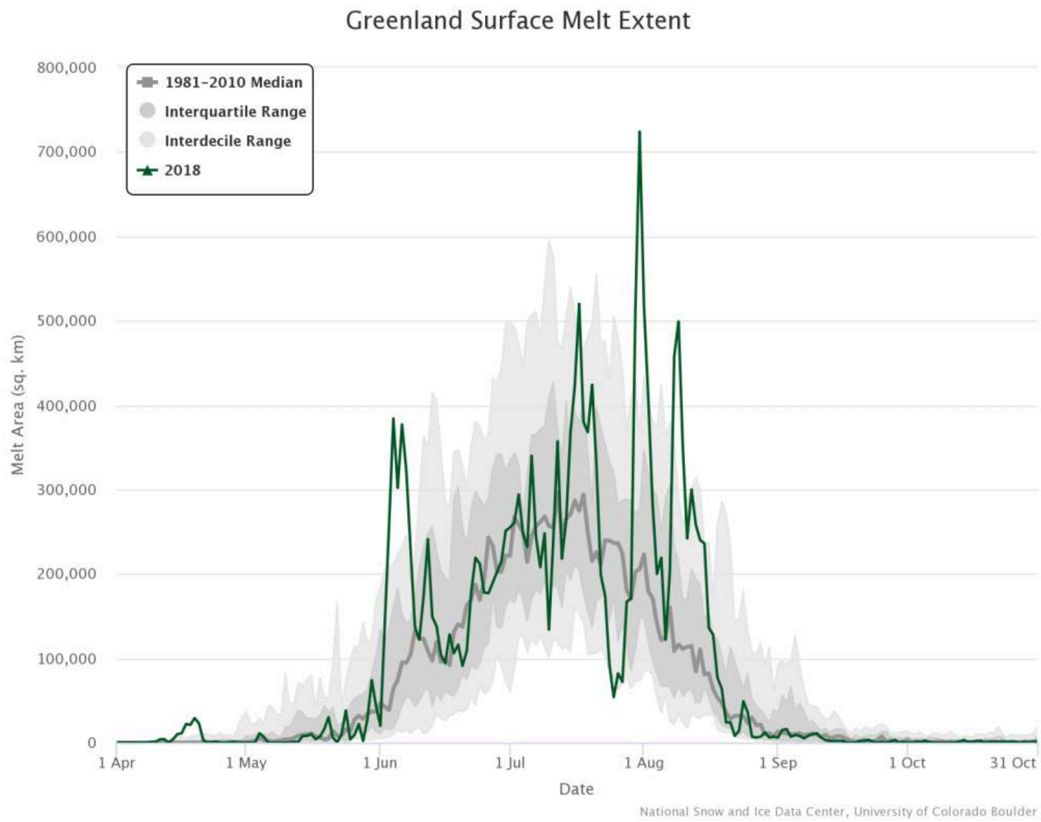


Figure 14.3.3 *2018 ice melt in Greenland. For much of the year, ice melt was significantly higher than the 1981-2010 average (National Snow and Ice Data Center).*

Additional links for more information:

- Get the latest updates on sea ice and glacial ice conditions at the National Snow and Ice Data Center website: https://nsidc.org/

Glossary

For each entry the chapter in which the word first appears is shown in parentheses

abyssal plain

the flat seafloor of the deep ocean, typically beyond the limits of the continental slopes (1.2)

abyssal zone

the region of the seafloor between 4000-6000 m (1.3)

abyssopelagic zone

the deeper parts of the open ocean, between 4000 and 6000 m; also known as the abyssalpelagic zone (1.3)

accretion (planetary)

the process by which solid celestial bodies are added to existing bodies during collisions (3.1)

active continental margin

where the boundary between the continent and the ocean is also a tectonic plate boundary (1.2)

aeolian

processes related to transportation and deposition of sediments by wind (12.2)

albedo

the reflectivity of a surface of a planet (expressed as the percentage of light that reflects from the surface) (8.1)

amphidromic circulation

gyre-sized tidal patterns where a tidal crest rotates around an ocean basin (11.2)

amphidromic point

the center point around which amphidromic circulation rotates; there is near-zero tidal range at the amphidromic point (also called a tidal node or amphidromic node) (11.2)

anoxia

conditions of zero or extremely low dissolved oxygen, usually below 0.5 mg/L (5.4)

Antarctic Bottom Water

water at abyssal depths in the ocean that forms from the sinking of dense cold water adjacent to Antarctica (9.8)

anthropogenic

resulting from the influence of humans (8.5)

aphotic zone

depths beyond 1000 m where there is no light penetration (1.3)

arch

a rock weathering remnant in the form of an arch (typically along a coast and resulting from wave erosion) (13.3)

asthenosphere

the part of the mantle, from about 100 to 200 km below surface, within which the mantle material is close to its melting point, and therefore relatively weak (3.2)

atmospheric wave

a wave formed in the atmosphere at the boundary between air masses of different densities (10.1)

atoll

a ring-shaped carbonate (or coral) reef or series of islands (4.10)

backshore

the region of the beach above the high tide line, which is only submerged under unusually high wave conditions (13.1)

backwash

the wash of wave water down the slope of a beach (13.1)

bar-built estuary

an estuary created when a sand bar or barrier island cuts off the estuary from mixing completely with seawater (13.6)

barrier island

a long, thin island parallel to the shore, created through the deposition of sand (13.4)

barrier reef

a reef that forms a barrier to waves along a coast; it is separated from land by a lagoon (4.10)

barycenter

the center of mass in the Earth-moon system around which they rotate (11.1)

basalt

a volcanic rock that makes up much of the oceanic crust (3.2)

bathyal zone

the region of the seafloor from the shelf break to 4000 m (1.3)

bathymetry

pertains to measuring the depths of the ocean (1.4)

bathypelagic zone

the moderately deep parts of the open ocean, between 1000 and 4000 m (1.3)

baymouth bar

a spit that extends across the mouth of a bay (13.4)

beach face

the area of a beach between the high and low tide lines (13.1)

Beaufort scale

a 0-12 scale describing the wind conditions at sea, often reflected in wave heights (10.2)

benthic

refers to the environment of the seafloor (1.3)

benthos

refers to the community of organisms living on or in the ocean floor (1.3)

berm

a flat area of a beach in the backshore area (above the high tide level) (13.1)

big-bang theory

the theory that the universe started with a giant expansion approximately 13.77 billion years ago (3.1)

biogenous sediment

sediment created from the remains of organisms (12.3)

blocky iceberg

iceberg with a flat top and steep sides, but their length to height ratio is not as great as it is for a tabular iceberg (14.2)

blowhole

a hole in the ceiling of an arch or sea cave through which water is ejected when waves approach (13.3)

body wave

a seismic wave that travels through rock (e.g., a P-wave or an S-wave) (3.3)

boulder

a sediment with a grain diameter of at least 256 mm (12.1)

boundary currents

ocean currents whose properties are influenced by the presence of a coastline (9.1)

Boyle's Law

the volume of a gas is inversely proportional to the pressure (6.1)

brackish

seawater of low salinity; part fresh water, part seawater (13.6)

breaker

an unstable wave that has collapsed (10.3)

breakwater

a structure built offshore in order to deflect the energy of waves (13.5)

buffer

a solution that moderates changes in pH when acids or alkalis are added to it (5.5)

calcareous sediment

sediments composed of calcium carbonate, often from the shells of marine organisms (12.3)

calving

when ice breaks off of the front of a glacier and collapses into the water (14.2)

capillary waves

small ripples that form on the water surface under light winds; their restoring force is surface tension (10.2)

carbonate compensation depth

the depth in the ocean (typically around 4000 m) below which carbonate minerals are soluble (12.6)

Carboniferous

a geologic period that spans 60 million years from the end of the Devonian Period 358.9 million years ago, to the beginning of the Permian Period, 298.9 Mya

celerity

the speed of a wave (10.1)

chemosynthesis

the creation of organic compounds using the energy from inorganic chemical reactions (4.11)

chip log

a device for determining a ship's speed at sea, by measuring the rate at which a line is unspooled when cast overboard (2.2)

clay

sediment particle that is less than 1/256 mm in diameter (12.1)

climate feedback

a process by which the physical effects of a climate forcing can have other effects (either negative or positive) on the climate (8.5)

climate forcing

a mechanism, such as a change in greenhouse gas levels, that forces the climate to change (8.5)

coastal plain estuary

an estuary formed when sea level rises and submerges a river valley (also known as a drowned river valley estuary) (13.6)

coastal straightening

the tendency for an irregular coast to be straightened over time by coastal erosion processes (13.3)

cobble

sediment particle that is between 64 and 256 mm in diameter (12.1)

coccolithophore

photosynthetic algae that makes its test (shell) out of calcium carbonate (7.2)

compensation depth

the depth where the rate of photosynthesis equals the rate of respiration (7.3)

conduction

the transfer of heat through direct contact (8.1)

conservative ions

ions whose proportions are the same regardless of overall salinity; the major ions in seawater (5.3)

constructive interference

where the interaction of multiple waves creates waves larger than any of the component waves (10.2)

continental crust

the Earth's crust underlying the continents (as opposed to ocean crust) (3.2)

continental drift

the idea that the continents have moved over the surface of the Earth over geological time (4.1)

continental margin

the region of transition from the land to the deep sea floor, i.e. between continental and oceanic crust (1.2)

continental rise

the area at the bottom of the continental slope, where it transitions to the abyssal sea floor (1.2)

continental shelf

the shallow (typically less than 200 m) and flat sub-marine extension of a continent (1.2)

continental slope

the steeper part of a continental margin, that slopes down from a continental shelf towards the abyssal plain (1.2)

convection cell

a rotating region in a fluid in which upward motion of warmer, low density fluid in the center is balanced by downward motion of cooler, denser fluid at the periphery (4.3)

convergent boundary

a plate boundary at which the two plates are moving towards each other (4.6)

core

the metallic interior part of the Earth, extending from a depth of 2900 km to the center (3.2)

core-mantle boundary

the boundary, at 2900 km depth, between the mantle and the core (3.2)

Coriolis Effect

the tendency for the path of moving bodies (e.g., ocean currents) to be deflected on the surface of the Earth, to the right in the Northern Hemisphere and to the left in the Southern Hemisphere (8.2)

cosmogenous sediment

sediment derived from extraterrestrial sources (12.5)

crest

the highest point on a wave (10.1)

Cretaceous

a geologic period that spans 79 million years from the end of the Jurassic Period 145 million years ago to the beginning of the Paleogene Period 66 mya

crust

the uppermost layer of the Earth, ranging in thickness from about 5 km (in the oceans) to over 50 km (on the continents) (3.2)

deep water wave

a wave above a water depth greater than half of its wavelength (10.1)

delta

large, often triangular accumulation of sediment near the mouth of a river (13.4)

denitrification

where nitrate is converted to molecular nitrogen through a series of intermediate nitrogen oxide products (5.6)

density

mass per unit volume of a substance (e.g., g/cubic cm) (6.3)

destructive interference

where the interaction of multiple waves creates waves smaller than any of the component waves (10.2)

diatom

photosynthetic algae that make their tests (shells) from silica (7.2)

diatomaceous earth

powdery sediment composed of silica diatom tests (12.3)

differentiation

the un-mixing of a magma, typically by the physical separation of minerals that crystallize early and settle towards the bottom (3.1)

dinoflagellate

photosynthetic algae characterized by the presence of flagella and a cellulose test (shell) (7.2)

discoaster

an extinct form of single-celled algae that produced calcareous tests that can still be found in some marine sediments (12.3)

diurnal tide

a tidal cycle with only one high and one low tide per day (11.3)

divergent boundary

a plate boundary at which the two plates are moving away from each other (4.5)

doldrums

areas of low pressure and weak winds along the equator (8.2)

domed iceberg

iceberg with a rounded top (14.2)

downwelling

process by which surface water is forced downwards (9.5)

drowned river valley estuary

an estuary formed when sea level rises and submerges a river valley (also known as a coastal plain estuary) (13.6)

drydock iceberg

iceberg with a water-covered channel running through it (14.2)

dysphotic zone

depths of the water column where there is some light penetration, but not enough to support photosynthesis; corresponds to the mesopelagic zone, 200-1000 m. Also known as the twilight zone (1.3)

ebb current

current created by an outgoing tide (11.3)

eccentricity

in the context of Milankovitch Cycles, the degree to which the Sun is offset from the geometric center of the Earth's orbit (8.5)

eddy

a rotating water mass (9.2)

Ekman spiral

where each layer of water is deflected relative to the layer above it, forming a spiral that extends down to about 100 m (9.3)

Ekman transport

bulk transport of water due to the Ekman spiral; the net movement Ekman transport is 90 degrees relative to the wind direction (9.3)

El Niño

a periodic climatic situation in which warm water extends all or most of the way to the eastern edge of the equatorial Pacific (9.6)

El Niño-Southern Oscillation (ENSO)

the fluctuating atmospheric conditions that lead to the localized ocean warming of El Niño (9.6)

electron

a sub-atomic particle of essentially no mass and a single negative charge (5.1)

Eocene

a geological epoch lasting from 56 to 33.9 million years ago

epicenter

the location on the surface vertically above the location (i.e., "hypocenter" or "focus") where an earthquake takes place (4.8)

epipelagic zone

the upper layer of water (0 to 200 m) in areas of the open ocean (1.3)

estuary

a partially enclosed body of water where seawater is diluted by freshwater input (13.6)

euphotic zone

the upper regions of the ocean where there is enough light to support photosynthesis; approximately 0-200 m; also called the photic zone (1.2)

eustatic sea level change

sea level change related to a change in the volume of the oceans, typically because of an increase or decrease in the amount of glacial ice on land (13.7)

evaporites

hydrogenous sediments that form when seawater evaporates (12.4)

fast ice

ice sheets that are attached to land (14.1)

fault

a boundary in rock or sediment along which displacement has taken place (4.7)

fecal express

small particles reach the seafloor much faster when incorporated into large fecal pellets than if they sank on their own (12.3)

feedback

a process by which the physical effects of a climate forcing can have other effects (either negative or positive) on the climate (8.5)

Ferrel Cell

the atmospheric convection cells between 30 and 60 degrees latitude (8.2)

fetch

the distance over which wind blows to form waves (10.2)

firn

the granular transitional state between snow and ice within a glacier (14.2)

fjord

a deep, U-shaped estuary that was carved out by advancing glaciers (13.6)

flood current

current created by an incoming tide (11.3)

flushing time

the time it would take for all of the fresh water in an estuary to be replaced by runoff of new water (13.6)

focus (earthquake)

the actual point below surface at which an earthquake takes place (equivalent to hypocenter) (4.6)

foraminifera

a single-celled protist with a shell that is typically made of calcium carbonate (12.3)

foreshore

the part of a beach between the high tide and low tide lines (13.1)

frazil

small, needle-like crystals in the first stages of sea ice formation (14.1)

frequency

the number of waves that pass a point in a given amount of time (10.1)

fringing reef

a reef adjacent to a shoreline where there is either a very narrow back reef area or none at all (in which case the reef is effectively attached to the shore) (4.10)

frost line

in the context of planetary systems the boundary beyond which volatile components (e.g., water, carbon dioxide, methane, ammonia etc.) are frozen (3.1)

Ga

(gigaannus) billions of years before the present

galaxy

a gravitationally-bound system of stars and interstellar matter (3.1)

gas giant

a large planet composed mostly of hydrogen and helium (e.g. Jupiter) (3.1)

geostrophic flow

circular currents created from the balance between gravity- and Ekman-driven flow (9.3)

giant impact hypothesis

the theory that the Moon formed when a Mars-sized planet (Theia) collided with the Earth at 4.5 billion years ago (3.1)

glacial groove

scratches and grooves carved into bedrock from rocks carried by moving glaciers (4.1)

glacial ice

ice formed from the accumulation and compression of snow into glaciers (14.1)

glacial period

a period of Earth's history during which glacial ice was present over a sufficient extent to have left recognizable evidence (4.1)

glacier

a long lasting (centuries or more) body of ice on land that moves under its own weight (4.1)

granite

an igneous (formed from cooling magma) rock that comprises much of the continental crust (3.2)

granule

a sedimentary particle ranging in size from 2 to 4 mm in diameter (12.1)

grease ice

an accumulation of frazil to create a slushy consistency in sea ice formation (14.1)

greenhouse effect

in the context of climate, the ability of an atmosphere to absorb infrared radiation due to the presence of greenhouse gases (8.1)

greenhouse gas

a gaseous molecule with 3 or more atoms that is able to absorb infrared radiation (8.1)

groin (groyne)

a man-made structure extending from the shore built to deflect the energy of waves (13.5)

groin field

a series of groins along a beach (13.5)

gross primary production

the total amount of organic material created by primary producers (7.1)

groundwater

water that lies beneath the surface of the ground (5.2)

Gulf Stream

the major surface current flowing northwards along the Atlantic coast of the U.S. and Canada (9.2)

guyot

a flat-topped seamount (also called a tablemount) (4.9)

gyre

a large circular ocean surface current (9.1)

hadal zone

the region of the seafloor below 6000 m (1.3)

Hadley Cell

the atmospheric convection cells between the equator and 30 degrees latitude (8.2)

hadopelagic (hadalpelagic) zone

region of the open ocean with water depths greater than 6000 m (1.3)

halite

NaCl, a mineral also known as table salt (12.4)

halocline

where there is a dramatic change in salinity over a small change in depth (5.3)

hard stabilization

the building of physical structures to prevent the erosion of beaches and shorelines (13.5)

harmful algal bloom (HAB)

when phytoplankton appear in very high concentrations with potentially hazardous consequences such as mass die-offs or toxicity (7.2)

headland

a point of land extending out to sea (13.3)

heat budget

the balance between the amount of heat entering and leaving the Earth (8.1)

heat capacity

the amount of heat needed to change a substance's temperature by one degree (5.1)

Henry's Law

as the pressure increases, a fluid will contain more dissolved gas (6.1)

high pressure

in atmospheric terms, a region of descending air, increasing the atmospheric pressure. Winds blow away from high pressure zones (8.3)

highly stratified estuary

a deep estuary with some mixing at the surface, but little mixing at depth (13.6)

homolosine projection

a map projection where area is retained, but there are interruptions to the continents or oceans (2.4)

horse latitudes

areas of high pressure and weak winds around 30 degrees latitude in both hemispheres (8.2)

hot spot

the surface area of volcanism and high heat flow above a mantle plume (4.9)

hydrogen bond

a weak bond between two molecules due to the electrostatic attraction of a proton in one molecule to the negative polar end of the other molecule (5.1)

hydrogenous sediment

sediments formed from the precipitation of dissolved substances (12.4)

hydrological cycle

the cycling of water through the ocean, atmosphere, lakes, organisms, and other reservoirs (5.2)

hydrothermal vent

area of the seafloor where superheated water seeps out of the crust (4.11)

hypocenter

the actual point below surface at which an earthquake takes place (equivalent to focus) (4.8)

hypoxia

a condition with low dissolved oxygen, usually defined as oxygen levels below 2 mg/L (5.4)

ice field

an area covered by ice floes (14.1)

ice floe

a relatively large piece of floating sea ice (14.1)

ice giant

a planet that is comprised mainly of gases heavier than hydrogen and helium, including oxygen, carbon, nitrogen, and sulfur (e.g., Uranus and Neptune) (3.1)

iceberg

a large, floating piece of glacial ice (14.2)

inner core

the solid metal mass at the center of the Earth, extending 1200 km from the center (3.2)

insolation

a measure of the intensity of solar energy at a specific location or time (expressed in W/square m) (8.5)

Intergovernmental Panel on Climate Change (IPCC)

an international body established in 1988 by the UN's World Meteorological Organization and the UN Environment Program to prepare periodic reports on the status of global climate change and its mitigation (8.5)

intermediate wave

a wave in a water depth between 1/2 and 1/20 its wavelength (10.1)

internal wave

waves that form below the surface at the interface between water masses of different densities (10.1)

intertidal zone

the region of a coast between the high and low tide lines. Also called the littoral zone (1.3)

Intertropical Convergence Zone

the area near the equator where the northeast and southeast trade winds converge; known for weak winds, it is also called the doldrums (8.2)

ion

an atom or molecule that has either gained or lost electrons and has thus become charged (5.1)

island arc

long chains of volcanic islands found along convergent tectonic plate boundaries (4.6)

isostasy

the equilibrium position reached by a block of crust floating on the underlying fluid mantle (3.2)

isostatic sea level change

the effect on relative sea level of a vertical movement of the crust resulting from a change in the mass of the crust (e.g., from losing or gaining ice) (13.7)

isotope

forms of the same element that contain equal numbers of protons but different numbers of neutrons in their nuclei

jetty

a long structure built to protect a harbor from filling with sand due to longshore transport (13.5)

Jovian planet

a gas giant (3.1)

ka

(kiloannus) thousands of years before the present

kinetic energy

the energy that an object possesses due to its motion (5.1)

knot

one knot (kt) = 1 nautical mile per hour = 1.15 mph = 1.85 kph

La Niña

a periodic climatic situation in which colder than normal water extends throughout the equatorial Pacific (9.6)

land breeze

winds blowing from land towards the ocean (8.3)

Langmuir circulation

corkscrew circulation patterns formed parallel to strong winds; they only extend a few meters below the surface (9.7)

latent heat of fusion

the heat required to change a substance from solid to liquid; 80 cal/g in the case of ice melting to water (5.1)

latent heat of vaporization

the heat required change a substance from liquid to gas; 540 cal/g to turn water into vapor (5.1)

latitude

the distance north or south of the equator, measured as an angle from the equator (2.1)

Laurentide Ice Sheet

the continental glacier that extended across central eastern North America during the Pleistocene, covering most of Canada and a significant part of the United States (3.2)

lithification

the conversion of unconsolidated sediments into rock by compaction and cementation (12.1)

lithogenous sediment

sediment derived from preexisting rock (12.2)

lithosphere

the rigid outer part of the Earth, including the crust and the mantle down to a depth of about 100 km (3.2)

littoral drift

the movement of sediment along a shoreline resulting from a longshore current and also from the swash and backwash on a beach face (another name for longshore transport) (13.2)

littoral zone

the region of a coast between the high and low tide lines. Also called the intertidal zone (1.3)

longitude

measurement of distance east or west of the prime meridian, expressed as an angle (2.1)

longshore bar

an offshore deposit of sand parallel to the shoreline (13.1)

longshore current

the movement of water parallel to a shoreline produced by the approach of waves at an angle to the shore (13.2)

longshore transport

the movement of sediment along a shoreline resulting from a longshore current and also from the swash and backwash on a beach face. Also known as littoral drift (13.2)

low pressure

in terms of the atmosphere, a region of rising air, lowering the atmospheric pressure. Winds blow towards low pressure regions, which are often characterized by precipitation from rising, cooling, condensing air (8.3)

low tide terrace

another name for the beach face (13.1)

lysocline

the depths where the rate of calcium carbonate dissolution increases dramatically over surface waters (12.6)

Ma

(Megaannus) millions of years before the present

magma

molten rock typically dominated by silica (3.2)

magnetic dip

the angle of the magnetic field within a rock, relative to the horizontal; may be used to infer the latitude where the rock was first formed (4.2)

magnitude

a measure of the amount of energy released by an earthquake (4.8)

major ions

the six ions that comprise over 99% of the ions in the ocean (chloride, sodium, sulfate, magnesium, calcium, potassium) (5.3)

manganese nodule

spherical accumulations of manganese and other metals that form slowly through precipitation on the seafloor (12.4)

mantle

the middle layer of the Earth, dominated by iron and magnesium rich silicate minerals and extending for about 2900 km from the base of the crust to the top of the core (3.2)

mantle convection

movements in the mantle from rising and sinking mantle material as it heats and cools (4.3)

mantle plume

a plume of hot rock (not magma) that rises through the mantle (either from the base or from part way up) and reaches the surface where it spreads out and also leads to hot-spot volcanism (4.9)

maturity

how long sediment particles have been transported by water or other vectors (12.1)

meander

the sinuous path taken by a current, such as the Gulf Stream (9.2)

Mercator projection

a map projection where latitude and longitude are both represented as straight, parallel lines intersecting at right angles (2.3)

mesopelagic zone

the upper middle zone of the open ocean extending from 200 to 1000 m depth (1.3)

Mesozoic

the geological era from about 252 to 66 million years ago

meteoroid

a fragment of either stony or metallic debris in space (12.5)

methane hydrate

a combination of water ice and methane in which the methane is trapped inside “cages” in the ice (12.4)

mid-ocean ridge

an underwater mountain system along divergent plate boundaries, formed by plate tectonics (4.5)

Milankovitch cycles

millennial-scale variations in the orbital and rotational parameters of the Earth that have subtle effects on the Earth’s climate (8.5)

mixed interference

where the interaction of multiple waves creates both constructive and destructive interference and an irregular surface pattern (10.2)

mixed layer

the topmost layer of the ocean, where winds, waves, and currents mix the water so that conditions are relatively constant; approximately the top 100 m (5.3)

mixed semidiurnal tide

a tidal cycle with two high and two low tides per day, each of different heights (11.3)

Mohorovičić discontinuity (Moho)

the boundary between the crust and the mantle (3.3)

nautical mile

a distance equal to one minute of latitude; equivalent to 1.15 land miles or 1.85 km (2.1)

neap tide

the period of minimum tidal range when the Earth is perpendicular to the sun and moon (11.1)

nearshore

the part of a beach from the low tide line to the depth where wave action is no longer influenced by the bottom, i.e. to where the depth exceeds the wave base (13.1)

nebula

a cloud of interstellar dust and gases (3.1)

negative feedback

a process that results in a decrease in that process (in the context of climate change it is a process that reduces the change in climate, such as the enhanced growth of vegetation in response to an increase in atmospheric carbon dioxide) (8.5)

neritic

the marine pelagic province from the low tide line to the shelf break (1.3)

net production

total primary production minus the organic compounds used up by respiration by the producers (7.1)

new production

primary production supported by nutrients brought in from outside of the local ecosystem (7.1)

nilas

a thin surface sheet of sea ice (14.1)

nitrification

the biological oxidation of ammonia or ammonium to nitrite followed by the conversion of the nitrite to nitrate (5.6)

nitrogen fixation

the process of converting atmospheric nitrogen gas into nitrogenous compounds like ammonia (5.6)

non-conservative ions

ions in seawater whose proportions fluctuate with changes in salinity (5.3)

non-tabular iceberg

an iceberg with any shape other than tabular (14.2)

North Atlantic Deep Water

deep Atlantic Ocean water that has descended in the far north of the basin in the area between Scandinavia and Greenland (9.8)

nutrient

in the context of primary production, substances required by photosynthetic organisms to undergo growth and reproduction (5.6)

nutrient-like element

elements that have a similar vertical profile to nutrients; low amounts at the surface, increased abundance at depth (5.7)

obliquity

in the context of Milankovitch Cycles, the angle of the tilt of the Earth's rotational axis with respect to the plane of its orbit around the Sun (8.5)

ocean acidification

where the overall pH of the ocean declines, likely due to an increased amount of carbon dioxide in the ocean (5.5)

oceanic

the marine pelagic province representing the open ocean regions, i.e. beyond the neritic zone (1.3)

oceanic crust

the Earth's crust underlying the oceans (as opposed to continental crust) (3.2)

offshore

the beach zone beyond the nearshore region (13.1)

oolite

a small (approximately 1 mm) sphere of calcite formed in areas of tropical shallow marine water (12.4)

ooze

a sediment composed of >30% biogenous material (12.3)

outer core

the layer of the inner Earth extending 2300 km from the top of the inner core to the bottom of the mantle, composed of fluid metal alloys (3.2)

outgassing

where dissolved substances in magmas are released as gases when the pressure is reduced (5.2)

overturning

the vertical cycling within a body of water, where denser water sinks and less dense water floats to the surface (5.1)

oxygen minimum layer

region of ocean depths where dissolved oxygen is at its lowest level; usually around 1000 m for the open ocean (5.4)

p-wave

a seismic body wave that is characterized by deformation of the rock in the same direction that the wave is propagating (compressional vibration) (3.3)

pack ice

free-floating ice floes (14.1)

Paleocene

a geological epoch that lasted from about 66 to 56 million years ago

paleomagnetic

past variations in the intensity and polarity of the Earth's magnetic field (4.2)

Paleozoic

the geologic era lasting from 541 to 252 million years ago

pancake ice

small, rounded, thin pieces of sea ice that will freeze together to form an ice floe (14.1)

Pangaea

the supercontinent that existed between approximately 300 and 180 Ma; it contained all of the modern continents combined into a single land mass (4.1)

partially mixed estuary

where salinity increases from the head to the mouth, but there is also a slight increase in salinity with depth at any point; also called a slightly stratified estuary (13.6)

passive continental margin

a boundary between a continent and an ocean at which there is no tectonic activity (e.g., the eastern edge of North America) (1.2)

pebble

a sedimentary particle ranging in size from 2 to 64 mm (includes granule) (12.1)

pelagic

relating to the open ocean (1.3)

period

the time it takes for a complete wave to pass a given point (10.1)

Permian

a geologic period which spans 47 million years from the end of the Carboniferous Period 298.9 million years ago, to the beginning of the Triassic period 251.902 Mya

Phanerozoic

the eon on the geological time scale covering time from the beginning of the Cambrian period 541 million years ago to the present, and comprising the Paleozoic, Mesozoic, and Cenozoic eras

phase change

the change of state between a solid, liquid, or gas (8.1)

photic zone

the upper regions of the ocean where there is enough light to support photosynthesis; approximately 0-200 m; also called the euphotic zone (1.2)

photosynthesis

the production of organic compounds from carbon dioxide and water, using sunlight as an energy source (5.5)

physiographic projection

map projection presenting bathymetry or altitude data as a 3D relief map (2.3)

phytoplankton

drifting, usually single-celled algae that undergo photosynthesis (7.1)

picoplankton

planktonic bacteria (7.2)

pinnacled iceberg

an iceberg with one or more tall spires (14.2)

planisphere projection

map projection that keeps latitude horizontal, but shows some convergence of longitude (2.3)

plankton

an organism that cannot swim effectively, so it drifts with the currents (7.1)

plate

a region of the lithosphere that is considered to be moving across the surface of the Earth as a single unit (4.1)

plate tectonics

the concept that the Earth's crust and upper mantle (lithosphere) is divided into a number of plates that move independently on the surface and interact with each other at their boundaries (4.1)

plunging breaker

a breaking wave on moderately-steep beaches that curls over on itself as it breaks (10.3)

Polar Cell

the atmospheric convection cells between 60 degrees latitude and the pole (8.2)

polar easterlies

the dominant wind bands between the poles and 60 degrees latitude (8.2)

polar front

the boundary between the polar cell and the Ferrel cell around the 60 latitude in each hemisphere (8.2)

polar molecule

a molecule where the electrons are not distributed equally, leading to a charge imbalance across the molecule; portions of the molecule are slightly positive while other portions are slightly negative (5.1)

polar wandering path

a path of varying magnetic pole positions defined by paleomagnetic data (in fact it is now understood that the continents have wandered, not the poles, so a more appropriate terms is "apparent polar wandering path") (4.2)

polyna

an area of persistent open water in areas otherwise covered with ice (14.1)

positive feedback

a process that results in an increase in that process (in the context of climate change it is a process that enhances the change in climate, such as the reduced reflectivity of the Earth's surface when ice melts) (8.5)

ppm

parts per million

ppt

parts per thousand

practical salinity unit (PSU)

a unitless measure of salinity equal to parts per thousand (5.3)

precession

in the context of Milankovitch Cycles, the variation in the direction at which the Earth's rotational axis is pointing (8.5)

pressure ridge

jagged ridges created from colliding and buckling ice floes (14.1)

primary production

the synthesis of organic compounds from aqueous carbon dioxide by plants, algae, and bacteria (7.1)

protoplanetary disk

a rotating cloud of gas and dust surrounding a young star (3.1)

pycnocline

a region in the water column where there is a large change in density over a small change in depth (6.3)

quartz

a mineral composed of silicon and oxygen atoms in the ratio of 1 Si:2 O; one of the most abundant minerals in the Earth's surface (12.1)

radiation

the emission of energy in the form of electromagnetic waves (8.1)

radiolaria

microscopic (0.1 to 0.2 mm) marine protozoa that produce silica shells (12.3)

rain shadow

arid conditions behind a mountain range, as rising air on the other side of the mountain caused rain, leaving only dry air to descend back down the mountain (8.3)

reef

a mound of carbonate formed in shallow tropical marine environments by corals, algae and a wide range of other organisms (4.10)

regenerated production

primary production resulting from the recycling of nutrients within an ecosystem (7.1)

remnant magnetism

magnetism of a body of rock that formed at the time the rock formed and is consistent with the magnetic field orientation that existed at that time and place (4.2)

residence time

the average amount of time an element will remain in the ocean before being removed (5.2)

restoring force

the force that opposes a wave-generating force and attempts to return the sea surface to the still water level (10.2)

ridge push

the concept that at least part of the mechanism of plate motion is the push of oceanic lithosphere down from a ridge area (4.3)

rift valley

a valley created when crust subsides along a divergent plate boundary (4.5)

rip current

a strong flow of water outward from a beach (13.2)

rogue wave

an exceptionally large wave arising among a series of smaller waves (10.2)

rule of constant proportions

the major ions in seawater are always found in the same proportions, regardless of overall salinity (5.3)

runoff

flow of water down a slope, either across the ground surface, or within a series of channels (12.2)

s-wave

a seismic body wave that is characterized by deformation of the rock perpendicular to the direction that the wave is propagating (3.3)

salinity

the concentration of dissolved ions in water (5.3)

salt wedge estuary

an estuary with mostly fresh surface water, and a wedge of seawater intruding along the bottom (13.6)

sand

a mineral or rock fragment ranging in size from 1/16th to 2 mm (12.1)

saturation

the amount of a substance currently dissolved in the water, relative to the maximum possible content (5.4)

scarp

a short, steep wall carved out by wave action between the foreshore and the berm of a beach (13.1)

scavenged element

elements whose vertical profiles show high abundance at the surface and declining concentrations at depth as they are removed by sinking particles (5.7)

sea breeze

winds blowing from the ocean towards the land (8.3)

sea cave

a shallow cave formed on a rocky shore by wave erosion (13.3)

sea cliff

a coastal escarpment that is typically eroding inland as a result of wave action (13.3)

sea ice

ice formed from the freezing of seawater (14.1)

sea stack

a prominent rocky island that is a remnant of the erosion of a headland (13.3)

sea state

describes the current wave conditions in an area (10.2)

seafloor spreading

the formation of new oceanic crust by volcanism at a divergent plate boundary (4.5)

seamount

a submerged mountain rising from the seafloor (4.9)

seawall

a wall built against a sea cliff or dune to prevent erosion from wave action (13.5)

sediments

unconsolidated particles of mineral or rock that settle to the seafloor (12.1)

seismic

pertaining to earthquakes (3.3)

seismology

the study of vibrations within the Earth (3.3)

semidiurnal tide

a tidal cycle with two high and two low tides per day, each of roughly equal heights (11.3)

shallow water wave

a wave in water with a depth less than 1/20 of the wavelength (10.1)

shelf break

the boundary between the continental shelf and continental slope, where the angle of the seafloor begins to get steeper (1.2)

significant wave height (Hs)

the mean height of the largest one-third of the waves in a wave spectrum (10.2)

silica

a form of the mineral quartz (7.2)

siliceous sediment

sediment dominated by particles of silica, often from the shells of marine organisms (7.2)

silt

sedimentary particles ranging is size from 1/256th to 1/16th of a mm (12.1)

slab pull

the concept that at least part of the mechanism of plate motion is the pull of oceanic lithosphere down into the mantle (4.3)

slack tide

period of little water movement between an incoming and outgoing tide (11.3)

slightly stratified estuary

where salinity increases from the head to the mouth, but there is also a slight increase in salinity with depth at any point; also called a partially mixed estuary (13.6)

snow line

in astronomy the radius around a star at which represents the boundary between gases (or liquids) and solids (3.1)

SOFAR channel

range of depths around 1000 m where sound travels the slowest, so sound waves are refracted back into the channel and can be propagated long distances (6.4)

solar system

a star and the planets surrounding it (3.1)

solar wind

a stream of ionized (charged) particles away from the Sun (3.1)

solubility

the amount of a dissolved substance that water can hold under a particular set of conditions, which are usually defined as 0 degrees C and 1 atmosphere of pressure (5.4)

sonar

acronym for sound navigation and ranging; a method of using sound echoes to detect objects (1.4)

sorting

how uniform the particles of a sediment are in terms of size (12.1)

sounding

a single measurement of ocean depth (1.4)

specific heat

the heat required to raise the temperature of 1 g of a substance by 1 degree C (5.1)

spherule

a microscopic piece of space dust (12.5)

spilling breaker

a breaker on relatively flat beaches that slowly increases its height and collapses (10.3)

spit

a sand or coarser deposit extending from shore out into open water (13.4)

splash wave

a wave formed when something falls into the ocean and creates a splash (10.1)

spring tide

the period of maximum tidal range when the moon, sun and Earth are aligned (11.1)

stable gas

unreactive gases that dissolve in seawater (5.7)

steady state

where a system shows no net change, as input equals output (5.2)

still water level

where the water surface would be if there were no waves present and the sea was completely calm (10.1)

storm surge

an area of high water that moves with storm systems (8.4)

subducted

when part of a plate is forced beneath another plate along a subduction zone (4.3)

subduction zone

the sloping region along which a tectonic plate descends into the mantle beneath another plate (4.6)

subittoral zone

the region of a coast from the low tide line to the end of the continental shelf (1.3)

supralittoral zone

the region of a coast above the high tide line (1.3)

surf beat

an irregular surface wave pattern caused by mixed interference (10.2)

surf zone

the near-shore zone where waves are breaking into surf (13.1)

surface tension

where a cohesive layer forms on the water surface due to attraction between water molecules (5.1)

surging breaker

waves that break on steep beaches with a very sudden increase in height and sudden collapse right on the beach (10.3)

swash

the upward motion of a wave on a beach (typically takes place at the same angle that the waves are approaching the shore) (13.1)

swell

regular, long-period waves that have sorted themselves based on speed (10.2)

tablemount

a flat-topped seamount (also called a guyot) (4.9)

tabular iceberg

flat-topped, steep-sided iceberg with a length greater than five times the height (14.2)

tectonic estuary

an estuary formed from flooding following the tectonic subsidence of land (13.6)

tectonic plate

a region of the lithosphere that is considered to be moving across the surface of the Earth as a single unit (4.4)

tektite

solidified glass fragments ejected during meteorite impacts (12.5)

terrestrial planet

a planet with a rocky mantle and crust and metallic core (e.g., Earth) (3.1)

terrigenous sediment

referring to sedimentary particles that originated on a continent (12.2)

test

the shell-like hard parts (either silica or carbonate) of small organisms such as radiolarians and foraminifera (12.3)

thermal expansion

the increase in the volume of a body water as its temperature rises and its density decreases (13.7)

thermocline

a region in the water column where there is a dramatic change in temperature over a small change in depth (6.2)

thermohaline circulation

deep ocean circulation driven by differences in water density (9.8)

tidal bore

a wave that moves up a river with an incoming tide (11.3)

tidal day

the amount of time between a tide on one day and the same tide the following day (11.2)

tidal range

the difference in height between the high and low tides (11.1)

tidal volume / tidal prism

the volume difference of an area between low and high tides (11.3)

tombolo

a sand or coarser deposit connecting an island or rocky prominence to a larger body of land (13.4)

trade winds

prevailing wind bands between the equator and 30 degrees latitude (8.2)

transform boundary

a boundary between two plates that are moving horizontally with respect to each other (4.5)

transform fault

a type of fault in which two pieces of crust slide past one another (4.5)

trough

the lowest point of a wave (10.1)

tsunami

a long-wavelength wave produced by the vertical motion of the floor of the ocean, typically related either to an earthquake or other submarine seismic event (10.1)

turbidity current

a current moving down downhill along the bottom, driven by the weight of the sediment within it (1.2)

twilight zone

depths of the water column where there is some light penetration, but not enough to support photosynthesis; corresponds to the mesopelagic zone, 200-1000 m. Also known as the dysphotic zone (1.3)

upwelling

process by which deeper water is brought to the surface (9.5)

vertically-mixed estuary

estuary with complete mixing of fresh and salt water, where salinity is constant at all depths in a particular location but increases towards the estuary mouth; also called a well-mixed estuary (13.6)

water mass

a volume of seawater with a distinctive density as a result of its unique profile of temperature and salinity (9.8)

wave base

the depth of water that is affected by the sub-surface orbital motion of wave action (approximately one-half of the wavelength) (10.1)

wave height

the distance between the crest and trough of a wave (10.1)

wave steepness

the ratio of wave height to wavelength (10.1)

wave-cut platform/terrace

a nearly-horizontal bench of rock eroded by waves within the surf zone (13.3)

wavelength

the distance between the crests of two waves (10.1)

weathering

a range of processes taking place in the surface environment, through which solid rock is transformed into sediment and ions in solution (12.2)

wedge iceberg

iceberg with a steep face next to a more gradually sloping side (14.2)

well-mixed estuary

estuary with complete mixing of fresh and salt water, where salinity is constant at all depths in a particular location but increases towards the estuary mouth; also called a vertically-mixed estuary (13.6)

westerlies

the dominant wind bands between 30 and 60 degrees latitude in each hemisphere (8.2)

western intensification

currents on the western side of a gyre are faster, deeper, and narrower than currents on the eastern side (9.4)

zooplankton

small, drifting carnivorous organisms (7.1)

About the Author

Paul Webb earned a BA in Biology from the University of Richmond, and a MS in Marine Science and PhD in Biology from the University of California-Santa Cruz. At UCSC his research focused on the behavioral and physiological ecology of northern elephant seals. For the past 20 years he has been a faculty member in the Department of Biology, Marine Biology, and Environmental Science at Roger Williams University in Bristol, Rhode Island, where he is a Professor of Marine Biology. At RWU he regularly teaches courses in Oceanography, Marine Mammalogy, Marine Vertebrate Zoology, Animal Behavior, Animal Physiology, Tropical Ecology (in Belize), and Herpetology.