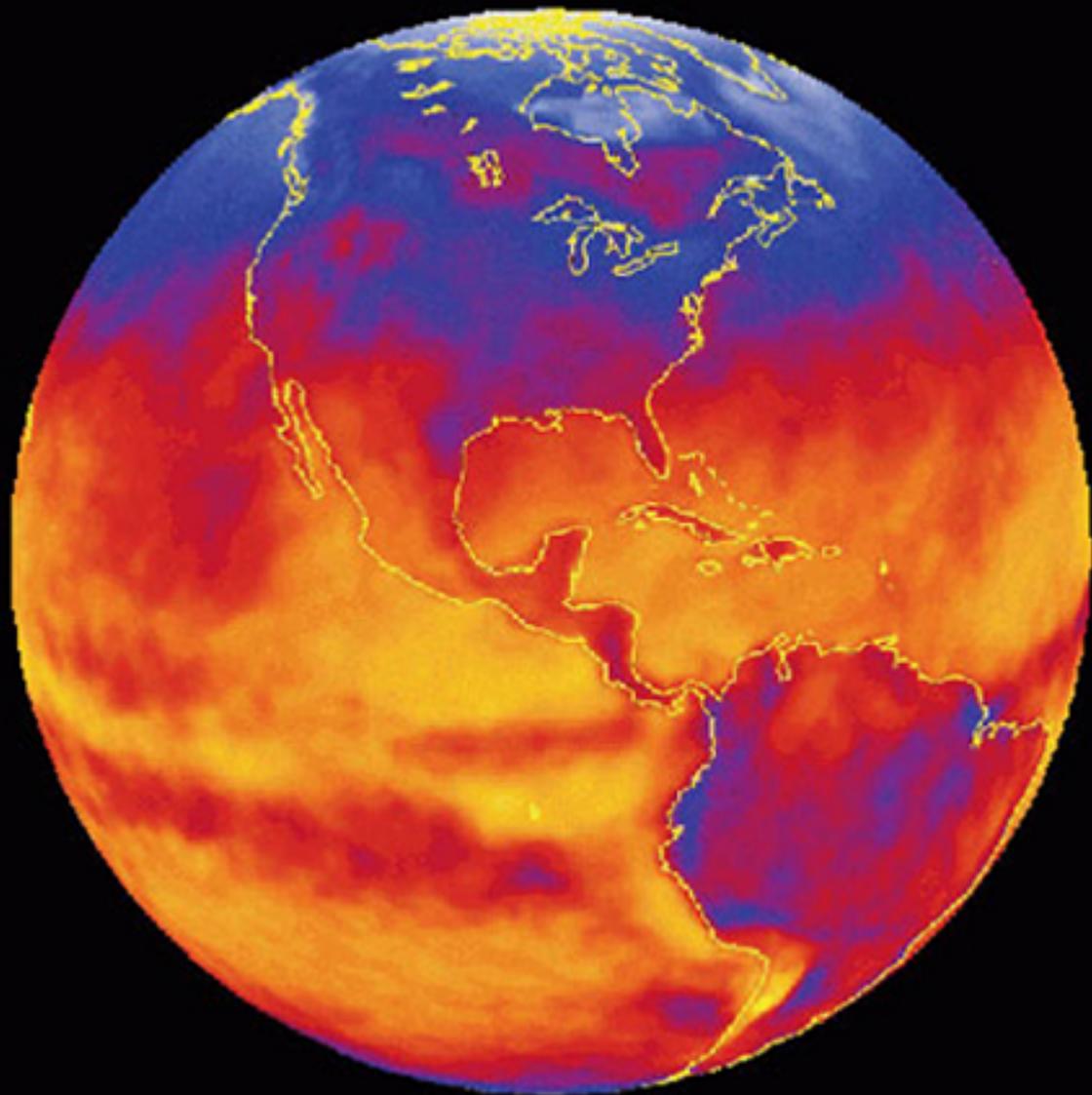


# Earth Science

(Components, Branches & Elements)



Leah Devito

First Edition, 2012

ISBN 978-81-323-1937-5

© All rights reserved.

*Published by:*

**Learning Press**

4735/22 Prakashdeep Bldg,

Ansari Road, Darya Ganj,

Delhi - 110002

Email: [info@wtbooks.com](mailto:info@wtbooks.com)

# Table of Contents

Chapter 1 - Introduction to Earth Science

Chapter 2 - Branches & Related Fields of Earth Science

Chapter 3 - Paleoclimatology

Chapter 4 - Biosphere & its Components

Chapter 5 - Earth Systems & Sciences

Chapter 6 - Earth's Interior & Related Theories

## Chapter 1

# Introduction to Earth Science

**Earth science** (also known as **geoscience**, **the geosciences** or **the Earth sciences**), is an all-embracing term for the sciences related to the planet Earth. It is arguably a special case in planetary science, the Earth being the only known life-bearing planet. There are both reductionist and holistic approaches to Earth sciences. The formal discipline of Earth sciences may include the study of the atmosphere, oceans and biosphere, as well as the solid earth. Typically Earth scientists will use tools from physics, chemistry, biology, chronology and mathematics to build a quantitative understanding of how the Earth system works, and how it evolved to its current state.



A volcano eruption is the release of stored energy from below the surface of Earth, originating from radioactive decay and gravitational sorting in the Earth's core and mantle, and residual energy gained during the Earth's formation.

## Earth's interior

Plate tectonics, mountain ranges, volcanoes, and earthquakes are geological phenomena that can be explained in terms of energy transformations in the Earth's crust.

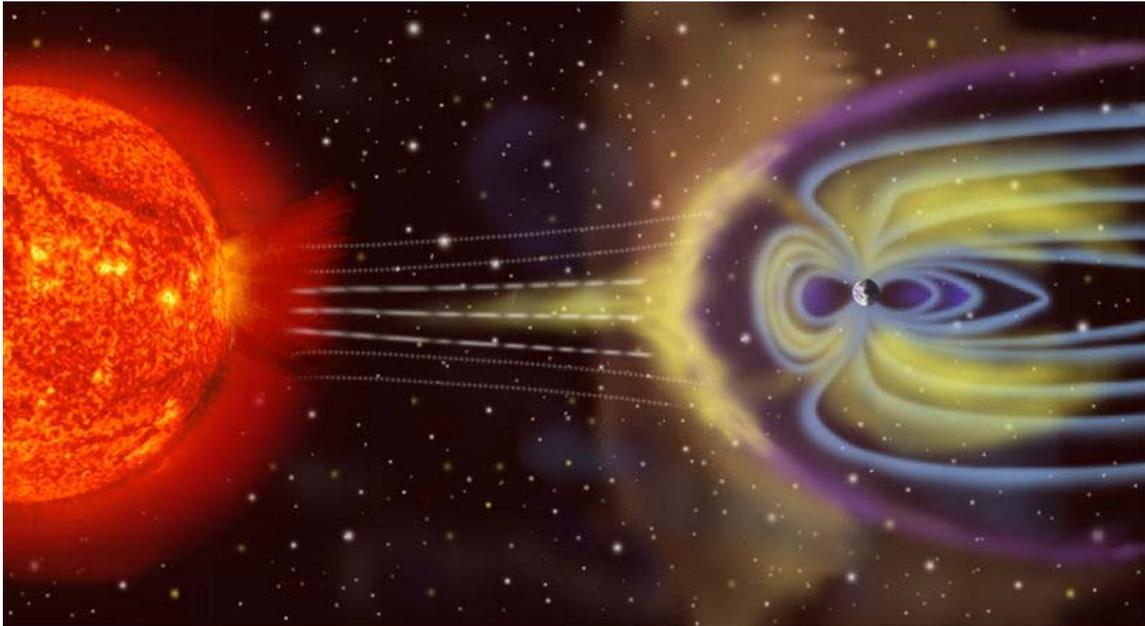
Beneath the Earth's crust lies the mantle which is heated by the radioactive decay of heavy elements. The mantle is not quite solid and consists of magma which is in a state of semi-perpetual convection. This convection process causes the lithospheric plates to move, albeit slowly. The resulting process is known as plate tectonics.

Plate tectonics might be thought of as the process by which the earth is resurfaced. Through a process called *spreading ridges* (or seafloor spreading), new earth crust is created by the flow of magma from underneath the lithosphere to the surface, through fissures, where it cools and solidifies. Through a process called subduction, crust is pushed underground—beneath the rest of the lithosphere—where it comes into contact with magma and melts—rejoining the mantle from which it originally came.

Areas of the crust where new crust is created are called *divergent boundaries*, and areas of the crust where it is brought back into the earth are called *convergent boundaries*. Earthquakes result from the movement of the lithospheric plates, and they often occur near convergent boundaries where parts of the crust are forced into the earth as part of subduction.

Volcanoes result primarily from the melting of subducted crust material. Crust material that is forced into the Asthenosphere melts, and some portion of the melted material becomes light enough to rise to the surface—giving birth to volcanoes.

# Atmosphere



The magnetosphere shields the surface of Earth from the charged particles of the solar wind. It is compressed on the day (Sun) side due to the force of the arriving particles, and extended on the night side. (Image not to scale.)

Earth is blanketed by an atmosphere consisting of 78.0% nitrogen, 20.9% oxygen, and 1% Argon. The atmosphere has five layers: troposphere, stratosphere, mesosphere, thermosphere, and exosphere; and 75% of the atmosphere's gases are in the bottom-most layer, the troposphere.

The magnetic field created by mantle's internal motions produces the magnetosphere which protects the Earth's atmosphere from the solar wind. It is theorized that the solar wind would strip away earth's atmosphere in a few million years were it not for the Earth's electromagnet. And since earth is 4.5 billion years old, earth would not have an atmosphere by now if there were no magnetosphere.

The atmosphere is composed of 78% nitrogen and 21% oxygen. The remaining one percent contains small amounts of other gases including CO<sub>2</sub> and water vapors. Water vapors and CO<sub>2</sub> allow the Earth's atmosphere to catch and hold the sun's energy through a phenomenon called the greenhouse effect. This allows earth's surface to be warm enough to have liquid water and support life.

In addition to storing heat, the atmosphere also protects living organisms by shielding the Earth's surface from cosmic rays. Note that the level of protection is high enough to prevent cosmic rays from destroying all life on Earth, yet low enough to aid the mutations that have an important role in pushing forward diversity in the biosphere.

## **Methodology**

Like all other scientists, Earth scientists apply the scientific method. They formulate hypotheses after observing events and gathering data about natural phenomena, and then they test hypotheses from such data.

A contemporary idea within earth science is uniformitarianism. Uniformitarianism says that "ancient geologic features are interpreted by understanding active processes that are readily observed". Simply stated, this means that features of the Earth can be explained by the actions of gradual processes operating over long periods of time; for example, a mountain need not be thought of as having been created in a moment, but instead it may be seen as the result of continuous subduction, causing magma to rise and form continental volcanic arcs.

## **Earth's spheres**

Earth science generally recognizes four spheres, the lithosphere, the hydrosphere, the atmosphere, and the biosphere; these correspond to rocks, water, air, and life. Some practitioners include, as part of the spheres of the Earth, the cryosphere (corresponding to ice) as a distinct portion of the hydrosphere, as well as the pedosphere (corresponding to soil) as an active and intermixed sphere.

## Chapter 2

# Branches & Related Fields of Earth Science

## Geology

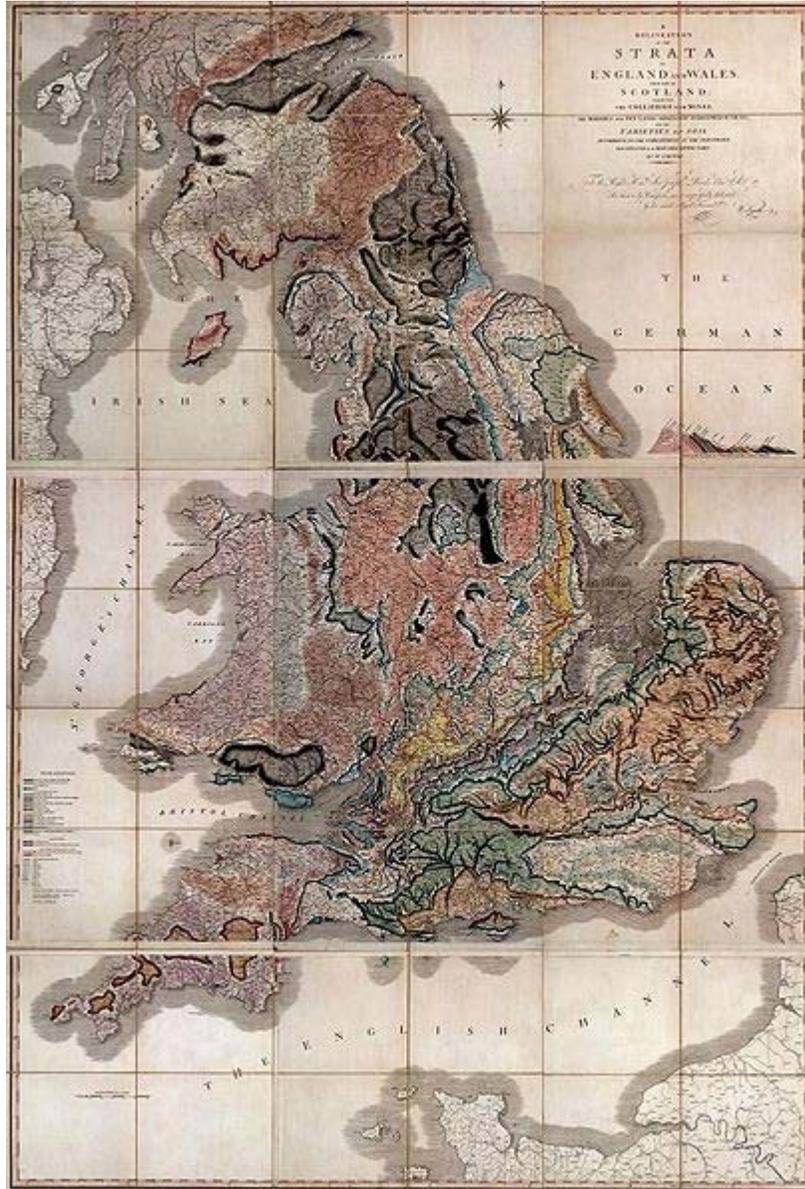


Students examining the Wasatch Fault near Salt Lake City, Utah

**Geology** is the science and study of the solid Earth and the processes by which it is shaped and changed. Geology provides primary evidence for plate tectonics, the history of life and evolution, and past climates. In modern times, geology is commercially important for mineral and hydrocarbon exploration, is publically important for predicting

and understanding natural hazards, plays an essential role in geotechnical engineering, and is a major academic discipline.

## History



William Smith's geologic map of England, Wales, and southern Scotland. Completed in 1815, it was the first national-scale geologic map, and by far the most accurate of its time.

The study of the physical material of the Earth dates back at least to ancient Greece when Theophrastus (372-287 BC) wrote the work *Peri Lithon* (*On Stones*). In the Roman

period, Pliny the Elder wrote in detail of the many minerals and metals then in practical use, and correctly noted the origin of amber.

Some modern scholars, such as Fielding H. Garrison, are of the opinion that modern geology began in the medieval Islamic world. Abu al-Rayhan al-Biruni (973–1048 AD) was one of the earliest Muslim geologists, whose works included the earliest writings on the geology of India, hypothesizing that the Indian subcontinent was once a sea. Islamic Scholar Ibn Sina (Avicenna, 981–1037) proposed detailed explanations for the formation of mountains, the origin of earthquakes, and other topics central to modern Geology, which provided an essential foundation for the later development of the science. In China, the polymath Shen Kua (1031–1095) formulated a hypothesis for the process of land formation: based on his observation of fossil animal shells in a geological stratum in a mountain hundreds of miles from the ocean, he inferred that the land was formed by erosion of the mountains and by deposition of silt.

Nicolas Steno (1638–1686) is credited with the law of superposition, the principle of original horizontality, and the principle of lateral continuity: three defining principles of stratigraphy.

The word *geology* was first used by Jean-André Deluc in 1778 and introduced as a fixed term by Horace-Bénédict de Saussure in 1779. The word is derived from the Greek γῆ, *gê*, meaning “earth” and λόγος, *logos*, meaning “speech”.

William Smith (1769–1839) drew some of the first geological maps and began the process of ordering rock strata (layers) by examining the fossils contained in them.

James Hutton is often viewed as the first modern geologist. In 1785 he presented a paper entitled *Theory of the Earth* to the Royal Society of Edinburgh. In his paper, he explained his theory that the Earth must be much older than had previously been supposed in order to allow enough time for mountains to be eroded and for sediments to form new rocks at the bottom of the sea, which in turn were raised up to become dry land. Hutton published a two-volume version of his ideas in 1795 (Vol. 1, Vol. 2).

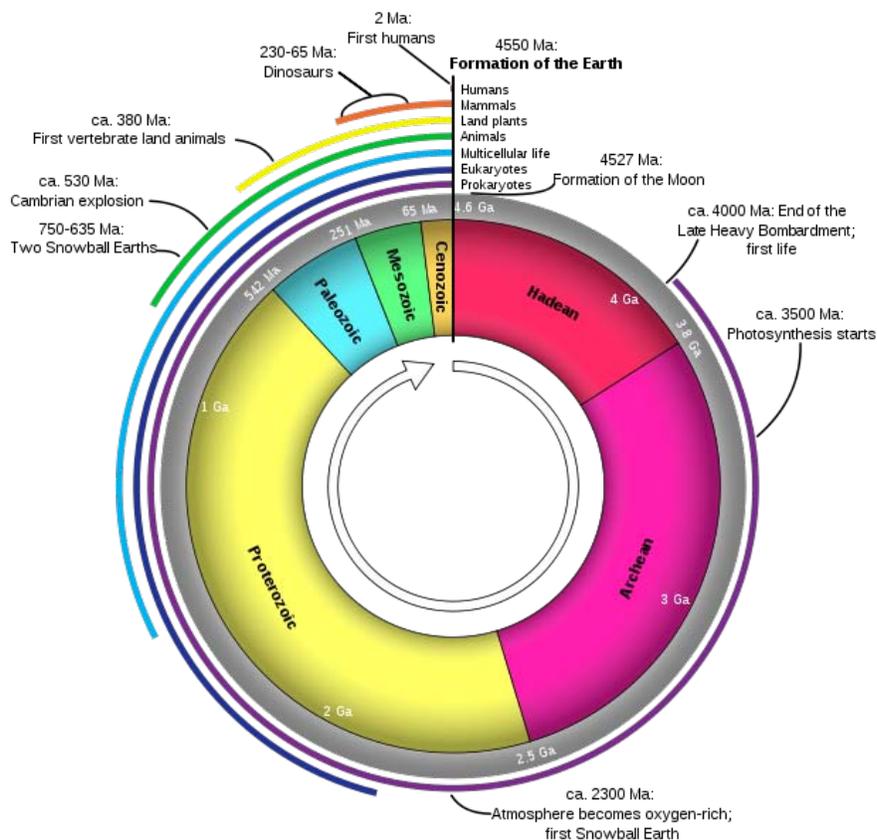
Followers of Hutton were known as *Plutonists* because they believed that some rocks were formed by vulcanism which is the deposition of lava from volcanoes, as opposed to the *Neptunists*, who believed that all rocks had settled out of a large ocean whose level gradually dropped over time.

Sir Charles Lyell first published his famous book, *Principles of Geology*, in 1830. The book, which influenced the thought of Charles Darwin, successfully promoted the doctrine of uniformitarianism. This theory states that slow geological processes have occurred throughout the Earth’s history and are still occurring today. In contrast, catastrophism is the theory that Earth’s features formed in single, catastrophic events and remained unchanged thereafter. Though Hutton believed in uniformitarianism, the idea was not widely accepted at the time.

Much of 19<sup>th</sup>-century geology revolved around the question of the Earth's exact age. Estimates varied from a few 100,000 to billions of years. By the early 20<sup>th</sup> century, radiometric dating allowed the Earth's age to be estimated at two billion years. The awareness of this vast amount of time opened the door to new theories about the processes that shaped the planet.

The most significant advances in 20<sup>th</sup> century geology have been the development of the theory of plate tectonics in the 1960s, and the refinement of estimates of the planet's age. Plate tectonic theory arose out of two separate geological observations: seafloor spreading and continental drift. The theory revolutionized the Earth sciences. Today the Earth is known to be approximately 4.5 billion years old.

## Geologic time



Geological time put in a diagram called a geological clock, showing the relative lengths of the eons of the Earth's history.

The geologic time scale encompasses the history of the Earth. It is bracketed at the old end by the dates of the earliest solar system material at 4.567 Ga, (gigaannum: billion years ago) and the age of the Earth at 4.54 Ga at the beginning of the informally recognized Hadean eon. At the young end of the scale, it is bracketed by the present day in the Holocene epoch.

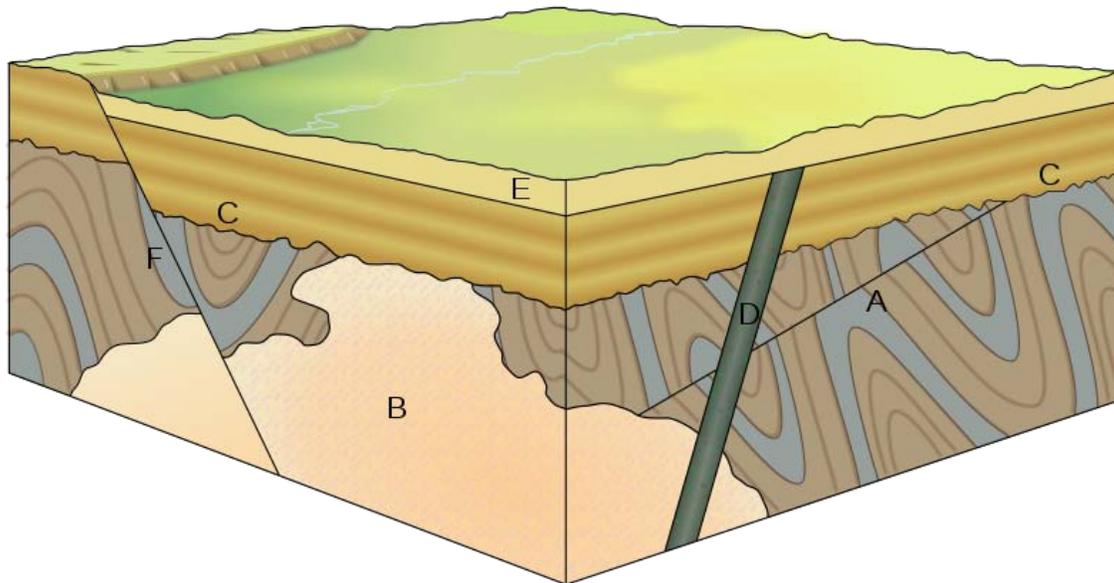
## Important milestones

- 4.567 Ga: Solar system formation
- 4.54 Ga: Accretion of Earth
- c. 4 Ga: End of Late Heavy Bombardment, first life
- c. 3.5 Ga: Start of photosynthesis
- c. 2.3 Ga: Oxygenated atmosphere, first snowball Earth
- 730–635 Ma (megaannum: million years ago): two snowball Earths
- 542± 0.3 Ma: Cambrian explosion – vast multiplication of hard-bodied life; first abundant fossils; start of the Paleozoic
- c. 380 Ma: First vertebrate land animals
- 250 Ma: Permian-Triassic extinction – 90% of all land animals die. End of Paleozoic and beginning of Mesozoic
- 65 Ma: Cretaceous-Tertiary extinction – Dinosaurs die; end of Mesozoic and beginning of Cenozoic
- c. 7 Ma – Present: Hominins
  - c. 7 Ma: First hominins appear
  - 3.9 Ma: First Australopithecus, direct ancestor to modern Homo sapiens, appear
  - 200 ka (kiloannum: thousand years ago): First modern Homo sapiens appear in East Africa

## Relative and absolute dating

Geological events can be given a precise date at a point in time, or they can be related to other events that came before and after them. Geologists use a variety of methods to give both relative and absolute dates to geological events. They then use these dates to find the rates at which processes occur.

## Relative dating



Cross-cutting relations can be used to determine the relative ages of rock strata and other geological structures. Explanations: A – folded rock strata cut by a thrust fault; B – large intrusion (cutting through A); C – erosional angular unconformity (cutting off A & B) on which rock strata were deposited; D – volcanic dyke (cutting through A, B & C); E – even younger rock strata (overlying C & D); F – normal fault (cutting through A, B, C & E).

Methods for relative dating were developed when geology first emerged as a formal science. Geologists still use the following principles today as a means to provide information about geologic history and the timing of geologic events.

**The principle of intrusive relationships** concerns crosscutting intrusions. In geology, when an igneous intrusion cuts across a formation of sedimentary rock, it can be determined that the igneous intrusion is younger than the sedimentary rock. There are a number of different types of intrusions, including stocks, laccoliths, batholiths, sills and dikes.

**The principle of cross-cutting relationships** pertains to the formation of faults and the age of the sequences through which they cut. Faults are younger than the rocks they cut; accordingly, if a fault is found that penetrates some formations but not those on top of it, then the formations that were cut are older than the fault, and the ones that are not cut must be younger than the fault. Finding the key bed in these situations may help determine whether the fault is a normal fault or a thrust fault.

**The principle of inclusions and components** states that, with sedimentary rocks, if inclusions (or *clasts*) are found in a formation, then the inclusions must be older than the formation that contains them. For example, in sedimentary rocks, it is common for gravel

from an older formation to be ripped up and included in a newer layer. A similar situation with igneous rocks occurs when xenoliths are found. These foreign bodies are picked up as magma or lava flows, and are incorporated, later to cool in the matrix. As a result, xenoliths are older than the rock which contains them.

**The principle of uniformitarianism** states that the geologic processes observed in operation that modify the Earth's crust at present have worked in much the same way over geologic time. A fundamental principle of geology advanced by the 18<sup>th</sup> century Scottish physician and geologist James Hutton, is that "the present is the key to the past." In Hutton's words: "the past history of our globe must be explained by what can be seen to be happening now."



The Permian through Jurassic stratigraphy of the Colorado Plateau area of southeastern Utah is a great example of both Original Horizontality and the Law of Superposition. These strata make up much of the famous prominent rock formations in widely spaced protected areas such as Capitol Reef National Park and Canyonlands National Park. From top to bottom: Rounded tan domes of the Navajo Sandstone, layered red Kayenta Formation, cliff-forming, vertically jointed, red Wingate Sandstone, slope-forming, purplish Chinle Formation, layered, lighter-red Moenkopi Formation, and white, layered Cutler Formation sandstone. Picture from Glen Canyon National Recreation Area, Utah.

**The principle of original horizontality** states that the deposition of sediments occurs as essentially horizontal beds. Observation of modern marine and non-marine sediments in a wide variety of environments supports this generalization (although cross-bedding is inclined, the overall orientation of cross-bedded units is horizontal).

**The principle of superposition** states that a sedimentary rock layer in a tectonically undisturbed sequence is younger than the one beneath it and older than the one above it.

Logically a younger layer cannot slip beneath a layer previously deposited. This principle allows sedimentary layers to be viewed as a form of vertical time line, a partial or complete record of the time elapsed from deposition of the lowest layer to deposition of the highest bed.

**The principle of faunal succession** is based on the appearance of fossils in sedimentary rocks. As organisms exist at the same time period throughout the world, their presence or (sometimes) absence may be used to provide a relative age of the formations in which they are found. Based on principles laid out by William Smith almost a hundred years before the publication of Charles Darwin's theory of evolution, the principles of succession were developed independently of evolutionary thought. The principle becomes quite complex, however, given the uncertainties of fossilization, the localization of fossil types due to lateral changes in habitat (facies change in sedimentary strata), and that not all fossils may be found globally at the same time.

### **Absolute dating**

Geologists can also give precise absolute dates to geologic events. These dates are useful on their own, and can also be used in conjunction with relative dating methods or to calibrate relative dating methods.

A large advance in geology in the advent of the 20<sup>th</sup> century was the ability to give precise absolute dates to geologic events through radioactive isotopes and other methods. The advent of radiometric dating changed the understanding of geologic time. Before, geologists could only use fossils to date sections of rock relative to one another. With isotopic dates, absolute dating became possible, and these absolute dates could be applied fossil sequences in which there was datable material, converting the old relative ages into new absolute ages.

For many geologic applications, isotope ratios are measured in minerals that give the amount of time that has passed since a rock passed through its particular closure temperature, the point at which different radiometric isotopes stop diffusing into and out of the crystal lattice. These are used in geochronologic and thermochronologic studies. Common methods include uranium-lead dating, potassium-argon dating and argon-argon dating, and uranium-thorium dating. These methods are used for a variety of applications. Dating of lavas and ash layers can help to date stratigraphy and calibrate relative dating techniques. These methods can also be used to determine ages of pluton emplacement. Thermochemical techniques can be used to determine temperature  $\square$ riticiz within the crust, the uplift of mountain ranges, and paleotopography.

Fractionation of the lanthanide series elements is used to compute ages since rocks were removed from the mantle.

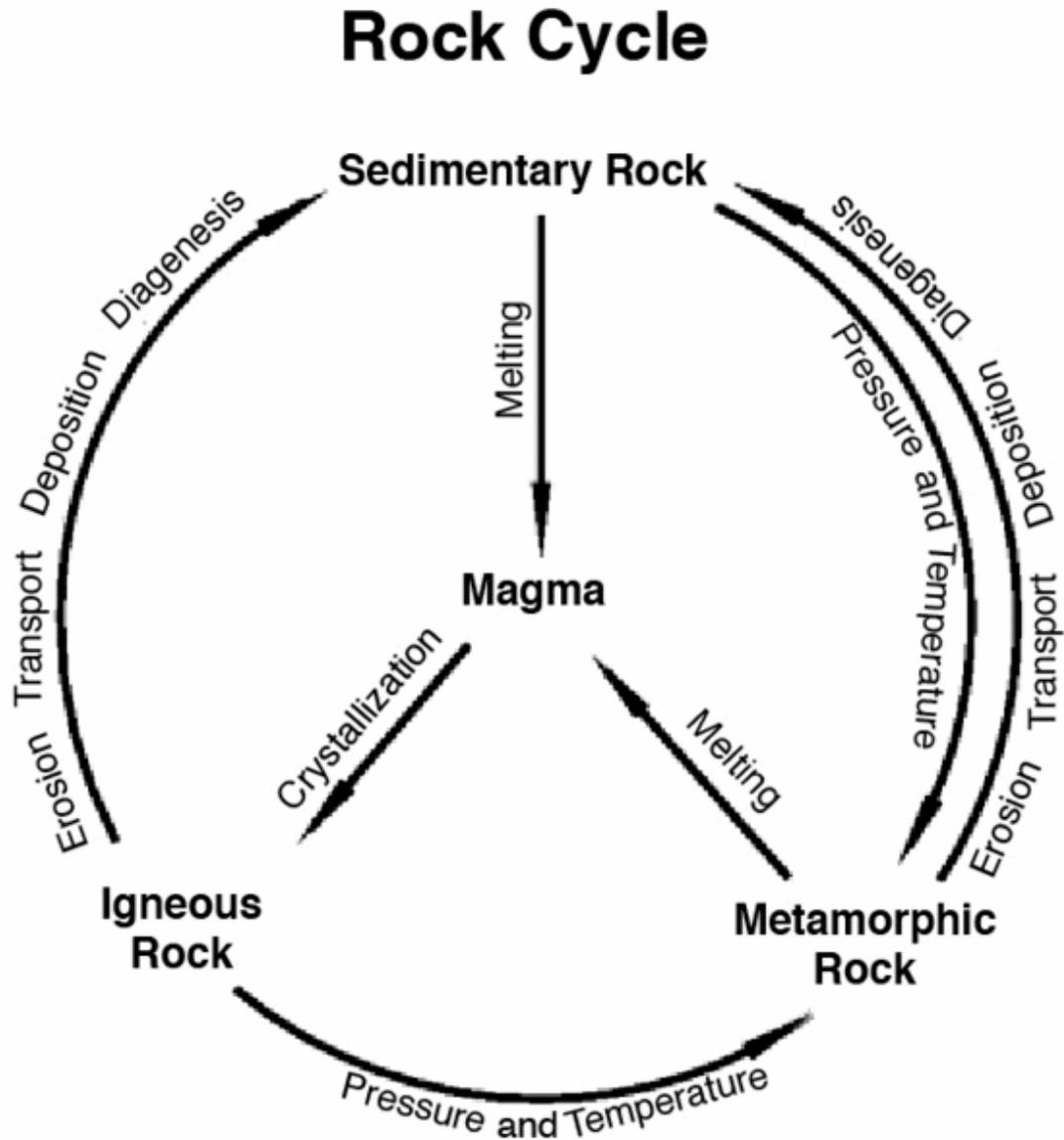
Other methods are used for more recent events. Optically stimulated luminescence and cosmogenic radionuclide dating are used to date surfaces and/or erosion rates.

Dendrochronology can also be used for the dating of landscapes. Radiocarbon dating is used for young organic material.

## Geologic materials

The majority of geological data come from research on solid Earth materials. These typically fall into one of two categories: rock and unconsolidated material.

### Rock



This schematic diagram of the rock cycle shows the relationship between magma and sedimentary, metamorphic, and igneous rock

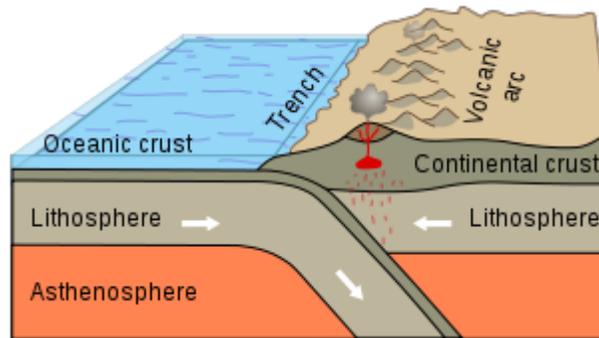
There are three major types of rock: igneous, sedimentary, and metamorphic. The rock cycle is an important concept in geology which illustrates the relationships between these three types of rock, and magma. When a rock crystallizes from melt (magma and/or lava), it is an igneous rock. This rock can be weathered and eroded, and then redeposited and lithified into a sedimentary rock, or be turned into a metamorphic rock due to heat and pressure that change the mineral content of the rock and give it a characteristic fabric. The sedimentary rock can then be subsequently turned into a metamorphic rock due to heat and pressure, and the metamorphic rock can be weathered, eroded, deposited, and lithified, becoming a sedimentary rock. Sedimentary rock may also be re-eroded and redeposited, and metamorphic rock may also undergo additional metamorphism. All three types of rocks may be re-melted; when this happens, a new magma is formed, from which an igneous rock may once again crystallize.

The majority of research in geology is associated with the study of rock, as rock provides the primary record of the majority of the geologic history of the Earth.

### **Unconsolidated material**

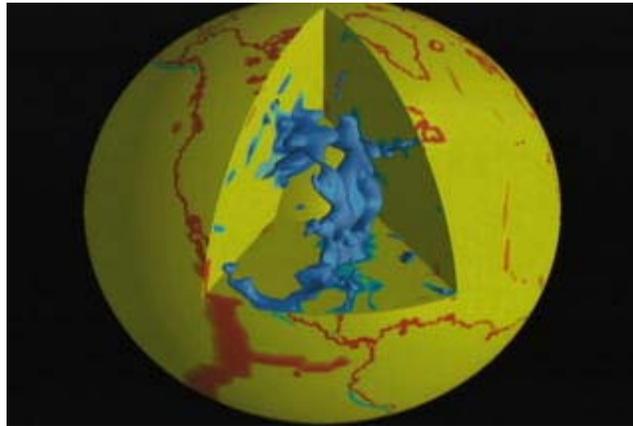
Geologists also study unlithified material, which typically comes from more recent deposits. Because of this, the study of such material is often known as Quaternary geology, after the recent Quaternary Period. This includes the study of sediment and soils, and is important to some (or many) studies in geomorphology, sedimentology, and paleoclimatology.

### **Whole-Earth structure**



Oceanic-continental convergence resulting in subduction and volcanic arcs illustrates one effect of plate tectonics.

## Plate tectonics

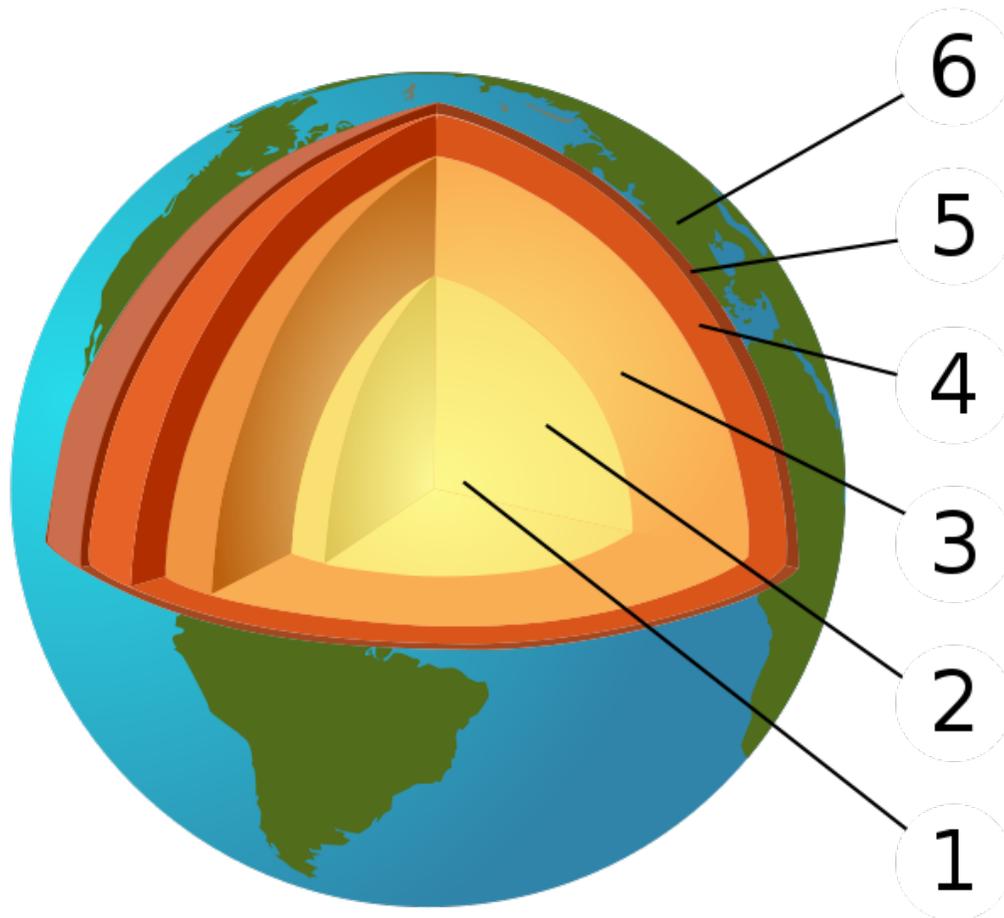


On this diagram, subducting slabs are in blue, and continental margins and a few plate boundaries are in red. The blue blob in the cutaway section is the seismically imaged Farallon Plate, which is subducting beneath North America. The remnants of this plate on the Surface of the Earth are the Juan de Fuca Plate and Explorer plate in the Northwestern USA / Southwestern Canada, and the Cocos Plate on the west coast of Mexico.

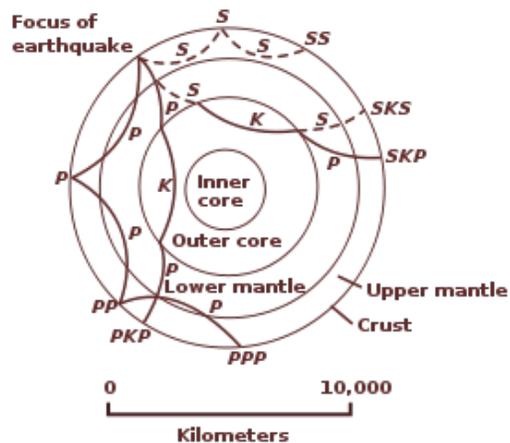
In the 1960s, a series of discoveries, the most important of which was seafloor spreading, showed that the Earth's lithosphere, which includes the crust and rigid uppermost portion of the upper mantle, is separated into a number of tectonic plates that move across the plastically deforming, solid, upper mantle, which is called the asthenosphere. There is an intimate coupling between the movement of the plates on the surface and the convection of the mantle: oceanic plate motions and mantle convection currents always move in the same direction, because the oceanic lithosphere is the rigid upper thermal boundary layer of the convecting mantle. This coupling between rigid plates moving on the surface of the Earth and the convecting mantle is called plate tectonics.

The development of plate tectonics provided a physical basis for many observations of the solid Earth. Long linear regions of geologic features could be explained as plate boundaries. Mid-ocean ridges, high regions on the seafloor where hydrothermal vents and volcanoes exist, were explained as divergent boundaries, where two plates move apart. Arcs of volcanoes and earthquakes were explained as convergent boundaries, where one plate subducts under another. Transform boundaries, such as the San Andreas fault system, resulted in widespread powerful earthquakes. Plate tectonics also provided a mechanism for Alfred Wegener's theory of continental drift, in which the continents move across the surface of the Earth over geologic time. They also provided a driving force for crustal deformation, and a new setting for the observations of structural geology. The power of the theory of plate tectonics lies in its ability to combine all of these observations into a single theory of how the lithosphere moves over the convecting mantle.

## Earth structure



The Earth's layered structure. (1) inner core; (2) outer core; (3) lower mantle; (4) upper mantle; (5) lithosphere; (6) crust



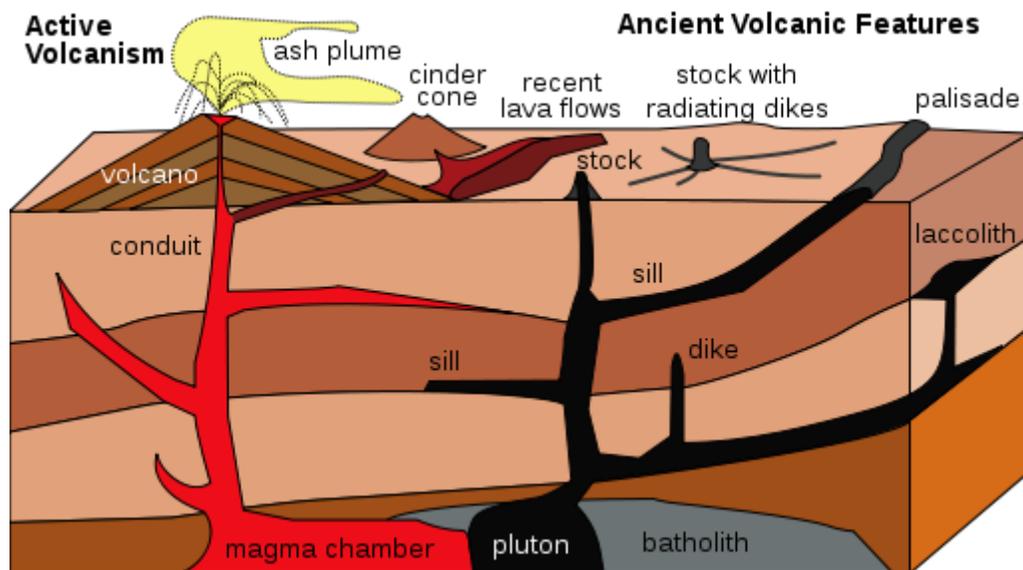
Earth layered structure. Typical wave paths from earthquakes like these gave early seismologists insights into the layered structure of the Earth

Advances in seismology, computer modeling, and mineralogy and crystallography at high temperatures and pressures give insights into the internal composition and structure of the Earth.

Seismologists can use the arrival times of seismic waves in reverse to image the interior of the Earth. Early advances in this field showed the existence of a liquid outer core (where shear waves were not able to propagate) and a dense solid inner core. These advances led to the development of a layered model of the Earth, with a crust and lithosphere on top, the mantle below (separated within itself by seismic discontinuities at 410 and 660 kilometers), and the outer core and inner core below that. More recently, seismologists have been able to create detailed images of wave speeds inside the earth in the same way a doctor images a body in a CT scan. These images have led to a much more detailed view of the interior of the Earth, and have replaced the simplified layered model with a much more dynamic model.

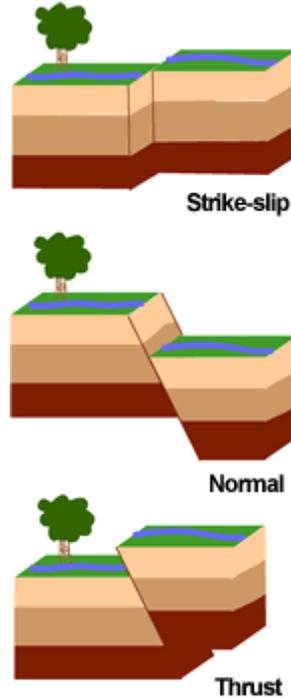
Mineralogists have been able to use the pressure and temperature data from the seismic and isotopic studies alongside knowledge of the elemental composition of the Earth at depth to reproduce these conditions in experimental settings and measure changes in crystal structure. These studies explain the chemical changes associated with the major seismic discontinuities in the mantle, and show the crystallographic structures expected in the inner core of the Earth.

## Geological evolution of an area



An originally horizontal sequence of sedimentary rocks (in shades of tan) are affected by igneous activity. Deep below the surface are a magma chamber and large associated igneous bodies. The magma chamber feeds the volcano, and sends off shoots of magma that will later crystallize into dikes and sills. Magma also advances upwards to form

intrusive igneous bodies. The diagram illustrates both a cinder cone volcano, which releases ash, and a composite volcano, which releases both lava and ash.



An illustration of the three types of faults. Strike-slip faults occur when rock units slide past one another, normal faults occur when rocks are undergoing horizontal extension, and thrust faults occur when rocks are undergoing horizontal shortening.

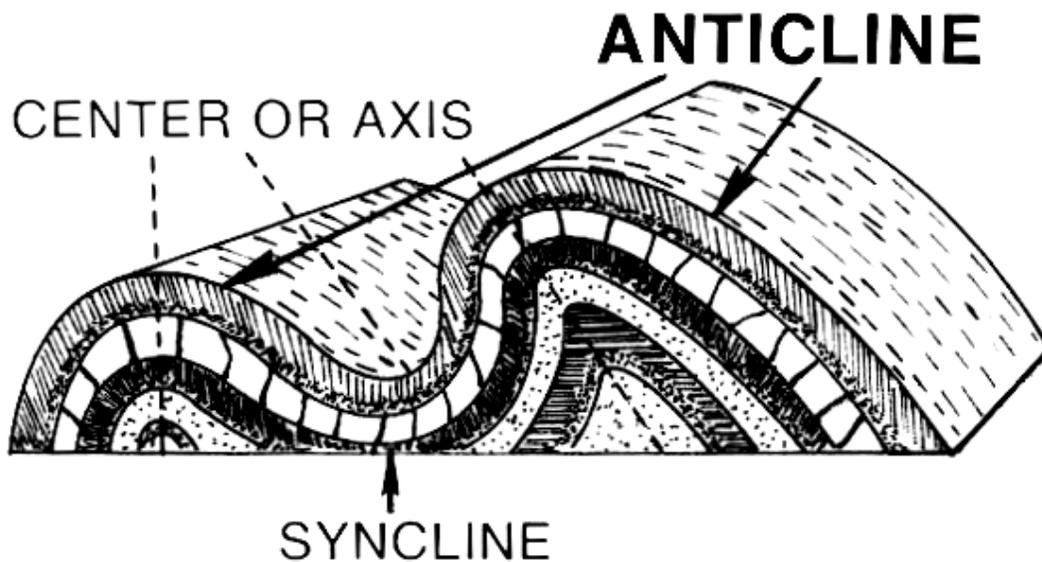
The geology of an area evolves through time as rock units are deposited and inserted and deformational processes change their shapes and locations.

Rock units are first emplaced either by deposition onto the surface or intrusion into the overlying rock. Deposition can occur when sediments settle onto the surface of the Earth and later lithify into sedimentary rock, or when as volcanic material such as volcanic ash or lava flows blanket the surface. Igneous intrusions such as batholiths, laccoliths, dikes, and sills, push upwards into the overlying rock, and crystallize as they intrude.

After the initial sequence of rocks has been deposited, the rock units can be deformed and/or metamorphosed. Deformation typically occurs as a result of horizontal shortening, horizontal extension, or side-to-side (strike-slip) motion. These structural regimes broadly relate to convergent boundaries, divergent boundaries, and transform boundaries, respectively, between tectonic plates.

When rock units are placed under horizontal compression, they shorten and become thicker. Because rock units, other than muds, do not significantly change in volume, this is accomplished in two primary ways: through faulting and folding. In the shallow crust, where brittle deformation can occur, thrust faults form, which cause deeper rock to move on top of shallower rock. Because deeper rock is often older, as noted by the principle of

superposition, this can result in older rocks moving on top of younger ones. Movement along faults can result in folding, either because the faults are not planar, or because the rock layers are dragged along, forming drag folds, as slip occurs along the fault. Deeper in the Earth, rocks behave plastically, and fold instead of faulting. These folds can either be those where the material in the center of the fold buckles upwards, creating “antiforms”, or where it buckles downwards, creating “synforms”. If the tops of the rock units within the folds remain pointing upwards, they are called anticlines and synclines, respectively. If some of the units in the fold are facing downward, the structure is called an overturned anticline or syncline, and if all of the rock units are overturned or the correct up-direction is unknown, they are simply called by the most general terms, antiforms and synforms.



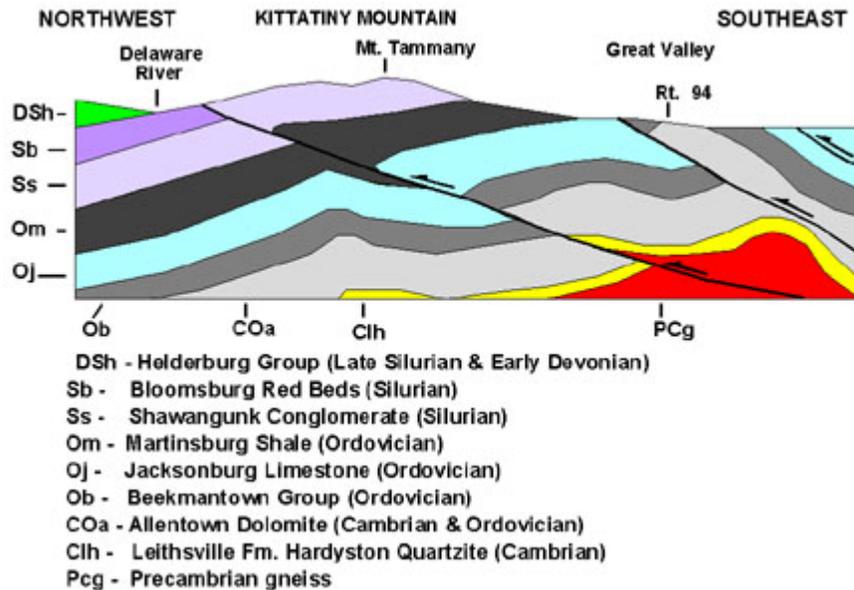
A diagram of folds, indicating an anticline and a syncline.

Even higher pressures and temperatures during horizontal shortening can cause both folding and metamorphism of the rocks. This metamorphism causes changes in the mineral composition of the rocks; creates a foliation, or planar surface, that is related to mineral growth under stress; and can remove signs of the original textures of the rocks, such as bedding in sedimentary rocks, flow features of lavas, and crystal patterns in crystalline rocks.

Extension causes the rock units as a whole to become longer and thinner. This is primarily accomplished through normal faulting and through the ductile stretching and thinning. Normal faults drop rock units that are higher below those that are lower. This typically results in younger units being placed below older units. Stretching of units can result in their thinning; in fact, there is a location within the Maria Fold and Thrust Belt in which the entire sedimentary sequence of the Grand Canyon can be seen over a length of less than a meter. Rocks at the depth to be ductilely stretched are often also

metamorphosed. These stretched rocks can also pinch into lenses, known as boudins, after the French word for “sausage”, because of their visual similarity.

Where rock units slide past one another, strike-slip faults develop in shallow regions, and become shear zones at deeper depths where the rocks deform ductilely.



Geologic cross-section of Kittatinny Mountain. This cross-section shows metamorphic rocks, overlain by younger sediments deposited after the metamorphic event. These rock units were later folded and faulted during the uplift of the mountain.

The addition of new rock units, both depositionally and intrusively, often occurs during deformation. Faulting and other deformational processes result in the creation of topographic gradients, causing material on the rock unit that is increasing in elevation to be eroded by hillslopes and channels. These sediments are deposited on the rock unit that is going down. Continual motion along the fault maintains the topographic gradient in spite of the movement of sediment, and continues to create accommodation space for the material to deposit. Deformational events are often also associated with volcanism and igneous activity. Volcanic ashes and lavas accumulate on the surface, and igneous intrusions enter from below. Dikes, long, planar igneous intrusions, enter along cracks, and therefore often form in large numbers in areas that are being actively deformed. This can result in the emplacement of dike swarms, such as those that are observable across the Canadian shield, or rings of dikes around the lava tube of a volcano.

All of these processes do not necessarily occur in a single environment, and do not necessarily occur in a single order. The Hawaiian Islands, for example, consist almost entirely of layered basaltic lava flows. The sedimentary sequences of the mid-continental United States and the Grand Canyon in the southwestern United States contain almost-undeformed stacks of sedimentary rocks that have remained in place since Cambrian time. Other areas are much more geologically complex. In the southwestern United

States, sedimentary, volcanic, and intrusive rocks have been metamorphosed, faulted, foliated, and folded. Even older rocks, such as the Acasta gneiss of the Slave craton in northwestern Canada, the oldest known rock in the world have been metamorphosed to the point where their origin is undiscernable without laboratory analysis. In addition, these processes can occur in stages. In many places, the Grand Canyon in the southwestern United States being a very visible example, the lower rock units were metamorphosed and deformed, and then deformation ended and the upper, undeformed units were deposited. Although any amount of rock emplacement and rock deformation can occur, and they can occur any number of times, these concepts provide a guide to understanding the geological history of an area.

## **Methods of geology**

Geologists use a number of field, laboratory, and numerical modeling methods to decipher Earth history and understand the processes that occur on and in the Earth. In typical geological investigations, geologists use primary information related to petrology (the study of rocks), stratigraphy (the study of sedimentary layers), and structural geology (the study of positions of rock units and their deformation). In many cases, geologists also study modern soils, rivers, landscapes, and glaciers; investigate past and current life and biogeochemical pathways, and use geophysical methods to investigate the subsurface.

## Field methods



A standard Brunton Geo compass, used commonly by geologists in mapping and surveying



A typical USGS field mapping camp in the 1950's



Today, handheld computers with GPS and geographic information systems software are often used in geological field work (digital geologic mapping).

Geological field work varies depending on the task at hand. Typical fieldwork could consist of:

- Geological mapping
  - Structural mapping: the locations of the major rock units and the faults and folds that led to their placement there.

- Stratigraphic mapping: the locations of sedimentary facies (lithofacies and biofacies) or the mapping of isopachs of equal thickness of sedimentary rock
  - Surficial mapping: the locations of soils and surficial deposits
- Surveying of topographic features
  - Creation of topographic maps
  - Work to understand change across landscapes, including:
    - Patterns of erosion and deposition
    - River channel change through migration and avulsion
    - Hillslope processes
- Subsurface mapping through geophysical methods
  - These methods include:
    - Shallow seismic surveys
    - Ground-penetrating radar
    - Electrical resistivity tomography
  - They are used for:
    - Hydrocarbon exploration
    - Finding groundwater
    - Locating buried archaeological artifacts
- High-resolution stratigraphy
  - Measuring and describing stratigraphic sections on the surface
  - Well drilling and logging
- Biogeochemistry and geomicrobiology
  - Collecting samples to:
    - Determine biochemical pathways
    - Identify new species of organisms. These organisms may help to show:
    - Identify new chemical compounds
  - And to use these discoveries to
    - Understand early life on Earth and how it functioned and metabolized
    - Find important compounds for use in pharmaceuticals.
- Paleontology: excavation of fossil material
  - For research into past life and evolution
  - For museums and education
- Collection of samples for geochronology and thermochronology
- Glaciology: measurement of characteristics of glaciers and their motion

## Laboratory methods



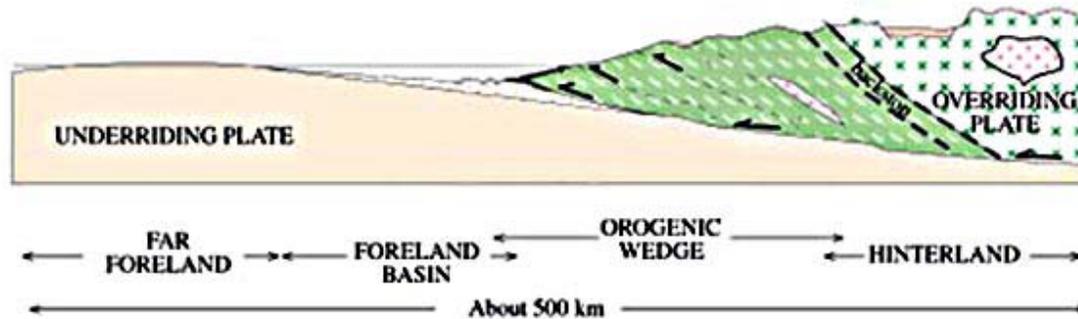
A petrographic microscope, which is an optical microscope fitted with cross-polarizing lenses, a conoscopic lens, and compensators (plates of anisotropic materials; gypsum plates and quartz wedges are common), for crystallographic analysis.

### Petrology

In addition to the field identification of rocks, petrologists identify rock samples in the laboratory. Two of the primary methods for identifying rocks in the laboratory are through optical microscopy and by using an electron microprobe. In an optical mineralogy analysis, thin sections of rock samples are analyzed through a petrographic microscope, where the minerals can be identified through their different properties in plane-polarized and cross-polarized light, including their birefringence, pleochroism, twinning, and interference properties with a conoscopic lens. In the electron microprobe, individual locations are analyzed for their exact chemical compositions and variation in composition within individual crystals. Stable and radioactive isotope studies provide insight into the geochemical evolution of rock units.

Petrologists use fluid inclusion data and perform high temperature and pressure physical experiments to understand the temperatures and pressures at which different mineral phases appear, and how they change through igneous and metamorphic processes. This research can be extrapolated to the field to understand metamorphic processes and the conditions of crystallization of igneous rocks. This work can also help to explain processes that occur within the Earth, such as subduction and magma chamber evolution.

## Structural geology



A diagram of an orogenic wedge. The wedge grows through faulting in the interior and along the main basal fault, called the *décollement*. It builds its shape into a critical taper, in which the angles within the wedge remain the same as failures inside the material balance failures along the *décollement*. It is analogous to a bulldozer pushing a pile of dirt, where the bulldozer is the overriding plate.

Structural geologists use microscopic analysis of oriented thin sections of geologic samples to observe the fabric within the rocks which gives information about strain within the crystal structure of the rocks. They also plot and combine measurements of geological structures in order to better understand the orientations of faults and folds in order to reconstruct the history of rock deformation in the area. In addition, they perform analog and numerical experiments of rock deformation in large and small settings.

The analysis of structures is often accomplished by plotting the orientations various features onto stereonet. A stereonet is a stereographic projection of a sphere onto a plane, in which planes are projected as lines and lines are projected as points. These can be used to find the locations of fold axes, relationships between several faults, and relationships between other geologic structures.

Among the most well-known experiments in structural geology are those involving orogenic wedges, which are zones in which mountains are built along convergent tectonic plate boundaries. In the analog versions of these experiments, horizontal layers of sand are pulled along a lower surface into a back stop, which results in realistic-looking patterns of faulting and the growth of a critically tapered (all angles remain the same) orogenic wedge. Numerical models work in the same way as these analog models, though they are often more sophisticated and can include patterns of erosion and uplift in the mountain belt. This helps to show the relationship between erosion and the shape of the mountain range. These studies can also give useful information about pathways for metamorphism through pressure, temperature, space, and time.

## Stratigraphy



Exploration geologists examining a freshly recovered drill core. Chile, 1994

In the laboratory, stratigraphers analyze samples of stratigraphic sections that can be returned from the field, such as those from drill cores. Stratigraphers also analyze data from geophysical surveys that show the locations of stratigraphic units in the subsurface. Geophysical data and well logs can be combined to produce a better view of the subsurface, and stratigraphers often use computer programs to do this in three dimensions. Stratigraphers can then use these data to reconstruct ancient processes occurring on the surface of the Earth, interpret past environments, and locate areas for water, coal, and hydrocarbon extraction.

In the laboratory, biostratigraphers analyze rock samples from outcrop and drill cores for the fossils found in them. These fossils help scientists to date the core and to understand the depositional environment in which the rock units formed. Geochronologists precisely date rocks within the stratigraphic section in order to provide better absolute bounds on the timing and rates of deposition. Magnetic stratigraphers look for signs of magnetic reversals in igneous rock units within the drill cores. Other scientists perform stable isotope studies on the rocks to gain information about past climate.

## Planetary geology



Surface of Mars as photographed by the Viking 2 lander December 9, 1977.

With the advent of space exploration in the twentieth century, geologists have begun to look at other planetary bodies in the same way as the Earth. This led to the establishment of the field of planetary geology, sometimes known as Astrogeology, in which geologic principles are applied to other bodies of the solar system.

Although the Greek-language-origin prefix *geo* refers to Earth, “geology” is often used in conjunction with the names of other planetary bodies when describing their composition and internal processes: examples are “the geology of Mars” and “Lunar geology”. Specialised terms such as *selenology* (studies of the Moon), *areology* (of Mars), etc., are also in use.

Although planetary geologists are interested in all aspects of the planets, a significant focus is in the search for past or present life on other worlds. This has led to many

missions whose purpose (or one of their purposes) is to examine planetary bodies for evidence of life. One of these is the Phoenix lander, which analyzed Martian polar soil for water and chemical and mineralogical constituents related to biological processes.

## **Applied geology**

### **Economic geology**

Economic geologists help locate and manage the Earth's natural resources, such as petroleum and coal, as well as mineral resources, which include metals such as iron, copper, and uranium.

### **Mining geology**

Mining geology consists of the extractions of mineral resources from the Earth. Some resources of economic interests include gemstones, metals, and many minerals such as asbestos, perlite, mica, phosphates, zeolites, clay, pumice, quartz, and silica, as well as elements such as sulfur, chlorine, and helium.

### **Petroleum geology**



Mud log in process, a common way to study the lithology when drilling oil wells.

Petroleum geologists study locations of the subsurface of the Earth which can contain extractable hydrocarbons, especially petroleum and natural gas. Because many of these reservoirs are found in sedimentary basins, they study the formation of these basins, as well as their sedimentary and tectonic evolution and the present-day positions of the rock units.

## **Engineering geology**

Engineering geology is the application of the geologic principles to engineering practice for the purpose of assuring that the geologic factors affecting the location, design, construction, operation and maintenance of engineering works are properly addressed.

In the field of civil engineering, geological principles and analyses are used in order to ascertain the mechanical principles of the material on which structures are built. This allows tunnels to be built without collapsing, bridges and skyscrapers to be built with sturdy foundations, and buildings to be built that will not settle in clay and mud.

## **Hydrology and environmental issues**

Geology and geologic principles can be applied to various environmental problems, such as stream restoration, the restoration of brownfields, and the understanding of the interactions between natural habitat and the geologic environment. Groundwater hydrology, or hydrogeology, is used to locate groundwater, which can often provide a ready supply of uncontaminated water and is especially important in arid regions, and to monitor the spread of contaminants in groundwater wells.

Geologists also obtain data through stratigraphy, boreholes, core samples, and ice cores. Ice cores and sediment cores are used to for paleoclimate reconstructions, which tell geologists about past and present temperature, precipitation, and sea level across the globe. These data are our primary source of information on global climate change outside of instrumental data.

## **Geophysics**

**Geophysics** is the physics of the Earth and its environment in space. Its subjects include the shape of the Earth, its gravitational and magnetic fields, the dynamics of the Earth as a whole and of its component parts, the Earth's internal structure, composition and tectonics, the generation of magmas, volcanism and rock formation, the hydrological cycle including snow and ice, all aspects of the oceans, the atmosphere, ionosphere, magnetosphere and solar-terrestrial relations, and analogous problems associated with the Moon and other planets.

Geophysics is also applied to societal needs, such as mineral resources, mitigation of natural hazards and environmental protection. Geophysical survey data are used to analyze potential petroleum reservoirs and mineral deposits, to locate groundwater, to locate archaeological finds, to find the thicknesses of glaciers and soils, and for environmental remediation.

## History



Replica of Zhang Heng's seismoscope.

## **Ancient and classical eras**

The magnetic compass existed in China back as far as the fourth century BC. It was used as much for feng shui as for navigation on land. It was not until good steel needles could be forged that compasses were used for navigation at sea; before that, they could not retain their magnetism for long. The first mention of a compass in Europe was in 1190.

In circa 240 BC, Erastotenes of Cyrene deduced that the Earth was round and measured the circumference of the Earth, using trigonometry and the angle of the Sun at more than one latitude in Egypt. He developed a system of latitude and longitude and measured the tilt of the Earth's axis.

Perhaps the earliest contribution to seismology was the invention of a seismoscope by the prolific inventor Zhang Heng in 132 CE. This instrument was designed to drop a bronze ball from the mouth of a dragon into the mouth of a toad. By looking at which of eight toads had the ball, one could determine the direction of the earthquake. It was 1571 years before the first design for a seismoscope was published in Europe, by Jean de la Hautefeuille. It was never built.

## **Beginnings of modern science**

One of the publications that marked the beginning of modern science was William Gilbert's *De Magnete* (1600), a report of a series of meticulous experiments in magnetism. Gilbert deduced that compasses point north because the Earth itself is magnetic.

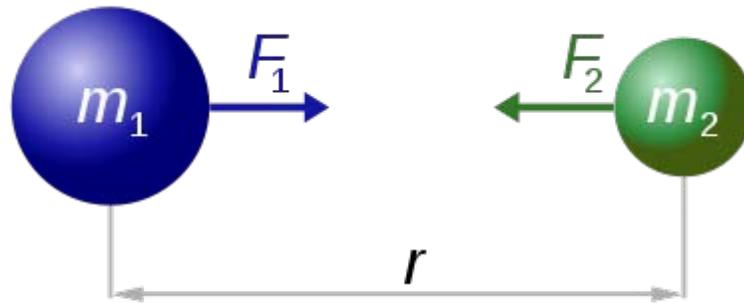
In 1687 Isaac Newton published his *Principia*, which not only laid the foundations for classical mechanics and gravitation but also explained a variety of geophysical phenomena such as the tides and the precession of the equinox.

The first seismometer, an instrument capable of keeping a continuous record of seismic activity, was built by James Forbes in 1844.

## **Physical phenomena**

Geophysics is a highly interdisciplinary subject and geophysicists contribute to every area of the Earth sciences. To provide a clearer idea of what constitutes geophysics, this section describes phenomena that are studied in physics and how they relate to the Earth and its surroundings.

## Gravity

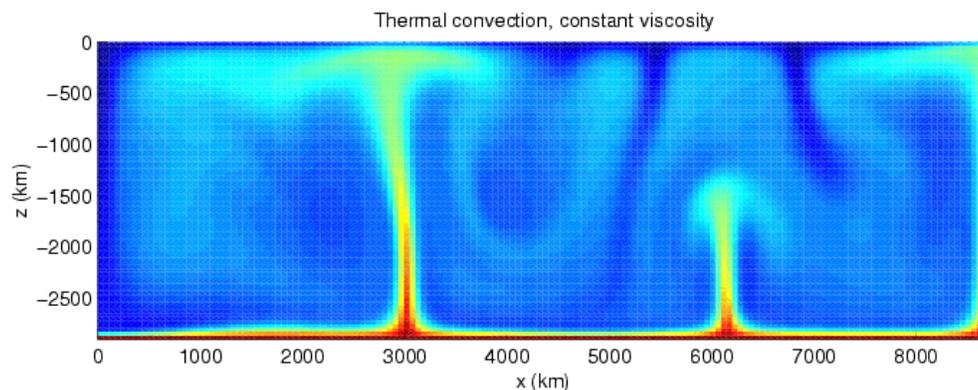


$$F_1 = F_2 = G \frac{m_1 \times m_2}{r^2}$$

The mechanism of Newton's law of universal gravitation.

The gravitational attraction of the Moon and Sun give rise to two high tides and two low tides a day. Gravitational forces cause rocks to press down on deeper rocks, making them increase in density as the depth increases. Measurements of gravitational acceleration and gravitational potential at the Earth's surface and above it can provide information on mineral deposits and the dynamics of plates. A particular geopotential surface called the geoid is one definition of the shape of the Earth: it would be the global mean sea level if the oceans were in equilibrium and could be extended through the continents (such as with very narrow canals).

## Heat flow

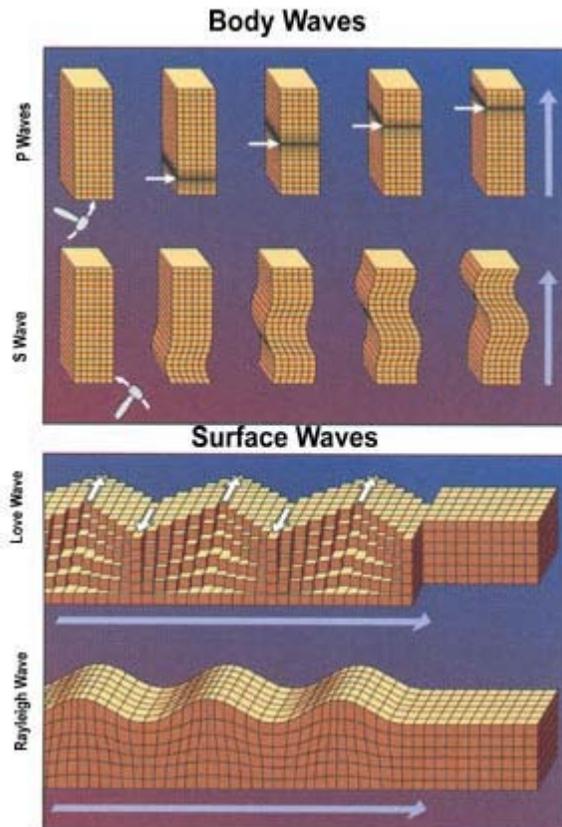


A model of thermal convection in the Earth's mantle.

The Earth is cooling, and the resulting heat flow generates the Earth's magnetic field through the geodynamo and plate tectonics through mantle convection. The main sources of heat are the primordial heat and radioactivity, although there are also contributions

from phase transitions. Heat is mostly carried to the surface by thermal convection, although there are two thermal boundary layers – the core-mantle boundary and the lithosphere – in which heat is transported by conduction. Some heat is carried up from the bottom of the mantle by mantle plumes. The heat flow at the Earth's surface is about  $4.2 \times 10^{13}$  W and it is a potential source of geothermal energy.

## Vibrations

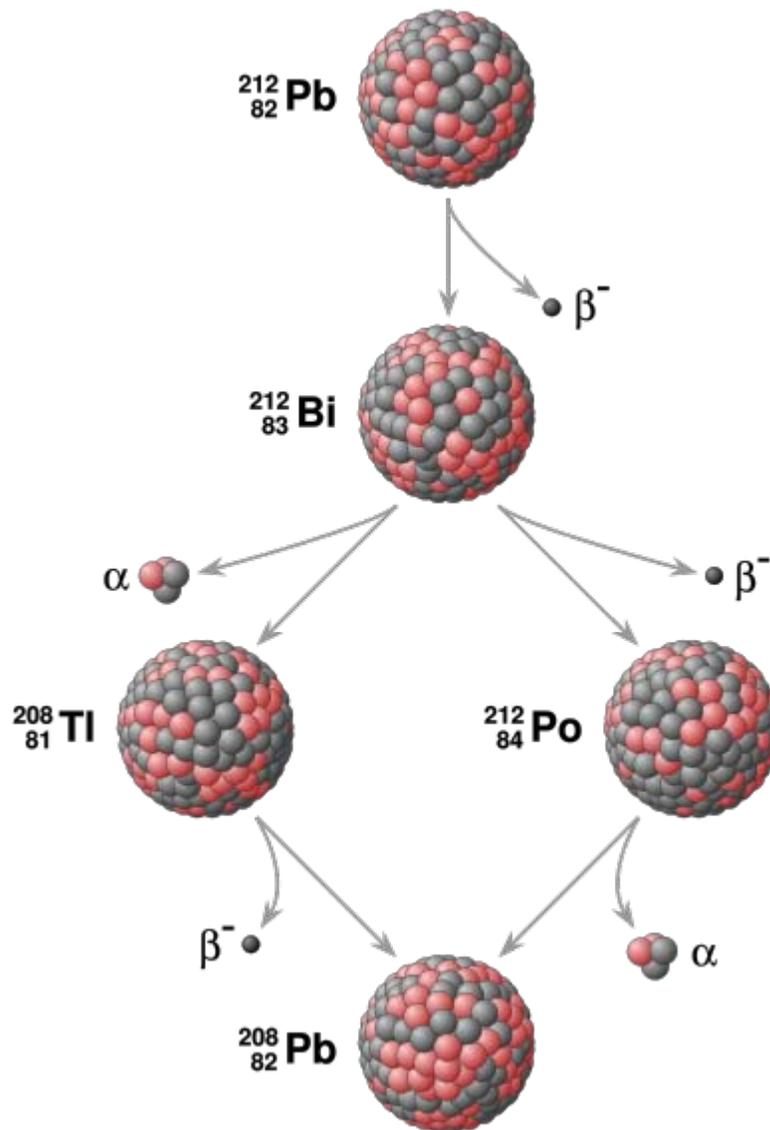


Body waves and surface waves

Vibrations of the Earth can take the form of seismic waves, which travel through the Earth's interior or along its surface, and normal modes or free oscillations. Seismic waves that are caused by localized sources such as earthquakes or explosions can be received at a distance by seismographs and the observed motion can provide information on the source as well as the Earth in between. Seismic reflections can provide information on near-surface structure while seismic refraction can be used to infer the deep structure of the Earth. The locations of earthquakes provide information on plate tectonics and mantle convection.

Earthquakes pose a risk to humans. Understanding their mechanisms (which depend on the type of earthquake, e.g., intraplate or deep focus can lead to better assessments of earthquake risk and improvements in earthquake engineering.

## Radioactivity



Example of a radioactive decay chain

Radioactive decay, in addition to being the main source of heat in the Earth, is an invaluable tool for geochronology. Unstable isotopes decay at predictable rates, and the decay rates of different isotopes cover several orders of magnitude, so radioactive decay can be used to accurately date both recent events and events in past geologic eras.

## Electricity

Although we mainly notice electricity during thunderstorms, there is always a downward electric field near the surface that averages  $120 \text{ V m}^{-1}$ . Relative to the solid Earth, the

atmosphere has a net positive charge due to bombardment by cosmic rays. A current of about 1800 A flows in the global circuit. It flows downward from the ionosphere over most of the Earth and back upwards through thunderstorms. The flow is manifested by lightning below the clouds and sprites above.

A variety of electric methods are used in geophysical survey. Some measure spontaneous potential a potential that arises in the ground because of man-made or natural disturbances. Telluric currents flow in Earth and the Oceans. They have two causes: electromagnetic induction by the time-varying, external-origin geomagnetic field and motion of conducting bodies (such as seawater) across the Earth's permanent magnetic field. The distribution of telluric current density can be used to detect variations in electrical resistivity of underground structures. Geophysicists can also provide the electric current themselves.

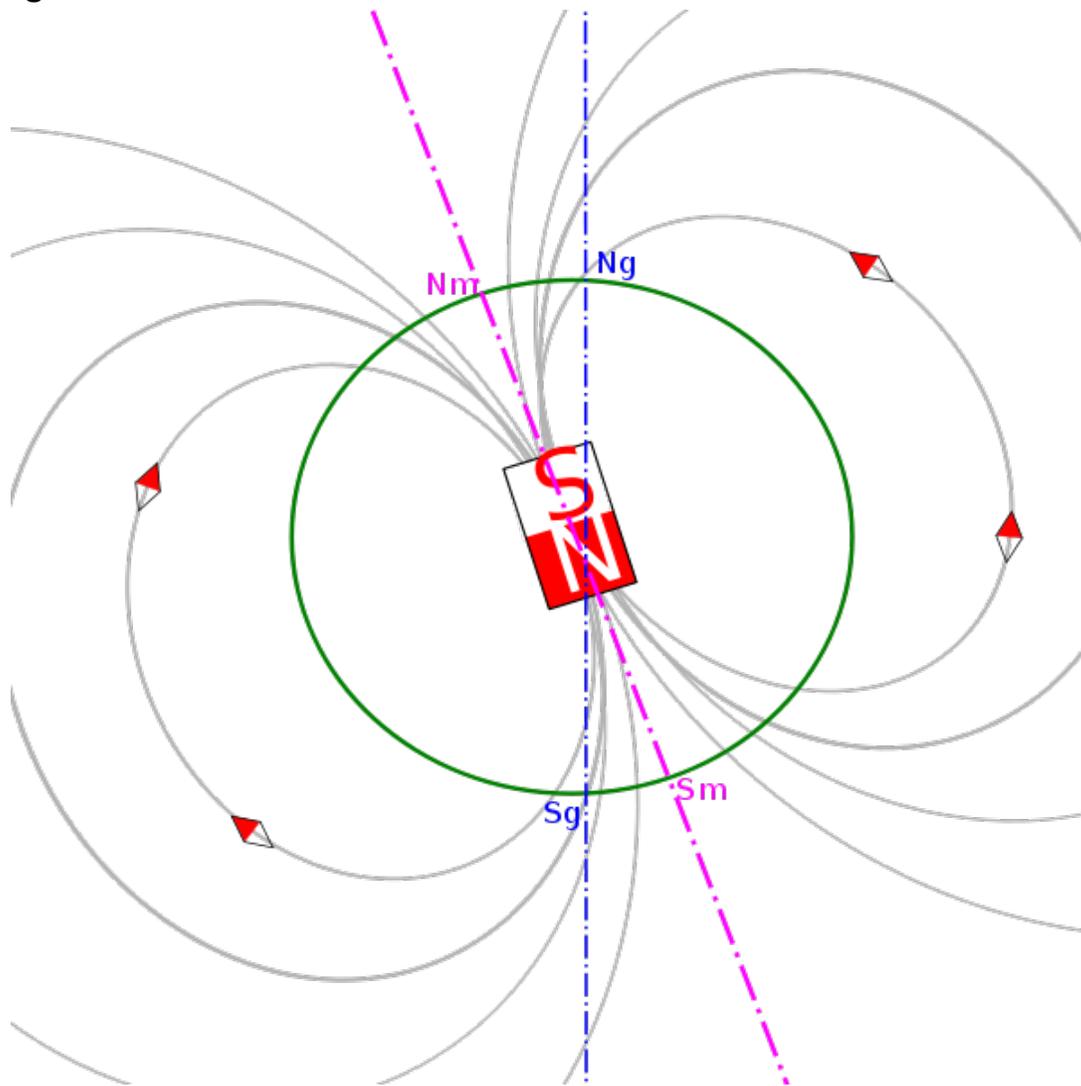
### **Electromagnetic waves**

Electromagnetic waves occur in the ionosphere and magnetosphere as well as the Earth's outer core. They have a variety of others. dawn chorus is caused by high-energy electrons that get caught in the Van Allen radiation belt. Whistlers are produced by lightning strikes. Hiss may be generated by both. Electromagnetic waves may also be generated by earthquakes.

In the Earth's outer core, electric currents in the highly conductive liquid iron create magnetic fields by magnetic induction. Alfvén waves are magnetohydrodynamic waves in the magnetosphere or the Earth's core. In the core, they probably have little observable effect on the geomagnetic field, but slower waves such as magnetic Rossby waves may be one source of secular variation.

Electromagnetic methods that are used for geophysical survey include transient electromagnetics and magnetotellurics.

## Magnetism



The variation between magnetic north and “true” north.

The Earth’s magnetic field protects the Earth from the deadly Solar wind and has long been used for navigation. It originates in the fluid motions of the Earth’s core. The magnetic field in the upper atmosphere gives rise to the auroras.

The Earth’s field is roughly like a tilted dipole, but it changes over time (a phenomenon called secular variation). Mostly the geomagnetic pole stays near the geographic pole, but at random intervals averaging a million years or so, the polarity of the Earth’s field reverses. These geomagnetic reversals are recorded in rocks and their signature can be seen in striped magnetic anomalies on the seafloor. These stripes provide quantitative information on seafloor spreading, a part of plate tectonics. In addition, the magnetization in rocks can be used to measure the motion of continents.

## **Fluid dynamics**

Fluid motions occur in the magnetosphere, atmosphere, ocean, mantle and core. Even the mantle, though it has an enormous viscosity, flows like a fluid over long time intervals. This flow is reflected in phenomena such as isostasy and post-glacial rebound. The mantle flow drives plate tectonics and the flow in the Earth's core drives the geodynamo.

Geophysical fluid dynamics is a primary tool in physical oceanography and meteorology. The rotation of the Earth has profound effects on the Earth's fluid dynamics, often due to the Coriolis effect. In the atmosphere it gives rise to large-scale patterns like Rossby waves and determines the basic circulation patterns of storms. In the ocean they drive large-scale circulation patterns as well as Kelvin waves and Ekman spirals at the ocean surface. In the Earth's core, the circulation of the molten iron is structured by Taylor columns.

Waves and other phenomena in the magnetosphere can be modeled using magnetohydrodynamics.

## **Condensed matter physics**

The physical properties of minerals must be understood to infer the composition of the Earth's interior from seismology, the geothermal gradient and other sources of information. Mineral physicists study the elastic properties of minerals as well as their high-pressure phase diagrams, melting points and equations of state at high pressure. Studies of creep determine how rocks that are brittle at the surface can flow deep down. These properties determine the rheology that determines the geodynamics.

Water is a very complex substance and its unique properties are essential for life. Its physical properties shape the hydrosphere and are an essential part of the water cycle and climate. Its thermodynamic properties determine evaporation and the thermal gradient in the atmosphere. The many types of precipitation involve a complex mixture of processes such as coalescence, supercooling and supersaturation. Some of the precipitated water becomes groundwater, and groundwater flow includes phenomena such as percolation, while the conductivity of water makes electrical and electromagnetic methods useful for tracking groundwater flow. Physical properties of water such as salinity have a large effect on its motion in the oceans.

The many phases of ice form the cryosphere and come in forms like ice sheets, glaciers, sea ice, freshwater ice, snow, and frozen ground (or permafrost).

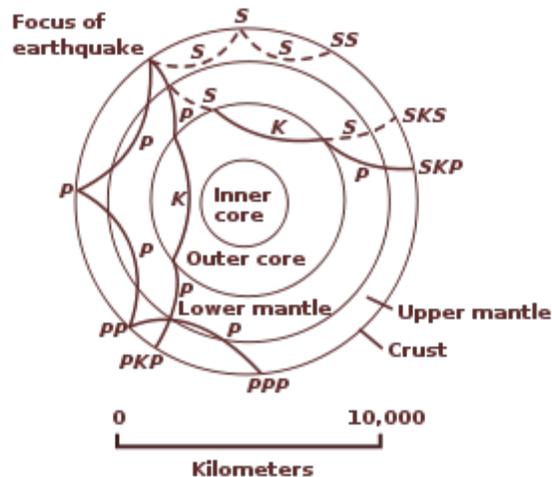
## **Regions of the Earth**

### **Size and form of the Earth**

The Earth is roughly spherical, but it bulges towards the Equator, so it is roughly in the shape of an ellipsoid. This bulge is due to its rotation and is nearly consistent with an

Earth in hydrostatic equilibrium. The detailed shape of the Earth, however, is also affected by the distribution of continents and ocean basins, and to some extent by the dynamics of the plates.

## Structure of the Earth



Mapping the interior of the Earth with earthquake waves.

Evidence from seismology, heat flow at the surface, and mineral physics is combined with the Earth's mass and moment of inertia to infer models of the Earth's interior – its composition, density, temperature, pressure. The Earth's mass is  $M = 5.975 \times 10^{24}$  kg and its mean radius is  $R = 6371$  km, so its mean specific gravity is  $\langle \rho \rangle = 5.515$ . This is substantially higher than the typical specific gravity (2.7–3.3) of rocks at the surface. Its moment of inertia is  $0.33 M R^2$ , whereas it would be  $0.4 M R^2$  if the earth was a sphere of constant density. Both lines of evidence point to a concentration of mass near the center. However, the density of the rock will increase with depth because of the increasing pressure. To determine how large this effect is, the Adams–Williamson equation is used to determine how density increases with pressure. The conclusion is that pressure alone cannot account for the increase in density. Instead, we know that the Earth's core is composed of an alloy of iron and other minerals.

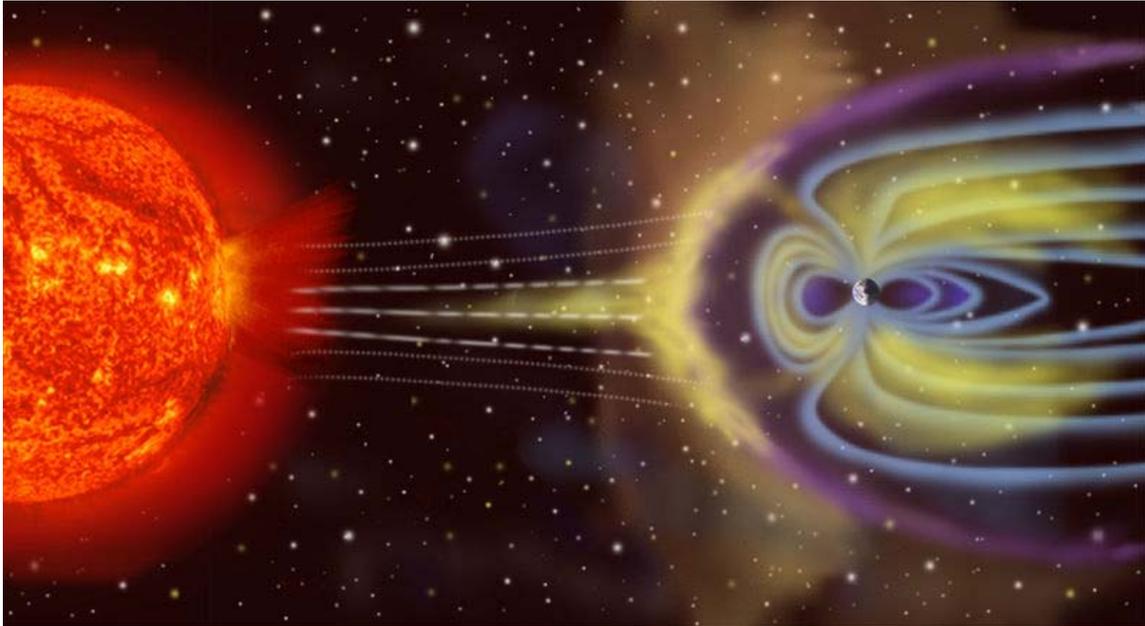
Reconstructions of seismic waves in the deep interior of the Earth show that there are no S-waves in the outer core. This indicates that the outer core is liquid, because liquids cannot support shear. The outer core is liquid, and the motion of this highly conductive fluid generates the Earth's field. The inner core, however, is solid because of the enormous pressure.

Reconstruction of seismic reflections in the deep interior indicate some major discontinuities in seismic velocities that demarcate the major zones of the Earth: inner core, outer core, mantle, lithosphere and crust. The mantle itself is divided into the upper mantle, transition zone, lower mantle and  $D''$  layer. Between the crust and the mantle is the Mohorovičić discontinuity.

The seismic model of the Earth does not by itself determine the composition of the layers. For a complete model of the Earth, mineral physics is needed to interpret seismic velocities in terms of composition. The mineral properties are temperature-dependent, so the geotherm must also be determined. This requires physical theory for thermal conduction and convection and the heat contribution of radioactive elements. The main model for the radial structure of the interior of the Earth is the Preliminary Reference Earth Model (PREM). Some parts of this model have been updated by recent findings in mineral physics and supplemented by seismic tomography. The mantle is mainly composed of silicates, and the boundaries between layers of the mantle are probably due to phase transitions.

The mantle acts as a solid for seismic waves, but under high pressures and temperatures it deforms so that over millions of years it acts like a liquid. This makes plate tectonics possible. Geodynamics is the study of the fluid flow in the mantle and core.

### The magnetosphere



The solar wind is deflected by the magnetosphere (not to scale)

If a planet's magnetic field is sufficiently strong, its interaction with the solar wind forms a magnetosphere around a planet. Early space probes discovered the gross dimensions of the terrestrial magnetic field, which extends about 10 Earth radii towards the Sun. The solar wind, a stream of charged particles, streams out and around the terrestrial magnetic field, and continues behind the magnetic tail, hundreds of Earth radii downstream. Inside the magnetosphere, there are relatively dense regions of solar wind particles, the Van Allen radiation belts.

# Other fields and related disciplines

## Fields

- Geodesy, measurement of the Earth: GPS, vertical and horizontal motions of the Earth's surface, navigation, the study of the Earth's gravitational field, and the size and form of the Earth
- The study of large-scale motions of the Earth's surface and interior, including:
  - Tectonophysics, the study of the physical processes that cause and result from plate tectonics
  - Geodynamics, the study of modes of transport deformation within the Earth: rock deformation, mantle flow and convection, heat flow, lithosphere dynamics
  - Shallow seismology is used in exploration geophysics (to find oil and gas) and for environmental characterization of the subsurface
- Geomagnetism, the study of the Earth's magnetic field, including its origin, telluric currents driven by the magnetic field, the Van Allen belts, and the interaction between the magnetosphere and the solar wind. This field is associated with paleomagnetism, or the measurement of the orientation of the Earth's magnetic field over the geologic past.
- Mathematical geophysics, The development and applications of mathematical methods and techniques for the solution of geophysical problems.
- Geophysical surveying:
  - Exploration and engineering geophysics, using surface methods to detect or infer the presence and position of concentrations of ore minerals and hydrocarbons
  - Archaeological geophysics, for archaeological imaging or mapping
  - Environmental and Engineering Geophysics, for locating underground storage tanks (USTs) or utilities, Unexploded ordnance (UXO), delineating landfills, locating voids or potential subsidence, finding depth to, P-wave or S-wave velocity of, or rippability of bedrock, or the pathway of groundwater movement

## Related disciplines

- Volcanology, the study of volcanoes, volcanic features (hot springs, geysers, fumaroles), volcanic rock, and heat flow related to volcanoes
- Atmospheric sciences, which includes:
  - Atmospheric electricity and the ionosphere
  - Aeronomy, the study of the physical structure and chemistry of the atmosphere.

- Meteorology and Climatology, which both involve studies of the weather.
- The study of water on the Earth, hydrology, physical oceanography and glaciology
- Geological and geophysical engineering and Engineering geology, applying geophysics to the engineering design of facilities including roads, tunnels, and mines
- The study of the rocks and minerals, including petrophysics and aspects of mineralogy such as physical mineralogy and crystal structure

## **Methods of geophysics**

### **Space probes**

Space probes made it possible to collect data not only the visible light region, but in other areas of the electromagnetic spectrum. The planets can be characterized by their force fields: gravity and their magnetic fields, which are studied through geophysics and space physics.

Measuring the changes in acceleration experienced by spacecraft as they orbit has allowed fine details of the gravity fields of the planets to be mapped. For example, in the 1970s, the gravity field disturbances above lunar maria were measured through lunar orbiters, which lead to the discovery of concentrations of mass, mascons, beneath the Imbrium, Serenitatis, Crisium, Nectaris and Humorum basins.

In 2002, NASA launched the Gravity Recovery and Climate Experiment, wherein two twin satellites map variations in Earth's gravity field by making measurements of the distance between the two satellites using GPS and a microwave ranging system. Gravity variations detected by GRACE include those caused by changes in ocean currents; runoff and ground water depletion; melting ice sheets and glaciers.

# Soil science



**Soil science** is the study of soil as a natural resource on the surface of the earth including soil formation, classification and mapping; physical, chemical, biological, and fertility properties of soils; and these properties in relation to the use and management of soils.

Sometimes terms which refer to branches of soil science, such as pedology (formation, chemistry, morphology and classification of soil) and edaphology (influence of soil on organisms, especially plants), are used as if synonymous with soil science. The diversity of names associated with this discipline is related to the various associations concerned. Indeed, engineers, agronomists, chemists, geologists, geographers, ecologists, biologists, microbiologists, sylviculturists, sanitarians, archaeologists, and specialists in regional planning, all contribute to further knowledge of soils and the advancement of the soil sciences.

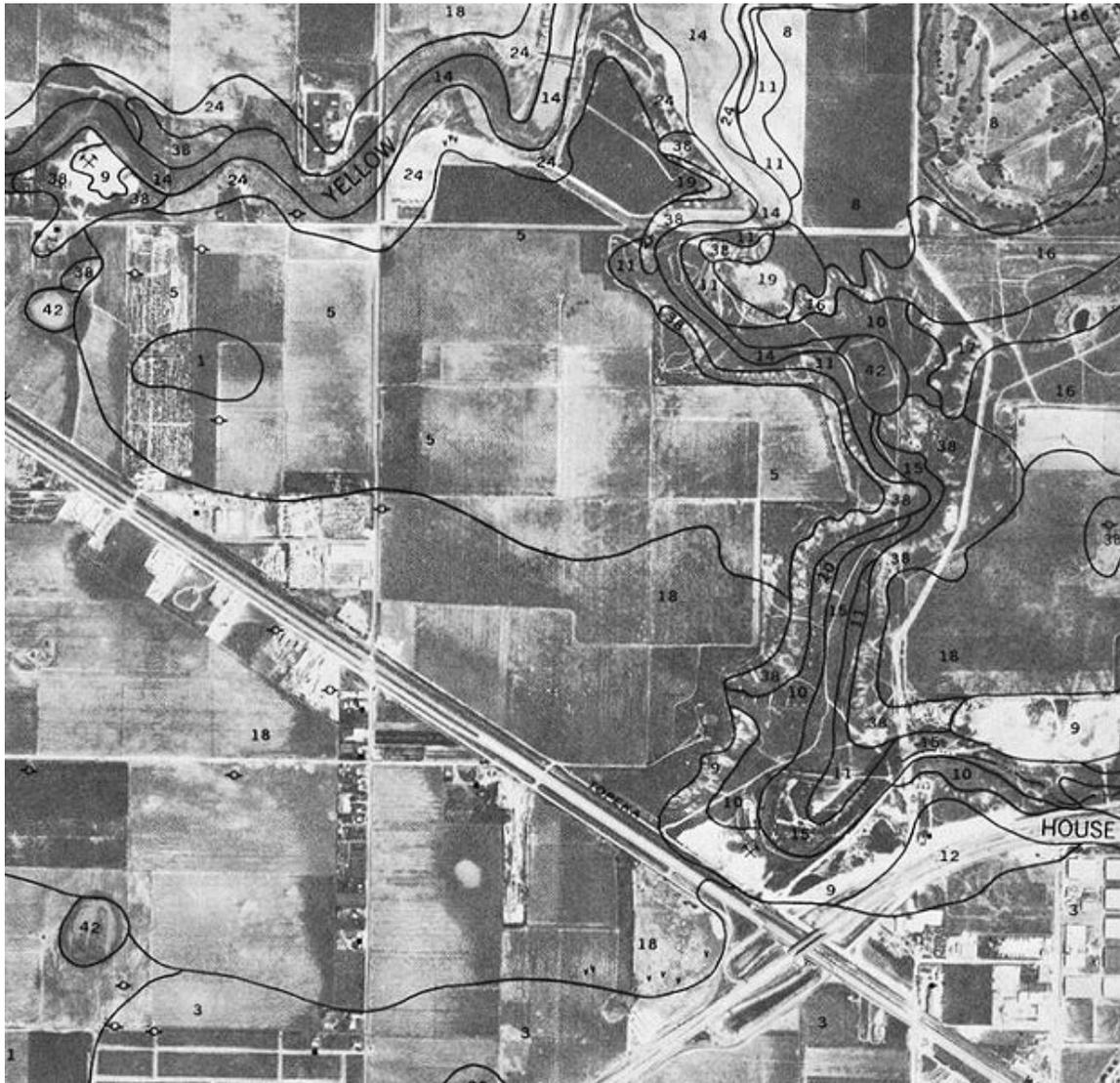
## **Fields of study**

Soil occupies the pedosphere, one of Earth's spheres that the geosciences use to organize the Earth conceptually. This is the conceptual perspective of pedology and edaphology, the two main branches of soil science. Pedology is the study of soil in its natural setting. Edaphology is the study of soil in relation to soil-dependent uses. Both branches apply a combination of soil physics, soil chemistry, and soil biology. Due to the numerous interactions between the biosphere, atmosphere and hydrosphere that are hosted within the pedosphere, more integrated, less soil-centric concepts are also valuable. Many concepts essential to understanding soil come from individuals not identifiable strictly as soil scientists. This highlights the interdisciplinary nature of soil concepts.

## **Research**

Dependence on and curiosity about soil, exploring the diversity and dynamic of this resource continues to yield fresh discoveries and insights. New avenues of soil research are compelled by a need to understand soil in the context of climate change, greenhouse gases, and carbon sequestration. Interest in maintaining the planet's biodiversity and in exploring past cultures has also stimulated renewed interest in achieving a more refined understanding of soil.

# Mapping

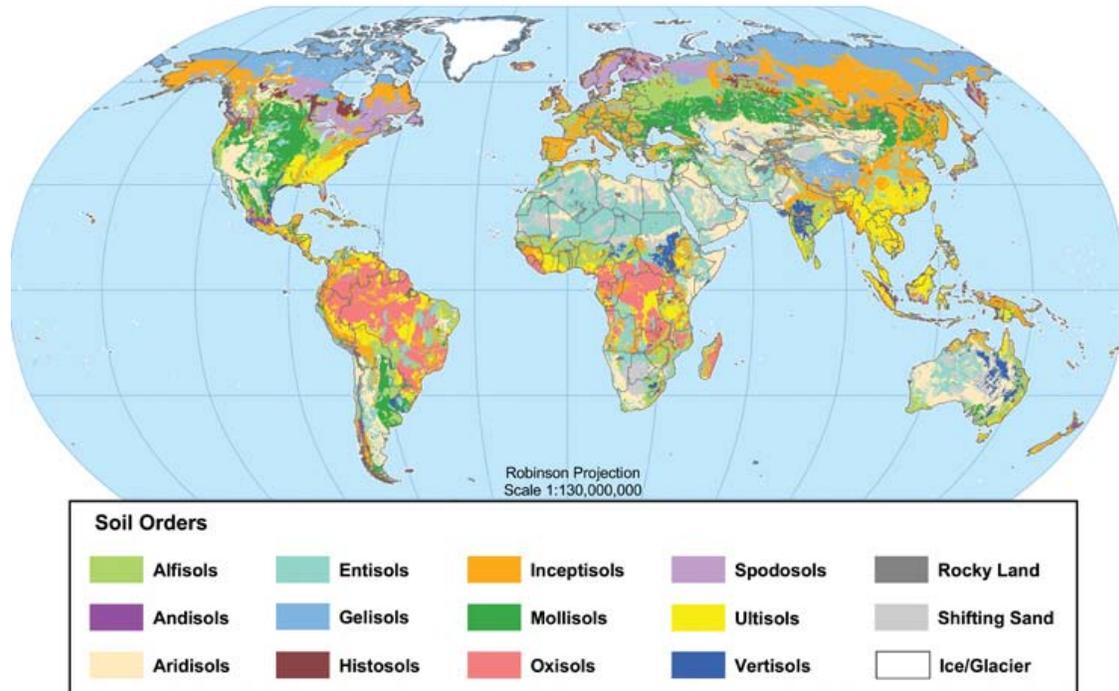


Sample of an aerial photo from a published soil survey

Most empirical knowledge of soil in nature comes from soil survey efforts. Soil survey, or soil mapping, is the process of determining the soil types or other properties of the soil cover over a landscape, and mapping them for others to understand and use. It relies heavily on distinguishing the individual influences of the five classic soil forming factors. This effort draws upon geomorphology, physical geography, and analysis of vegetation and land-use patterns. Primary data for the soil survey are acquired by field sampling and supported by remote sensing.

# Classification

## Global Soil Regions



US Department of Agriculture  
Natural Resources  
Conservation Service

Soil Survey Division  
World Soil Resources  
[soils.usda.gov/use/worldsoils](http://soils.usda.gov/use/worldsoils)

November 2005

Map of global soil regions from the USDA

As of 2006, the World Reference Base for Soil Resources, via its Land & Water Development division, is the pre-eminent soil classification system. It replaces the previous FAO soil classification.

The WRB borrows from modern soil classification concepts, including USDA soil taxonomy. The classification is based mainly on soil morphology as an expression pedogenesis. A major difference with USDA soil taxonomy is that soil climate is not part of the system, except insofar as climate influences soil profile characteristics.

Many other classification schemes exist, including vernacular systems. The structure in vernacular systems are either nominal, giving unique names to soils or landscapes, or descriptive, naming soils by their characteristics such as red, hot, fat, or sandy. Soils are distinguished by obvious characteristics, such as physical appearance (e.g., color, texture, landscape position), performance (e.g., production capability, flooding), and accompanying vegetation. A vernacular distinction familiar to many is classifying texture as heavy or light. Light soil content and better structure, take less effort to turn and

cultivate. Contrary to popular belief, light soils do not weigh less than heavy soils on an air dry basis nor do they have more porosity.

## **History**

Vasily Dokuchaev, a Russian geologist, geographer and early soil scientist, is credited with identifying soil as a resource whose distinctness and complexity deserved to be separated conceptually from geology and crop production and treated as a whole.

Previously, soil had been considered a product of chemical transformations of rocks, a dead substrate from which plants derive nutritious elements. Soil and bedrock were in fact equated. Dokuchaev considers the soil as a natural body having its own genesis and its own history of development, a body with complex and multiform processes taking place within it. The soil is considered as different from bedrock. The latter becomes soil under the influence of a series of soil-formation factors (climate, vegetation, country, relief and age). According to him, soil should be called the “daily” or outward horizons of rocks regardless of the type; they are changed naturally by the common effect of water, air and various kinds of living and dead organisms.

A 1914 encyclopedic definition: “the different forms of earth on the surface of the rocks, formed by the breaking down or weathering of rocks”. serves to illustrate the historic view of soil which persisted from the 19<sup>th</sup> century. Dokuchaev’s late 19<sup>th</sup> century soil concept developed in the 20<sup>th</sup> century to one of soil as earthy material that has been altered by living processes. A corollary concept is that soil without a living component is simply a part of earth’s outer layer.

Further refinement of the soil concept is occurring in view of an appreciation of energy transport and transformation within soil. The term is popularly applied to the material on the surface of the Earth’s moon and Mars, a usage acceptable within a portion of the scientific community. Accurate to this modern understanding of soil is Nikiforoff’s 1959 definition of soil as the “excited skin of the sub aerial part of the earth’s crust”.

## **Areas of practice**

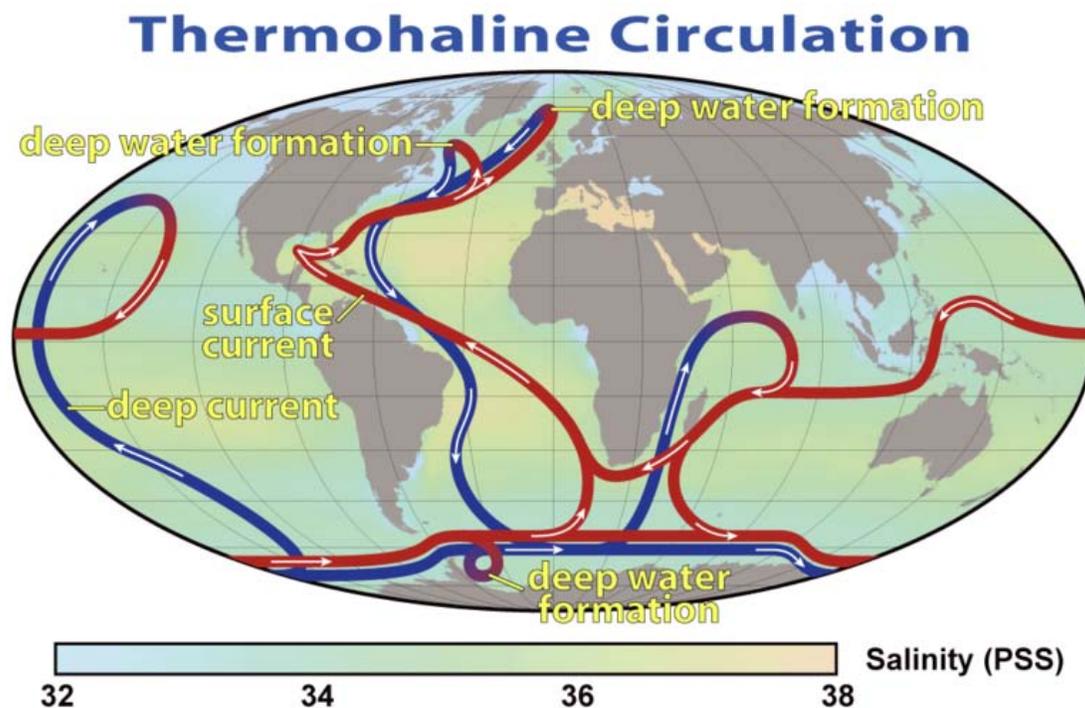
Academically, soil scientists tend to be drawn to one of five areas of specialization: microbiology, pedology, edaphology, physics or chemistry. Yet the work specifics are very much dictated by the challenges facing our civilization’s desire to sustain the land that supports it, and the distinctions between the sub-disciplines of soil science often blur in the process. Soil science professionals commonly stay current in soil chemistry, soil physics, soil microbiology, pedology, and applied soil science in related disciplines

One interesting effort drawing in soil scientists in the USA as of 2004 is the Soil Quality Initiative. Central to the Soil Quality Initiative is developing indices of soil health and then monitoring them in a way that gives us long term (decade-to-decade) feedback on our performance as stewards of the planet. The effort includes understanding the

functions of soil microbiotic crusts and exploring the potential to sequester atmospheric carbon in soil organic matter. The concept of soil quality, however, has not been without its share of controversy and criticism, including critiques by Nobel Laureate Norman Borlaug and World Food Prize Winner Pedro Sanchez.

A more traditional role for soil scientists has been to map soils. Most every area in the United States now has a published soil survey, which includes interpretive tables as to how soil properties support or limit activities and uses. An internationally accepted soil taxonomy allows uniform communication of soil characteristics and functions. National and international soil survey efforts have given the profession unique insights into landscape scale functions. The landscape functions that soil scientists are called upon to address in the field seem to fall roughly into six areas:

## Oceanography



Thermohaline circulation

**Oceanography** (compound of the Greek words *ωκεανός* meaning “ocean” and *γράφω* meaning “to write”), also called **oceanology** or **marine science**, is the branch of Earth science that studies the ocean. It covers a wide range of topics, including marine organisms and ecosystem dynamics; ocean currents, waves, and geophysical fluid

dynamics; plate tectonics and the geology of the sea floor; and fluxes of various chemical substances and physical properties within the ocean and across its boundaries. These diverse topics reflect multiple disciplines that oceanographers blend to further knowledge of the world ocean and understanding of processes within it: biology, chemistry, geology, meteorology, and physics as well as geography.

## History



Map of the Gulf Stream by Benjamin Franklin, 1769-1770. Courtesy of the NOAA Photo Library.

Humans first acquired knowledge of the waves and currents of the seas and oceans in pre-historic times. Observations on tides are recorded by Aristotle and Strabo. Early modern exploration of the oceans was primarily for cartography and mainly limited to its surfaces and of the creatures that fishermen brought up in nets, though depth soundings by lead line were taken.

Although Juan Ponce de León in 1513 first identified the Gulf Stream, and the current was well-known to mariners, Benjamin Franklin made the first scientific study of it and gave it its name. Franklin measured water temperatures during several Atlantic crossings and correctly explained the Gulf Stream's cause. Franklin and Timothy Folger printed the first map of the Gulf Stream in 1769-1770.

When Louis Antoine de Bougainville, who voyaged between 1766 and 1769, and James Cook, who voyaged from 1768 to 1779, carried out their explorations in the South Pacific, information on the oceans themselves formed part of the reports. James Rennell wrote the first scientific textbooks about currents in the Atlantic and Indian oceans during the late 18<sup>th</sup> and at the beginning of 19<sup>th</sup> century. Sir James Clark Ross took the first modern sounding in deep sea in 1840, and Charles Darwin published a paper on reefs and the formation of atolls as a result of the second voyage of HMS *Beagle* in 1831-6. Robert FitzRoy published a report in four volumes of the three voyages of the *Beagle*. In 1841–1842 Edward Forbes undertook dredging in the Aegean Sea that founded marine ecology.

As first superintendent of the United States Naval Observatory (1842–1861) Matthew Fontaine Maury devoted his time to the study of marine meteorology, navigation, and charting prevailing winds and currents. His *Physical Geography of the Sea*, 1855 was the first textbook of oceanography. Many nations sent oceanographic observations to Maury at the Naval Observatory, where he and his colleagues evaluated the information and gave the results worldwide distribution.

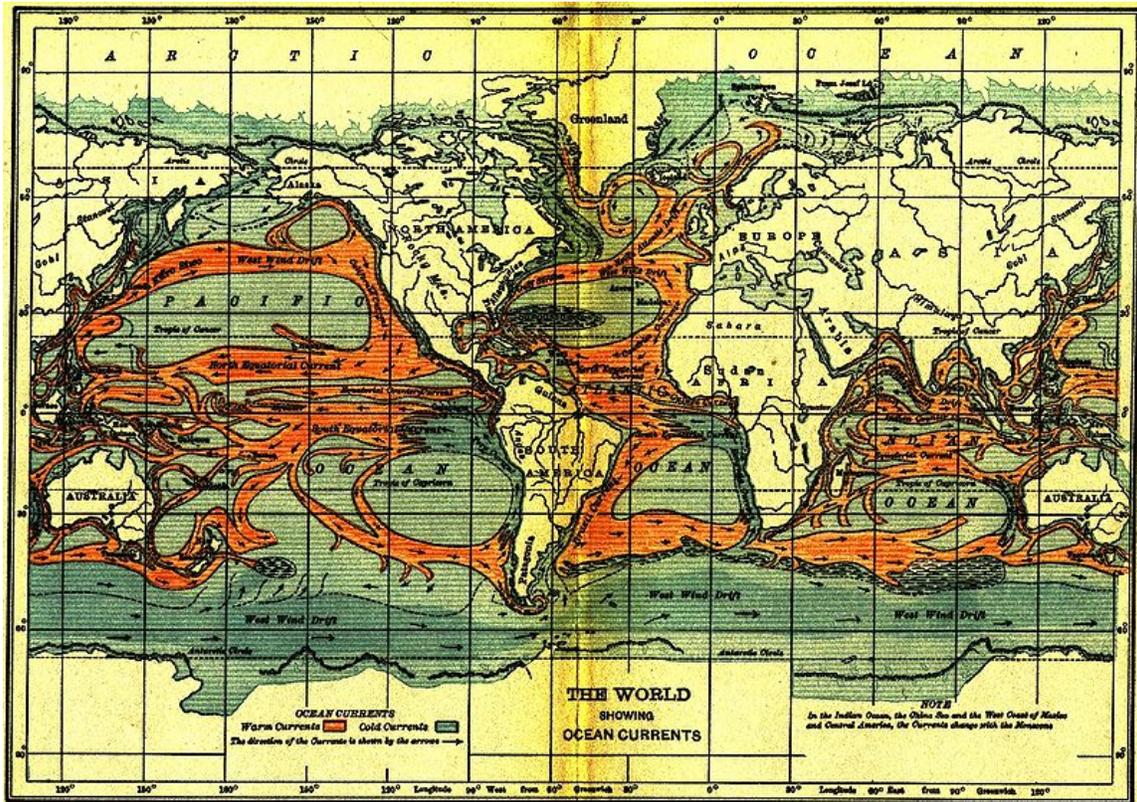
The steep slope beyond the continental shelves was discovered in 1849. The first successful laying of transatlantic telegraph cable in August 1858 confirmed the presence of an underwater “telegraphic plateau” mid-ocean ridge. After the middle of the 19<sup>th</sup> century, scientific societies were processing a flood of new terrestrial botanical and zoological information.

In 1871, under the recommendations of the Royal Society of London, the British government sponsored an expedition to explore world's oceans and conduct scientific investigations. Under that sponsorship the Scots Charles Wyville Thompson and Sir John Murray launched the Challenger expedition (1872–1876). The results of this were published in 50 volumes covering biological, physical and geological aspects. 4417 new species were discovered.

Other European and American nations also sent out scientific expeditions (as did private individuals and institutions). The first purpose built oceanographic ship, the “Albatros” was built in 1882. The four-month 1910 North Atlantic expedition headed by Sir John Murray and Johan Hjort was at that time the most ambitious research oceanographic and marine zoological project ever, and led to the classic 1912 book *The Depths of the Ocean*.

Oceanographic institutes dedicated to the study of oceanography were founded. In the United States, these included the Scripps Institution of Oceanography in 1892, Woods Hole Oceanographic Institution in 1930, Virginia Institute of Marine Science in 1938, Lamont-Doherty Earth Observatory at Columbia University, and the School of

Oceanography at University of Washington. In Britain, there is a major research institution: National Oceanography Centre, Southampton which is the successor to the Institute of Oceanography. In Australia, CSIRO Marine and Atmospheric Research, known as CMAR, is a leading center. In 1921 the International Hydrographic Bureau (IHB) was formed in Monaco.



Ocean currents (1911)

In 1893, Fridtjof Nansen allowed his ship “Fram” to be frozen in the Arctic ice. As a result he was able to obtain oceanographic data as well as meteorological and astronomical data. The first international organization of oceanography was created in 1902 as the International Council for the Exploration of the Sea.

The first acoustic measurement of sea depth was made in 1914. Between 1925 and 1927 the “Meteor” expedition gathered 70,000 ocean depth measurements using an echo sounder, surveying the Mid Atlantic ridge. The Great Global Rift, running along the Mid Atlantic Ridge, was discovered by Maurice Ewing and Bruce Heezen in 1953 while the mountain range under the Arctic was found in 1954 by the Arctic Institute of the USSR. The theory of seafloor spreading was developed in 1960 by Harry Hammond Hess. The Ocean Drilling Project started in 1966. Deep sea vents were discovered in 1977 by John Corlis and Robert Ballard in the submersible “Alvin”.

In the 1950s, Auguste Piccard invented the bathyscaphe and used the “Trieste” to investigate the ocean’s depths. The nuclear submarine Nautilus made the first journey under the ice to the North Pole in 1958. In 1962 there was the first deployment of FLIP (Floating Instrument Platform), a 355 foot spar buoy.

Then, in 1966, the U.S. Congress created a *National Council for Marine Resources and Engineering Development*. NOAA was put in charge of exploring and studying all aspects of Oceanography in the USA. It also enabled the National Science Foundation to award *Sea Grant College* funding to multi-disciplinary researchers in the field of oceanography.

From the 1970s, there has been much emphasis on the application of large scale computers to oceanography to allow numerical predictions of ocean conditions and as a part of overall environmental change prediction. An oceanographic buoy array was established in the Pacific to allow prediction of El Niño events.

1990 saw the start of the World Ocean Circulation Experiment (WOCE) which continued until 2002. Geosat seafloor mapping data became available in 1995.

In 1942, Sverdrup and Fleming published “The Ocean” which was a major landmark. “The Sea” (in three volumes covering physical oceanography, seawater and geology) edited by M.N. Hill was published in 1962 while the “Encyclopedia of Oceanography” by Rhodes Fairbridge was published in 1966.

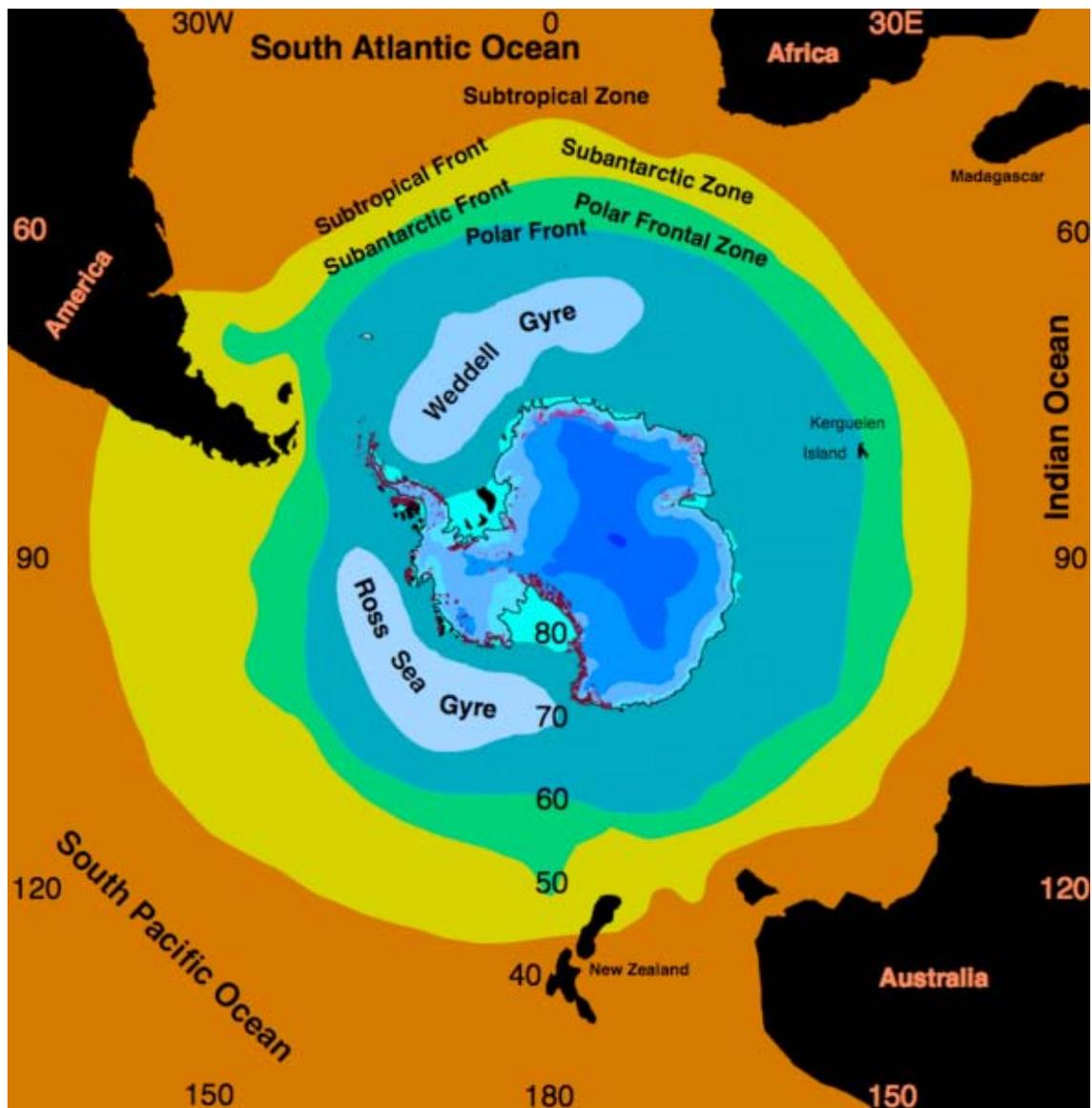
## **Connection to the atmosphere**

The study of the oceans is linked to understanding global climate changes, potential global warming and related biosphere concerns. The atmosphere and ocean are linked because of evaporation and precipitation as well as thermal flux (and solar insolation). Wind stress is a major driver of ocean currents while the ocean is a sink for atmospheric carbon dioxide.

Our planet is invested with two great oceans; one visible, the other invisible; one underfoot, the other overhead; one entirely envelopes it, the other covers about two thirds of its surface.

—Matthew F. Maury, *The Physical Geography of the Seas and Its Meteorology* (1855)

## Branches



Oceanographic frontal systems on the Southern Hemisphere

The study of oceanography is divided into branches:

- **Biological oceanography**, or **marine biology**, is the study of the plants, animals and microbes of the oceans and their ecological interaction with the ocean;
- **Chemical oceanography**, or **marine chemistry**, is the study of the chemistry of the ocean and its chemical interaction with the atmosphere;
- **Geological oceanography**, or **marine geology**, is the study of the geology of the ocean floor including plate tectonics;

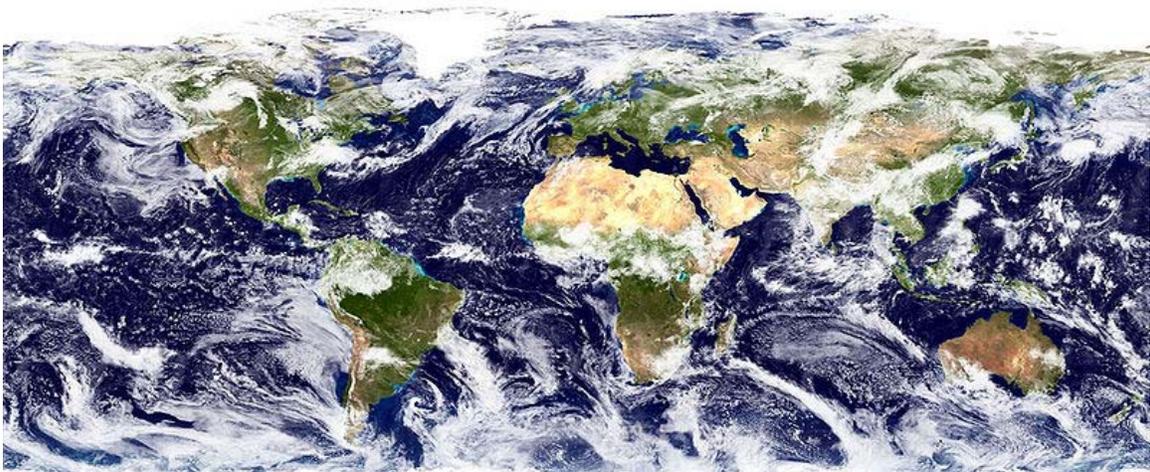
- **Physical oceanography**, or **marine physics**, studies the ocean's physical attributes including temperature-salinity structure, mixing, waves, internal waves, surface tides, internal tides, and currents. Of particular interest is the behavior of sound (acoustical oceanography), light (optical oceanography) and radio waves in the ocean.

These branches reflect the fact that many oceanographers are first trained in the exact sciences or mathematics and then focus on applying their interdisciplinary knowledge, skills and abilities to oceanography.

Data derived from the work of Oceanographers is used in **marine engineering**, in the design and building of oil platforms, ships, harbours, and other structures that allow us to use the ocean safely.

Oceanographic data management is the discipline ensuring that oceanographic data both past and present are available to researchers.

## Hydrology



Water covers 70% of the Earth's surface.

**Hydrology** is the study of the movement, distribution, and quality of water throughout the Earth, including the hydrologic cycle, water resources and environmental watershed sustainability. A practitioner of hydrology is a hydrologist, working within the fields of either earth or environmental science, physical geography, geology or civil and environmental engineering.

Domains of hydrology include hydrometeorology, surface hydrology, hydrogeology, drainage basin management and water chemistry, where water plays the central role.

Oceanography and meteorology are not included because water is only one of many important aspects.

Hydrological research can inform environmental engineering, policy and planning.

The term *hydrology* is from Greek: ὕδωρ, *hydōr*, “water”; and λόγος, *logos*, “study”.

## History of hydrology

Hydrology has been a subject of investigation and engineering for millennia. For example, about 4000 B.C. the Nile was dammed to improve agricultural productivity of previously barren lands. Mesopotamian towns were protected from flooding with high earthen walls. Aqueducts were built by the Greeks and Ancient Romans, while the History of China shows they built irrigation and flood control works. The ancient Sinhalese used hydrology to build complex irrigation Works in Sri Lanka, also known for invention of the Valve Pit which allowed construction of large reservoirs, anicuts and canals which still function.

Marcus Vitruvius, in the first century B.C., described a philosophical theory of the hydrologic cycle, in which precipitation falling in the mountains infiltrated the Earth’s surface and led to streams and springs in the lowlands. With adoption of a more scientific approach, Leonardo da Vinci and Bernard Palissy independently reached an accurate representation of the hydrologic cycle. It was not until the 17<sup>th</sup> century that hydrologic variables began to be quantified.

Pioneers of the modern science of hydrology include Pierre Perrault, Edme Mariotte and Edmund Halley. By measuring rainfall, runoff, and drainage area, Perrault showed that rainfall was sufficient to account for flow of the Seine. Marriotte combined velocity and river cross-section measurements to obtain discharge, again in the Seine. Halley showed that the evaporation from the Mediterranean Sea was sufficient to account for the outflow of rivers flowing into the sea.

Advances in the 18<sup>th</sup> century included the Bernoulli piezometer and Bernoulli’s equation, by Daniel Bernoulli, the Pitot tube. The 19<sup>th</sup> century saw development in groundwater hydrology, including Darcy’s law, the Dupuit-Thiem well formula, and Hagen-Poiseuille’s capillary flow equation.

Rational analyses began to replace empiricism in the 20<sup>th</sup> century, while governmental agencies began their own hydrological research programs. Of particular importance were Leroy Sherman’s unit hydrograph, the infiltration theory of Robert E. Horton, and C.V. Theis’s Aquifer test/equation describing well hydraulics.

Since the 1950s, hydrology has been approached with a more theoretical basis than in the past, facilitated by advances in the physical understanding of hydrological processes and by the advent of computers and especially Geographic Information Systems (GIS).

## Hydrologic cycle

The central theme of hydrology is that water circulates throughout the Earth through different pathways and at different rates. The most vivid image of this is in the evaporation of water from the ocean, which forms clouds. These clouds drift over the land and produce rain. The rainwater flows into lakes, rivers, or aquifers. The water in lakes, rivers, and aquifers then either evaporates back to the atmosphere or eventually flows back to the ocean, completing a cycle. Water changes its state of being several times throughout this cycle

## Branches of hydrology

Chemical hydrology is the study of the chemical characteristics of water.

Ecohydrology is the study of interactions between organisms and the hydrologic cycle.

Hydrogeology is the study of the presence and movement of ground water.

Hydroinformatics is the adaptation of information technology to hydrology and water resources applications.

Hydrometeorology is the study of the transfer of water and energy between land and water body surfaces and the lower atmosphere.

Isotope hydrology is the study of the isotopic signatures of water.

Surface hydrology is the study of hydrologic processes that operate at or near Earth's surface.

## Hydrologic measurements

Measurement is fundamental for assessing water resources and understanding the processes involved in the hydrologic cycle. Because the hydrologic cycle is so diverse, hydrologic measurement methods span many disciplines: including soils, oceanography, atmospheric science, geology, geophysics and limnology, to name a few. Here, hydrologic measurement methods are organized by hydrologic sub-disciplines. Each of these subdisciplines is addressed briefly with a practical discussion of the methods used to date and a bibliography of background information.

Quantifying groundwater flow and transport

- Aquifer characterization
  - Flow direction
    - Piezometer – groundwater pressure and, by inference, groundwater depth
    - Conductivity, storativity, transmissivity
    - Geophysical methods
- Vadose zone characterization
  - Infiltration
    - Infiltrometer – infiltration

- Soil moisture
  - Capacitance probe-soil moisture
  - Time domain reflectometer – soil moisture
  - Tensiometer – soil moisture
  - Solute sampling
  - Geophysical methods

#### Quantifying surface water flow and transport

- Direct and indirect discharge measurements
  - Stream gauge – stream flow
  - Tracer techniques
  - Chemical transport
  - Sediment transport and erosion
  - Stream-aquifer exchange

#### Quantifying exchanges at the land-atmospheric boundary

- Precipitation
  - Bulk rain events
    - Disdrometer – precipitation characteristics
    - Radar – cloud properties, rain rate estimation, hail and snow detection
    - Rain gauge – rain and snowfall
    - Satellite – rainy area identification, rain rate estimation, land-cover/land-use, soil moisture
    - Sling psychrometer – humidity
  - Snow, hail and ice
  - Dew, mist and fog
- Evaporation
  - from water surfaces
  - Evaporation – Symon’s evaporation pan
  - from plant surfaces
  - through the boundary layer
- Transpiration
  - Natural ecosystems
  - Agronomic ecosystems
- Momentum
- Heat flux
  - Energy budgets

#### Uncertainty analyses

#### Remote sensing of hydrologic processes

- Land based sensors
- Airborne Sensors

- Satellite sensors

#### Water quality

- Sample collection
- In-situ methods
- Physical measurements (includes sediment concentration)
- Collection of samples to quantify Organic Compounds
- Collection of samples to quantify Inorganic Compounds
- Analysis of aqueous Organic Compounds
- Analysis of aqueous Inorganic Compounds
- Microbiological sampling and analysis

#### Integrating measurement and modeling

- Budget analyses
- Parameter estimation
- Scaling in time and space
- Data assimilation
- Quality control of data

## Hydrologic prediction

Observations of hydrologic processes are used to make predictions of the future behaviour of hydrologic systems (water flow, water quality). One of the major current concerns in hydrologic research is the Prediction in Ungauged Basins (PUB), i.e. in basins where no or only very few data exist.

### Statistical hydrology

By criticizing the statistical properties of hydrologic records, such as rainfall or river flow, hydrologists can estimate future hydrologic phenomena, assuming the characteristics of the processes remain unchanged.

These estimates are important for engineers and economists so that proper risk analysis can be performed to influence investment decisions in future infrastructure and to determine the yield reliability characteristics of water supply systems. Statistical information is criticized to formulate operating rules for large dams forming part of systems which include agricultural, industrial and residential demands.

### Hydrologic modeling

Hydrologic models are simplified, conceptual representations of a part of the hydrologic cycle. They are primarily used for hydrologic prediction and for understanding hydrologic processes. Two major types of hydrologic models can be distinguished:

- Models based on data. These models are black box systems, using mathematical and statistical concepts to link a certain input (for instance rainfall) to the model output (for instance runoff). Commonly used techniques are regression, transfer functions, and system identification. The simplest of these models may be linear models, but it is common to deploy non-linear components to represent some general aspects of a catchment's response without going deeply into the real physical processes involved. An example of such an aspect is the well-known behaviour that a catchment will respond much more quickly and strongly when it is already wet than when it is dry..
- Models based on process descriptions. These models try to represent the physical processes observed in the real world. Typically, such models contain representations of surface runoff, subsurface flow, evapotranspiration, and channel flow, but they can be far more complicated. These models are known as deterministic hydrology models. Deterministic hydrology models can be subdivided into single-event models and continuous simulation models.

Recent research in hydrologic modeling tries to have a more global approach to the understanding of the behaviour of hydrologic systems to make better predictions and to face the major challenges in water resources management.

## **Hydrologic transport**

Water movement is a significant means by which other material, such as soil or pollutants, are transported from place to place. Initial input to receiving waters may arise from a point source discharge or a line source or area source, such as surface runoff. Since the 1960s rather complex mathematical models have been developed, facilitated by the availability of high speed computers. The most common pollutant classes analyzed are nutrients, pesticides, total dissolved solids and sediment.

## **Applications of hydrology**

- Determining the water balance of a region.
- Determining the agricultural water balance.
- Designing riparian restoration projects.
- Mitigating and predicting flood, landslide and drought risk.
- Real-time flood forecasting and flood warning.
- Designing irrigation schemes and managing agricultural productivity.
- Part of the hazard module in catastrophe modeling.
- Providing drinking water.
- Designing dams for water supply or hydroelectric power generation.
- Designing bridges.
- Designing sewers and urban drainage system.
- Analyzing the impacts of antecedent moisture on sanitary sewer systems.
- Predicting geomorphological changes, such as erosion or sedimentation.

- Assessing the impacts of natural and anthropogenic environmental change on water resources.
- Assessing contaminant transport risk and establishing environmental policy guidelines.

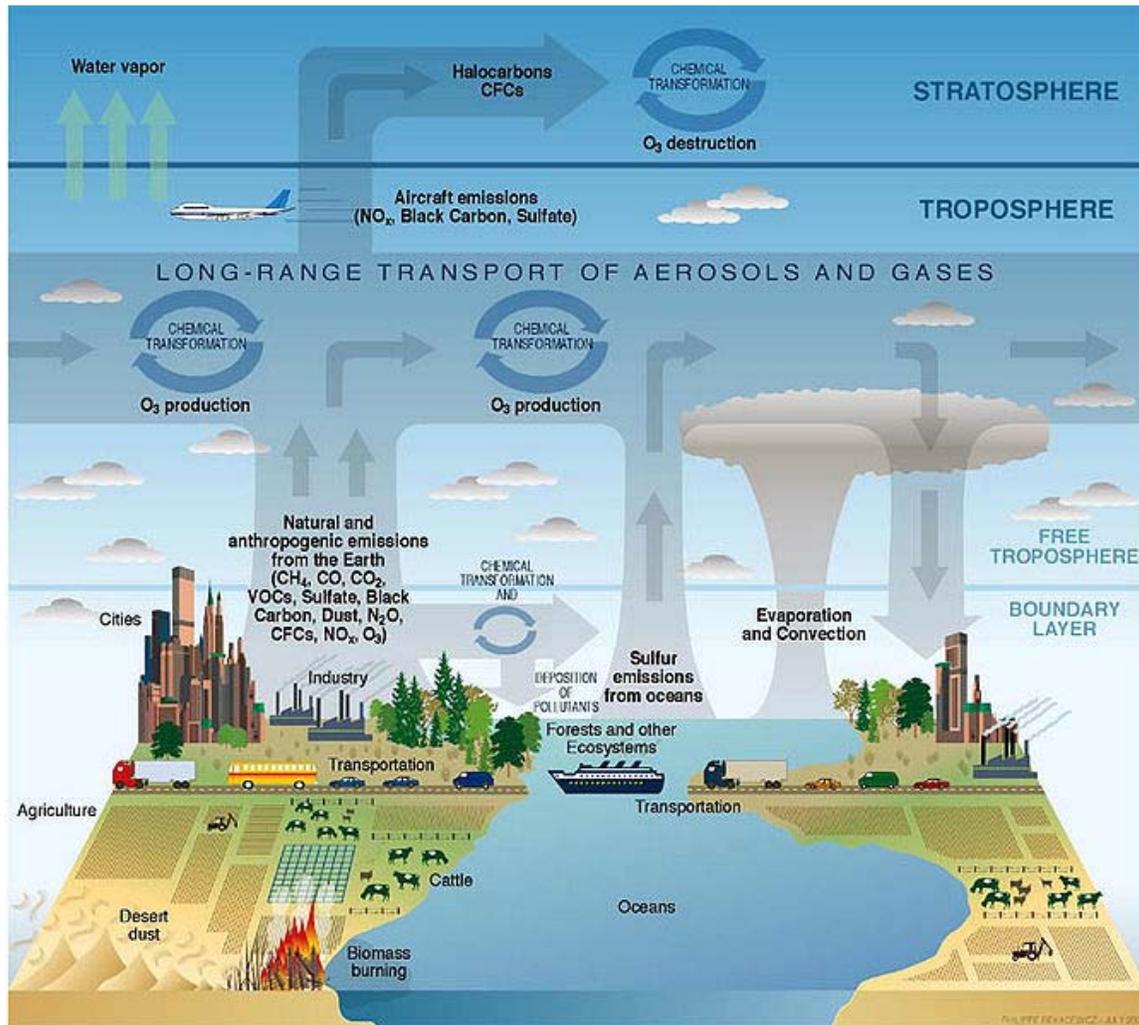
## **Atmospheric sciences**

*Atmospheric sciences* is an umbrella term for the study of the atmosphere, its processes, the effects other systems have on the atmosphere, and the effects of the atmosphere on these other systems. Meteorology includes atmospheric chemistry and atmospheric physics with a major focus on weather forecasting. Climatology is the study of atmospheric changes (both long and short-term) that define average climates and their change over time, due to both natural and anthropogenic climate variability. Aeronomy is the study of the upper layers of the atmosphere, where dissociation and ionization are important. Atmospheric science has been extended to the field of planetary science and the study of the atmospheres of the planets of the solar system.

Experimental instruments used in atmospheric sciences include satellites, rocketsondes, radiosondes, weather balloons, and lasers.

The term *aerology* is sometimes used as an alternative term for the study of Earth's atmosphere.

# Atmospheric chemistry



Atmospheric chemistry is a branch of atmospheric science in which the chemistry of the Earth's atmosphere and that of other planets is studied. It is a multidisciplinary field of research and draws on environmental chemistry, physics, meteorology, computer modeling, oceanography, geology and volcanology and other disciplines. Research is increasingly connected with other areas of study such as climatology.

The composition and chemistry of the atmosphere is of importance for several reasons, but primarily because of the interactions between the atmosphere and living organisms. The composition of the Earth's atmosphere has been changed by human activity and some of these changes are harmful to human health, crops and ecosystems. Examples of problems which have been addressed by atmospheric chemistry include acid rain, photochemical smog and global warming. Atmospheric chemistry seeks to understand the causes of these problems, and by obtaining a theoretical understanding of them, allow possible solutions to be tested and the effects of changes in government policy evaluated.

## **Atmospheric dynamics**

Atmospheric dynamics involves the study of observations and theory dealing with all motion systems of meteorological importance. The list includes diverse phenomena as thunderstorms, tornadoes, gravity waves, tropical cyclones, extratropical cyclones, jet streams, and global-scale circulations. The goal of dynamical studies is to explain the observed circulations on the basis of fundamental principles from physics. The objectives of such studies include improving weather forecasting, developing methods for predicting seasonal and interannual climate fluctuations, and understanding the implications of human-induced perturbations (e.g., increased carbon dioxide concentrations or depletion of the ozone layer) on the global climate.

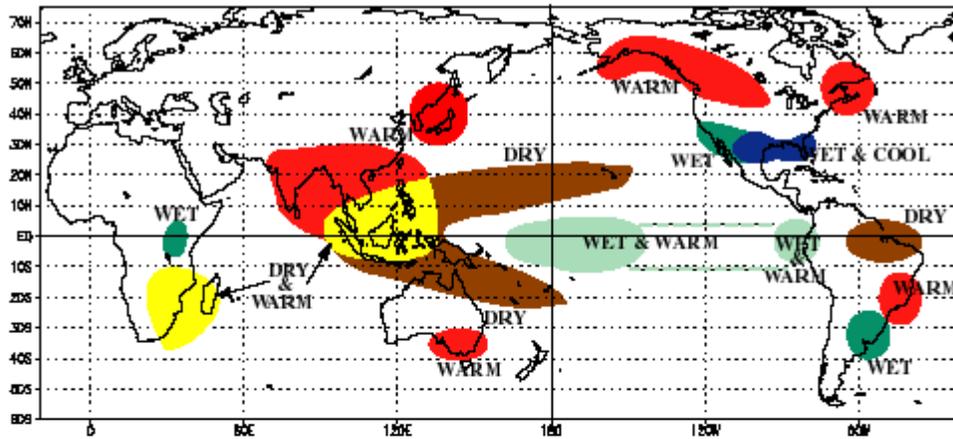
## **Atmospheric physics**

Atmospheric physics is the application of physics to the study of the atmosphere. Atmospheric physicists attempt to model Earth's atmosphere and the atmospheres of the other planets using fluid flow equations, chemical models, radiation balancing, and energy transfer processes in the atmosphere and underlying oceans. In order to model weather systems, atmospheric physicists employ elements of scattering theory, wave propagation models, cloud physics, statistical mechanics and spatial statistics which are highly mathematical and related to physics. It has close links to meteorology and climatology and also covers the design and construction of instruments for studying the atmosphere and the interpretation of the data they provide, including remote sensing instruments.

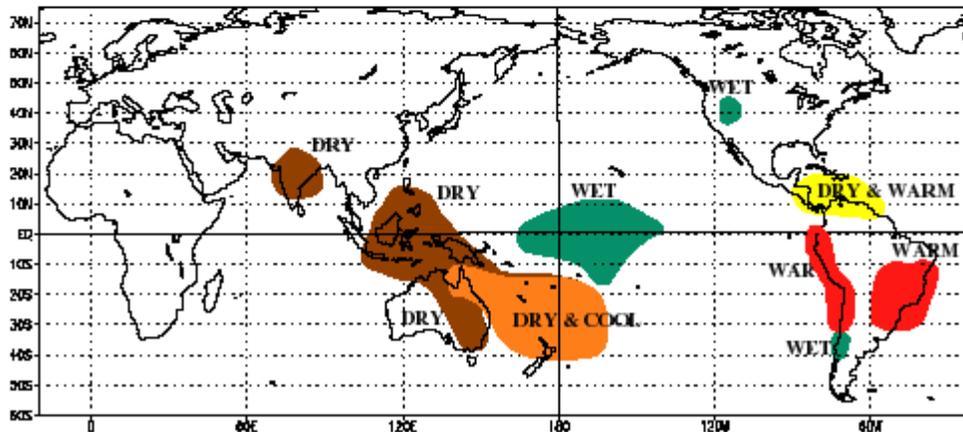
In the UK, atmospheric studies are underpinned by the Meteorological Office. Divisions of the U.S. National Oceanic and Atmospheric Administration (NOAA) oversee research projects and weather modeling involving atmospheric physics. The U.S. National Astronomy and Ionosphere Center also carries out studies of the high atmosphere.

# Climatology

## WARM EPISODE RELATIONSHIPS DECEMBER - FEBRUARY



## WARM EPISODE RELATIONSHIPS JUNE - AUGUST



Regional impacts of warm ENSO episodes (El Niño).

In contrast to meteorology, which studies short term weather systems lasting up to a few weeks, climatology studies the frequency and trends of those systems. It studies the periodicity of weather events over years to millennia, as well as changes in long-term average weather patterns, in relation to atmospheric conditions. **Climatologists**, those who practice climatology, study both the nature of climates – local, regional or global – and the natural or human-induced factors that cause climates to change. Climatology considers the past and can help predict future climate change.

Phenomena of climatological interest include the atmospheric boundary layer, circulation patterns, heat transfer (radiative, convective and latent), interactions between the atmosphere and the oceans and land surface (particularly vegetation, land use and topography), and the chemical and physical composition of the atmosphere. Related

disciplines include astrophysics, atmospheric physics, chemistry, ecology, geology, geophysics, glaciology, hydrology, oceanography, and volcanology.

## **Atmospheres on other planets**



Earth's atmosphere

All of the Solar System planets have atmospheres as their large masses mean gravity is strong enough to keep gaseous particles close to the surface. The larger gas giants are massive enough to keep large amounts of the light gases hydrogen and helium close by, while the smaller planets lose these gases into space. The composition of the Earth's atmosphere is different from the other planets because the various life processes that have transpired on the planet have introduced free molecular oxygen. The only solar planet without a true atmosphere is Mercury which had it mostly, although not entirely, blasted away by the solar wind. The only moon that has retained a dense atmosphere is Titan. There is a thin atmosphere on Triton, and a trace of an atmosphere on the Moon.

Planetary atmospheres are affected by the varying degrees of energy received from either the Sun or their interiors, leading to the formation of dynamic weather systems such as hurricanes, (on Earth), planet-wide dust storms (on Mars), an Earth-sized anticyclone on Jupiter (called the Great Red Spot), and holes in the atmosphere (on Neptune). At least one extrasolar planet, HD 189733 b, has been claimed to possess such a weather system, similar to the Great Red Spot but twice as large.

Hot Jupiters have been shown to be losing their atmospheres into space due to stellar radiation, much like the tails of comets. These planets may have vast differences in temperature between their day and night sides which produce supersonic winds, although the day and night sides of HD 189733b appear to have very similar temperatures, indicating that that planet's atmosphere effectively redistributes the star's energy around the planet.

## Meteorology

**Meteorology** is the interdisciplinary scientific study of the atmosphere that focuses on weather processes and short term forecasting (in contrast with climatology). Studies in the field stretch back millennia, though significant progress in meteorology did not occur until the eighteenth century. The nineteenth century saw breakthroughs occur after observing networks developed across several countries. Breakthroughs in weather forecasting were achieved in the latter half of the twentieth century, after the development of the computer.

Meteorological phenomena are observable weather events which illuminate and are explained by the science of meteorology. Those events are bound by the variables that exist in Earth's atmosphere: They are temperature, air pressure, water vapor, and the gradients and interactions of each variable, and how they change in time. The majority of Earth's observed weather is located in the troposphere. Different spatial scales are studied to determine how systems on local, region, and global levels impact weather and climatology. Meteorology, climatology, atmospheric physics, and atmospheric chemistry are sub-disciplines of the atmospheric sciences. Meteorology and hydrology compose the interdisciplinary field of hydrometeorology. Interactions between Earth's atmosphere and the oceans are part of coupled ocean-atmosphere studies. Meteorology has application in many diverse fields such as the military, energy production, transport, agriculture and construction.

The word "meteorology" is from Greek μετέωρος, *metéōros*, "high in the sky"; and –λογία, *-logia*.

## History

In 350 BC, Aristotle wrote *Meteorology*. Aristotle is considered the founder of meteorology. One of the most impressive achievements described in the *Meteorology* is the description of what is now known as the hydrologic cycle. The Greek scientist Theophrastus compiled a book on weather forecasting, called the *Book of Signs*. The work of Theophrastus remained a dominant influence in the study of weather and in weather forecasting for nearly 2,000 years. In 25 AD, Pomponius Mela, a geographer for the Roman Empire, formalized the climatic zone system. Around the 9<sup>th</sup> century, Al-Dinawari, a Kurdish naturalist, writes the *Kitab al-Nabat (Book of Plants)*, in which he deals with the application of meteorology to agriculture during the Muslim Agricultural Revolution. He describes the meteorological character of the sky, the planets and

constellations, the sun and moon, the lunar phases indicating seasons and rain, the *anwa* (heavenly bodies of rain), and atmospheric phenomena such as winds, thunder, lightning, snow, floods, valleys, rivers, lakes, wells and other sources of water.

### Research of visual atmospheric phenomena



Twilight at Baker Beach

In 1021, Ibn al-Haytham (Alhazen) wrote on the atmospheric refraction of light. He showed that the twilight is due to atmospheric refraction and only begins when the Sun is 19 degrees below the horizon, and uses a complex geometric demonstration to measure the height of the Earth's atmosphere as 52,000 *passuum* (49 miles (79 km)), which is very close to the modern measurement of 50 miles (80 km). He also realized that the atmosphere also reflects light, from his observations of the sky brightening even before the Sun rises.

St. Albert the Great was the first to propose that each drop of falling rain had the form of a small sphere, and that this form meant that the rainbow was produced by light interacting with each raindrop. Roger Bacon was the first to calculate the angular size of the rainbow. He stated that the rainbow summit can not appear higher than 42 degrees above the horizon. In the late 13<sup>th</sup> century and early 14<sup>th</sup> century, Theodoric of Freiberg and Kamāl al-Dīn al-Fārisī continued the work of Ibn al-Haytham, and they were the first to give the correct explanations for the primary rainbow phenomenon. Nevertheless, Theodoric went further and also explained the secondary rainbow. In 1716, Edmund

Halley suggests that aurorae are caused by “magnetic effluvia” moving along the Earth’s magnetic field lines.

### Instruments and classification scales



A hemispherical cup anemometer

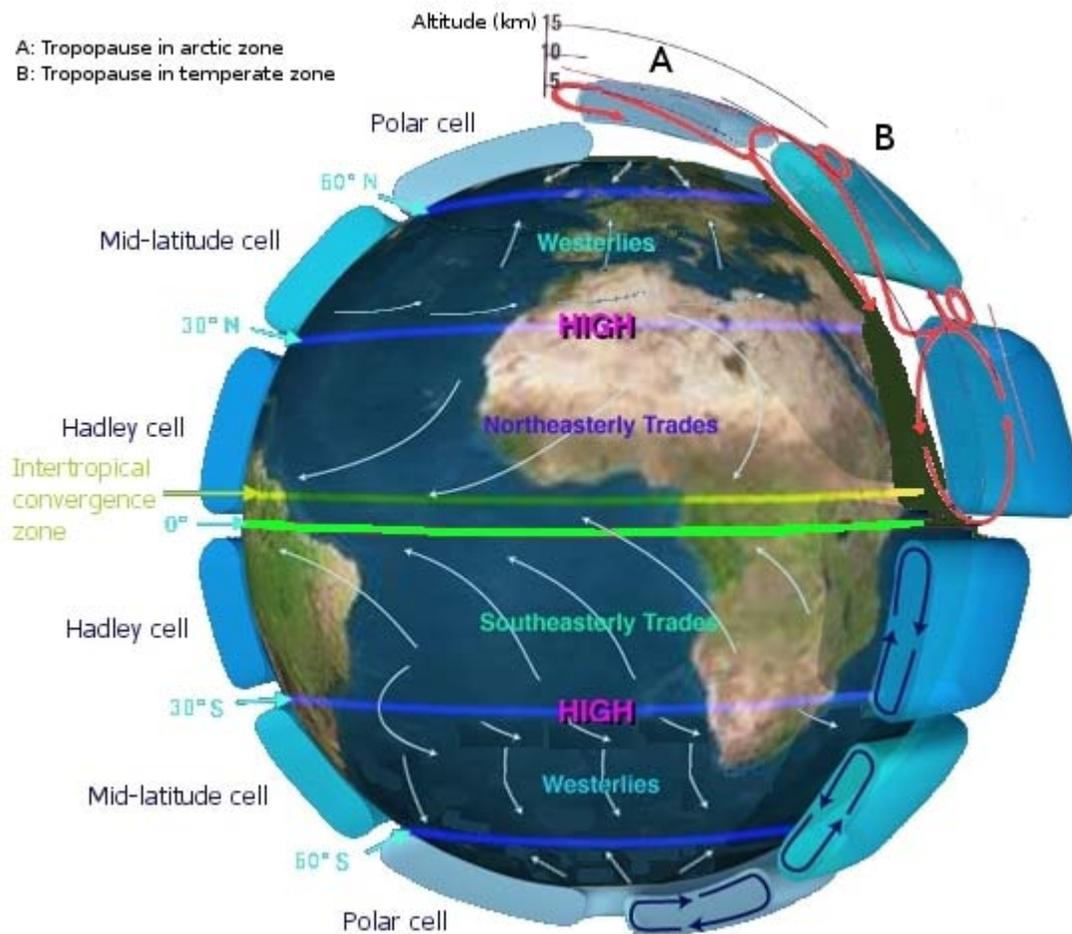
In 1441, King Sejong's son, Prince Munjong, invented the first standardized rain gauge. These were sent throughout the Joseon Dynasty of Korea as an official tool to assess land taxes based upon a farmer's potential harvest. In 1450, Leone Battista Alberti developed a swinging-plate anemometer, and is known as the first *anemometer*. In 1607, Galileo Galilei constructs a thermoscope. In 1611, Johannes Kepler writes the first scientific treatise on snow crystals: “Strena Seu de Nive Sexangula (A New Year’s Gift of Hexagonal Snow)”. In 1643, Evangelista Torricelli invents the mercury barometer. In

1662, Sir Christopher Wren invented the mechanical, self-emptying, tipping bucket rain gauge. In 1714, Gabriel Fahrenheit creates a reliable scale for measuring temperature with a mercury-type thermometer. In 1742, Anders Celsius, a Swedish astronomer, proposed the 'centigrade' temperature scale, the predecessor of the current Celsius scale. In 1783, the first hair hygrometer is demonstrated by Horace-Bénédict de Saussure. In 1802-1803, Luke Howard writes *On the Modification of Clouds* in which he assigns cloud types Latin names. In 1806, Francis Beaufort introduced his system for classifying wind speeds. Near the end of the 19<sup>th</sup> century the first cloud atlases were published, including the *International Cloud Atlas*, which has remained in print ever since. The April 1960 launch of the first successful weather satellite, TIROS-1, marked the beginning of the age where weather information became available globally.

### **Atmospheric composition research**

In 1648, Blaise Pascal rediscovers that atmospheric pressure decreases with height, and deduces that there is a vacuum above the atmosphere. In 1738, Daniel Bernoulli publishes *Hydrodynamics*, initiating the kinetic theory of gases and established the basic laws for the theory of gases. In 1761, Joseph Black discovers that ice absorbs heat without changing its temperature when melting. In 1772, Black's student Daniel Rutherford discovers nitrogen, which he calls *phlogisticated air*, and together they developed the phlogiston theory. In 1777, Antoine Lavoisier discovers oxygen and develops an explanation for combustion. In 1783, in Lavoisier's book *Reflexions sur le phlogistique*, he deprecates the phlogiston theory and proposes a caloric theory. In 1804, Sir John Leslie observes that a matte black surface radiates heat more effectively than a polished surface, suggesting the importance of black body radiation. In 1808, John Dalton defends caloric theory in *A New System of Chemistry* and describes how it combines with matter, especially gases; he proposes that the heat capacity of gases varies inversely with atomic weight. In 1824, Sadi Carnot analyzes the efficiency of steam engines using caloric theory; he develops the notion of a reversible process and, in postulating that no such thing exists in nature, lays the foundation for the second law of thermodynamics.

## Research into cyclones and air flow

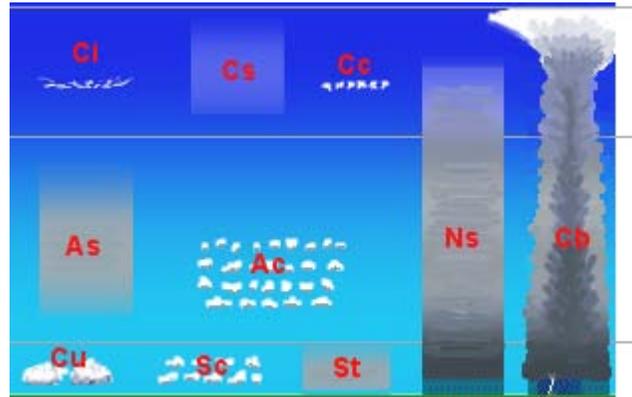


The westerlies and trade winds are part of the Earth's atmospheric circulation

In 1494, Christopher Columbus experiences a tropical cyclone, leads to the first written European account of a hurricane. In 1686, Edmund Halley presents a systematic study of the trade winds and monsoons and identifies solar heating as the cause of atmospheric motions. In 1735, an *ideal* explanation of global circulation through study of the Trade winds was written by George Hadley. In 1743, when Benjamin Franklin is prevented from seeing a lunar eclipse by a hurricane, he decides that cyclones move in a contrary manner to the winds at their periphery. Understanding the kinematics of how exactly the rotation of the Earth affects airflow was partial at first. Gaspard-Gustave Coriolis published a paper in 1835 on the energy yield of machines with rotating parts, such as waterwheels. In 1856, William Ferrel proposed the existence of a circulation cell in the mid-latitudes with air being deflected by the Coriolis force to create the prevailing westerly winds. Late in the 19<sup>th</sup> century the full extent of the large scale interaction of pressure gradient force and deflecting force that in the end causes air masses to move *along* isobars was understood. By 1912, this deflecting force was named the Coriolis effect. Just after World War II, a group of meteorologists in Norway led by Vilhelm Bjerknes developed the Norwegian cyclone model that explains the generation,

intensification and ultimate decay (the life cycle) of mid-latitude cyclones, introducing the idea of fronts, that is, sharply defined boundaries between air masses. The group included Carl-Gustaf Rossby (who was the first to explain the large scale atmospheric flow in terms of fluid dynamics), Tor Bergeron (who first determined the mechanism by which rain forms) and Jacob Bjerknes.

## Observation networks and weather forecasting



Cloud classification by altitude of occurrence

In 1654, Ferdinando II de Medici establishes the first *weather observing* network, that consisted of meteorological stations in Florence, Cutigliano, Vallombrosa, Bologna, Parma, Milan, Innsbruck, Osnabrück, Paris and Warsaw. Collected data was centrally sent to Florence at regular time intervals. In 1832, an electromagnetic telegraph was created by Baron Schilling. The arrival of the electrical telegraph in 1837 afforded, for the first time, a practical method for quickly gathering surface weather observations from a wide area. This data could be used to produce maps of the state of the atmosphere for a region near the Earth's surface and to study how these states evolved through time. To make frequent weather forecasts based on these data required a reliable network of observations, but it was not until 1849 that the Smithsonian Institution began to establish an observation network across the United States under the leadership of Joseph Henry. Similar observation networks were established in Europe at this time. In 1854, the United Kingdom government appointed Robert FitzRoy to the new office of *Meteorological Statist to the Board of Trade* with the role of gathering weather observations at sea. FitzRoy's office became the United Kingdom Meteorological Office in 1854, the first national meteorological service in the world. The first daily weather forecasts made by FitzRoy's Office were published in *The Times* newspaper in 1860. The following year a system was introduced of hoisting storm warning cones at principal ports when a gale was expected.

Over the next 50 years many countries established national meteorological services. The India Meteorological Department (1875) was established following tropical cyclone and monsoon related famines in the previous decades. The Finnish Meteorological Central Office (1881) was formed from part of Magnetic Observatory of Helsinki University. Japan's Tokyo Meteorological Observatory, the forerunner of the Japan Meteorological

Agency, began constructing surface weather maps in 1883. The United States Weather Bureau (1890) was established under the United States Department of Agriculture. The Australian Bureau of Meteorology (1906) was established by a Meteorology Act to unify existing state meteorological services.

## Numerical weather prediction



A meteorologist at the console of the IBM 7090 in the Joint Numerical Weather Prediction Unit. C. 1965

In 1904, Norwegian scientist Vilhelm Bjerknes first argued in his paper *Weather Forecasting as a Problem in Mechanics and Physics* that it should be possible to forecast weather from calculations based upon natural laws.

It was not until later in the 20<sup>th</sup> century that advances in the understanding of atmospheric physics led to the foundation of modern numerical weather prediction. In 1922, Lewis Fry Richardson published “Weather Prediction By Numerical Process”, after finding notes and derivations he worked on as an ambulance driver in World War I. He described therein how small terms in the prognostic fluid dynamics equations governing atmospheric flow could be neglected, and a finite differencing scheme in time and space could be devised, to allow numerical prediction solutions to be found. Richardson envisioned a large auditorium of thousands of people performing the calculations and

passing them to others. However, the sheer number of calculations required was too large to be completed without the use of computers, and the size of the grid and time steps led to unrealistic results in deepening systems. It was later found, through numerical analysis, that this was due to numerical instability.

Starting in the 1950s, numerical forecasts with computers became feasible. The first weather forecasts derived this way used barotropic (that means, single-vertical-level) models, and could successfully predict the large-scale movement of midlatitude Rossby waves, that is, the pattern of atmospheric lows and highs.

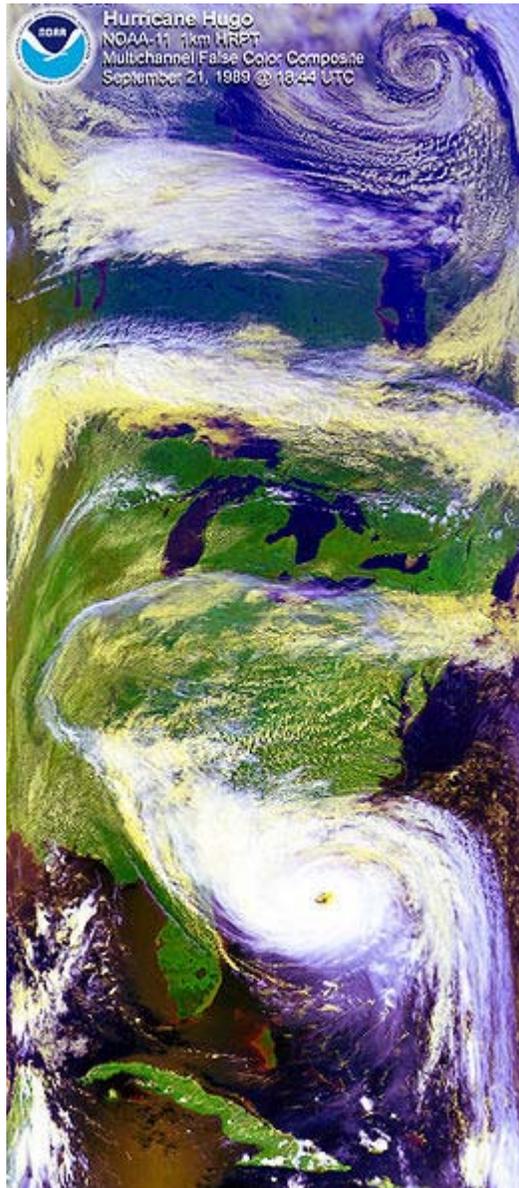
In the 1960s, the chaotic nature of the atmosphere was first observed and mathematically described by Edward Lorenz, founding the field of chaos theory. These advances have led to the current use of ensemble forecasting in most major forecasting centers, to take into account uncertainty arising from the chaotic nature of the atmosphere. In recent years, climate models have been developed that feature a resolution comparable to older weather prediction models. These climate models are used to investigate long-term climate shifts, such as what effects might be caused by human emission of greenhouse gases.

## **Meteorologists**

**Meteorologists** are scientists who study meteorology. Meteorologists work in government agencies, private consulting and research services, industrial enterprises, utilities, radio and television stations, and in education. In the United States, meteorologists held about 8,800 jobs in 2006.

Meteorologists are best-known for forecasting the weather. Many radio and television weather forecasters are professional meteorologists, while others are merely reporters with no formal meteorological training. The American Meteorological Society and National Weather Association issue “Seals of Approval” to weather broadcasters who meet certain requirements.

# Equipment



Satellite image of Hurricane Hugo with a polar low visible at the top of the image.

Each science has its own unique sets of laboratory equipment. In the atmosphere, there are many things or qualities of the atmosphere that can be measured. Rain, which can be observed, or seen anywhere and anytime was one of the first ones to be measured historically. Also, two other accurately measured *qualities* are wind and humidity. Neither of these can be *seen* but can be felt. The devices to measure these three sprang up in the mid-15<sup>th</sup> century and were respectively the rain gauge, the anemometer, and the hygrometer.

Sets of surface measurements are important data to meteorologists. They give a snapshot of a variety of weather conditions at one single location and are usually at a weather station, a ship or a weather buoy. The measurements taken at a weather station can include any number of atmospheric observables. Usually, temperature, pressure, wind measurements, and humidity are the variables that are measured by a thermometer, barometer, anemometer, and hygrometer, respectively. Upper air data are of crucial importance for weather forecasting. The most widely used technique is launches of radiosondes. Supplementing the radiosondes a network of aircraft collection is organized by the World Meteorological Organization.

Remote sensing, as used in meteorology, is the concept of collecting data from remote weather events and subsequently producing weather information. The common types of remote sensing are Radar, Lidar, and satellites (or photogrammetry). Each collects data about the atmosphere from a remote location and, usually, stores the data where the instrument is located. RADAR and LIDAR are not passive because both use EM radiation to illuminate a specific portion of the atmosphere. Weather satellites along with more general-purpose Earth-observing satellites circling the earth at various altitudes have become an indispensable tool for studying a wide range of phenomena from forest fires to El Niño.

## **Spatial scales**

In the study of the atmosphere, meteorology can be divided into distinct areas of emphasis depending on the temporal scope and spatial scope of interest. At one extreme of this scale is climatology. In the timescales of hours to days, meteorology separates into micro-, meso-, and synoptic scale meteorology. Respectively, the geospatial size of each of these three scales relates directly with the appropriate timescale.

Other subclassifications are available based on the need by or by the unique, local or broad effects that are studied within that sub-class.

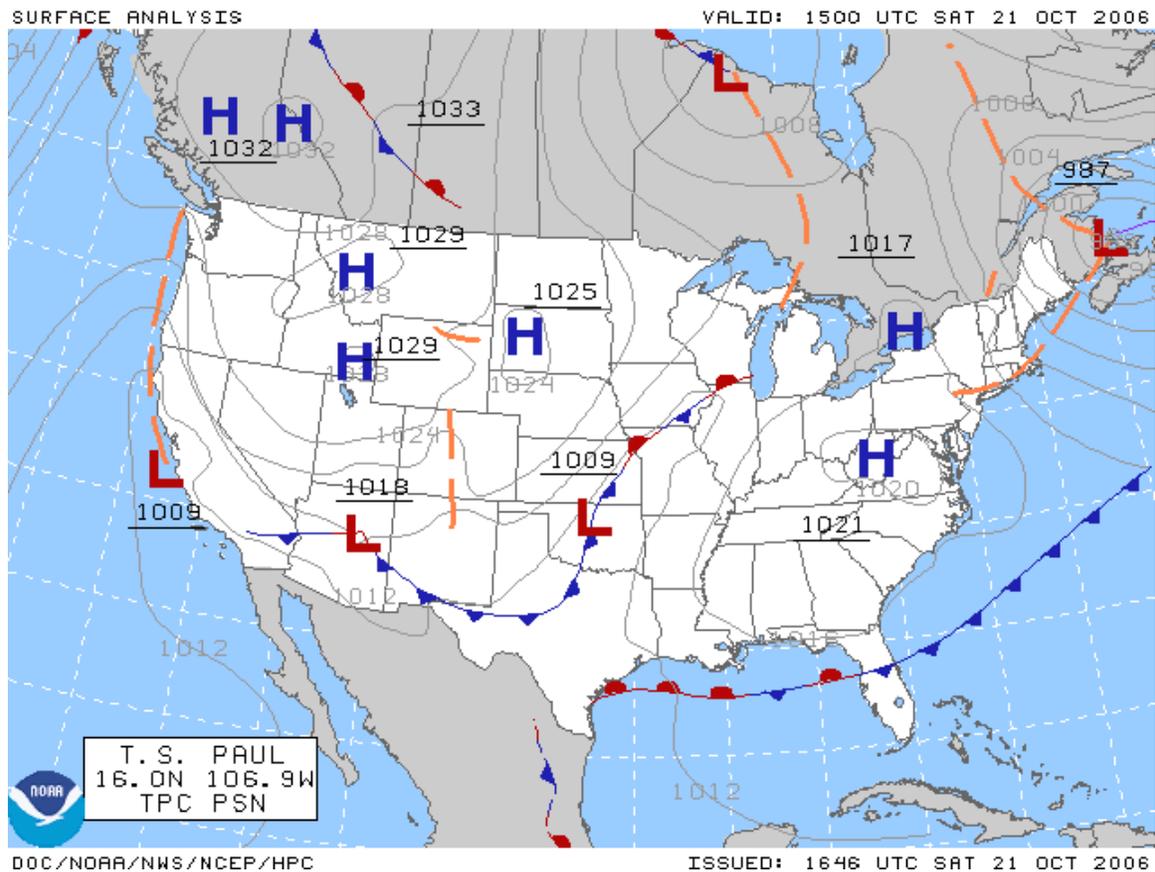
### **Microscale**

Microscale meteorology is the study of atmospheric phenomena of about 1 km or less. Individual thunderstorms, clouds, and local turbulence caused by buildings and other obstacles, such as individual hills fall within this category.

### **Mesoscale**

Mesoscale meteorology is the study of atmospheric phenomena that has horizontal scales ranging from microscale limits to synoptic scale limits and a vertical scale that starts at the Earth's surface and includes the atmospheric boundary layer, troposphere, tropopause, and the lower section of the stratosphere. Mesoscale timescales last from less than a day to the lifetime of the event, which in some cases can be weeks. The events typically of interest are thunderstorms, squall lines, fronts, precipitation bands in tropical

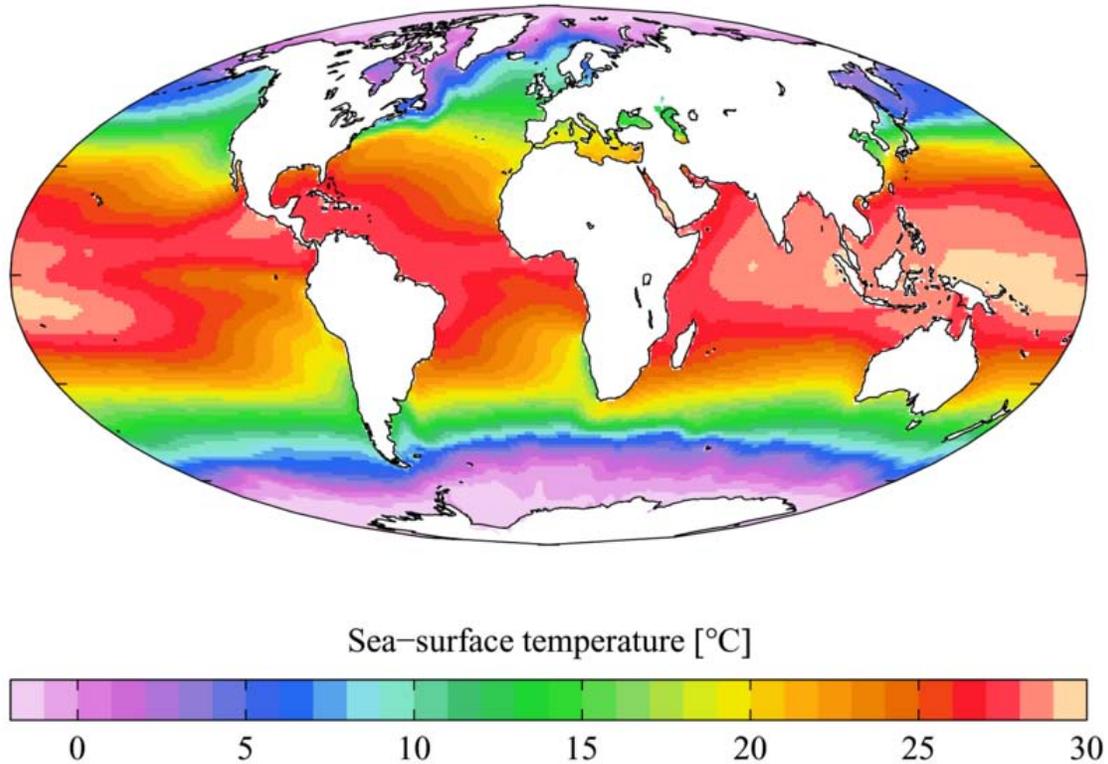
and extratropical cyclones, and topographically generated weather systems such as mountain waves and sea and land breezes.



NOAA: Synoptic scale weather analysis.

## Synoptic scale

Synoptic scale meteorology is generally large area dynamics referred to in horizontal coordinates and with respect to time. The phenomena typically described by synoptic meteorology include events like extratropical cyclones, baroclinic troughs and ridges, frontal zones, and to some extent jet streams. All of these are typically given on weather maps for a specific time. The minimum horizontal scale of synoptic phenomena are limited to the spacing between surface observation stations.



Annual mean sea surface temperatures.

## Global scale

Global scale meteorology is study of weather patterns related to the transport of heat from the tropics to the poles. Also, very large scale oscillations are of importance. Those oscillations have time periods typically longer than a full annual seasonal cycle, such as ENSO, PDO, MJO, etc. Global scale pushes the thresholds of the perception of meteorology into climatology. The traditional definition of climate is pushed in to larger timescales with the further understanding of how the global oscillations cause both climate and weather disturbances in the synoptic and mesoscale timescales.

Numerical Weather Prediction is a main focus in understanding air-sea interaction, tropical meteorology, atmospheric predictability, and tropospheric/stratospheric processes. Currently (2007) Naval Research Laboratory in Monterey produces the atmospheric model called **NOGAPS**, a global scale atmospheric model, this model is run operationally at Fleet Numerical Meteorology and Oceanography Center. There are several other global atmospheric models.

# Some meteorological principles

## Boundary layer meteorology

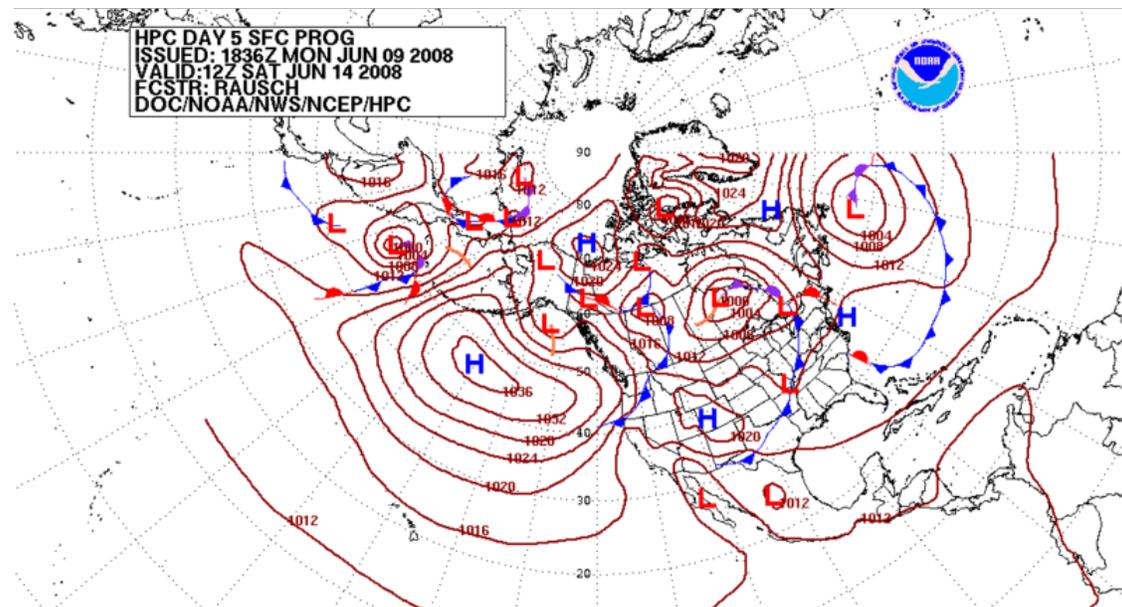
Boundary layer meteorology is the study of processes in the air layer directly above Earth's surface, known as the atmospheric boundary layer (ABL). The effects of the surface – heating, cooling, and friction – cause turbulent mixing within the air layer. Significant fluxes of heat, matter, or momentum on time scales of less than a day are advected by turbulent motions. Boundary layer meteorology includes the study of all types of surface-atmosphere boundary, including ocean, lake, urban land and non-urban land.

## Dynamic meteorology

Dynamic meteorology generally focuses on the fluid dynamics of the atmosphere. The idea of air parcel is used to define the smallest element of the atmosphere, while ignoring the discrete molecular and chemical nature of the atmosphere. An air parcel is defined as a point in the fluid continuum of the atmosphere. The fundamental laws of fluid dynamics, thermodynamics, and motion are used to study the atmosphere. The physical quantities that characterize the state of the atmosphere are temperature, density, pressure, etc. These variables have unique values in the continuum.

## Applications

### Weather forecasting



Forecast of surface pressures five days into the future for the north Pacific, North America, and north Atlantic ocean.

Weather forecasting is the application of science and technology to predict the state of the atmosphere for a future time and a given location. Human beings have attempted to predict the weather informally for millennia, and formally since at least the nineteenth century. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific understanding of atmospheric processes to project how the atmosphere will evolve.

Once an all human endeavor based mainly upon changes in barometric pressure, current weather conditions, and sky condition, forecast models are now used to determine future conditions. Human input is still required to pick the best possible forecast model to base the forecast upon, which involves pattern recognition skills, teleconnections, knowledge of model performance, and knowledge of model biases. The chaotic nature of the atmosphere, the massive computational power required to solve the equations that describe the atmosphere, error involved in measuring the initial conditions, and an incomplete understanding of atmospheric processes mean that forecasts become less accurate as the difference in current time and the time for which the forecast is being made (the *range* of the forecast) increases. The use of ensembles and model consensus help narrow the error and pick the most likely outcome.

There are a variety of end uses to weather forecasts. Weather warnings are important forecasts because they are used to protect life and property. Forecasts based on temperature and precipitation are important to agriculture, and therefore to commodity traders within stock markets. Temperature forecasts are used by utility companies to estimate demand over coming days. On an everyday basis, people use weather forecasts to determine what to wear on a given day. Since outdoor activities are severely curtailed by heavy rain, snow and the wind chill, forecasts can be used to plan activities around these events, and to plan ahead and survive them.

### **Aviation meteorology**

Aviation meteorology deals with the impact of weather on air traffic management. It is important for air crews to understand the implications of weather on their flight plan as well as their aircraft, as noted by the Aeronautical Information Manual:

*The effects of ice on aircraft are cumulative-thrust is reduced, drag increases, lift lessens, and weight increases. The results are an increase in stall speed and a deterioration of aircraft performance. In extreme cases, 2 to 3 inches of ice can form on the leading edge of the airfoil in less than 5 minutes. It takes but ½ inch of ice to reduce the lifting power of some aircraft by 50 percent and increases the frictional drag by an equal percentage.*

### **Agricultural meteorology**

Meteorologists, soil scientists, agricultural hydrologists, and agronomists are persons concerned with studying the effects of weather and climate on plant distribution, crop yield, water-use efficiency, phenology of plant and animal development, and the energy

balance of managed and natural ecosystems. Conversely, they are interested in the role of vegetation on climate and weather.

## **Hydrometeorology**

Hydrometeorology is the branch of meteorology that deals with the hydrologic cycle, the water budget, and the rainfall statistics of storms. A hydrometeorologist prepares and issues forecasts of accumulating (quantitative) precipitation, heavy rain, heavy snow, and highlights areas with the potential for flash flooding. Typically the range of knowledge that is required overlaps with climatology, mesoscale and synoptic meteorology, and other geosciences.

## **Nuclear meteorology**

Nuclear meteorology investigates the distribution of radioactive aerosols and gases in the atmosphere.

## **Maritime meteorology**

Maritime meteorology deals with air and wave forecasts for ships operating at sea. Organizations such as the Ocean Prediction Center, Honolulu National Weather Service forecast office, United Kingdom Met Office, and JMA prepare high seas forecasts for the world's oceans.

# **Climatology**

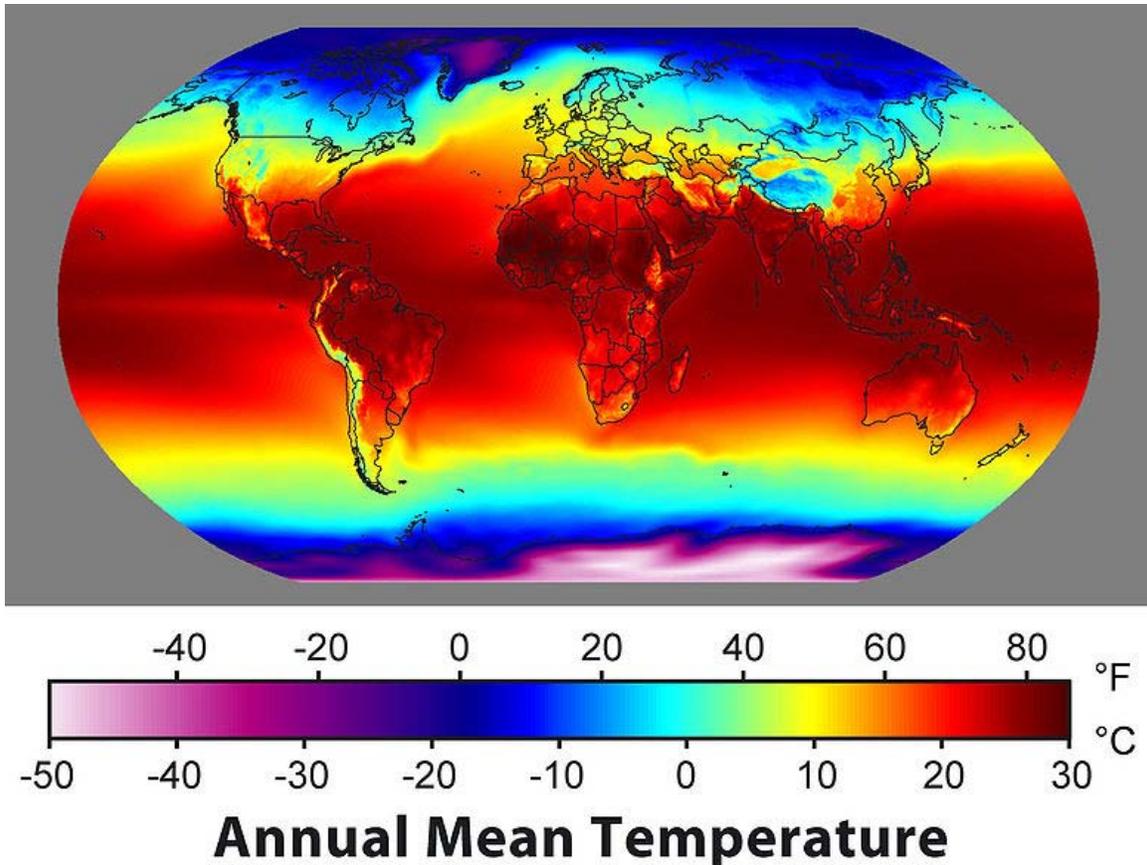
**Climatology** is the study of climate, scientifically defined as weather conditions averaged over a period of time, and is a branch of the atmospheric sciences. Basic knowledge of climate can be used within shorter term weather forecasting using analog techniques such as the El Niño – Southern Oscillation (ENSO), the Madden-Julian Oscillation (MJO), the North Atlantic Oscillation (NAO), the Northern Annular Mode (NAM), the Arctic oscillation (AO), the Northern Pacific (NP) Index, the Pacific Decadal Oscillation (PDO), and the Interdecadal Pacific Oscillation (IPO). Climate models are used for a variety of purposes from study of the dynamics of the weather and climate system to projections of future climate.

## **History**

The earliest person to hypothesize the concept of climate change may have been the medieval Chinese scientist Shen Kuo (1031-1095 AD). Shen Kuo theorized that climates naturally shifted over an enormous span of time, after observing petrified bamboos found underground near Yanzhou (modern day Yan'an, Shaanxi province), a dry climate area unsuitable for the growth of bamboo.

Early climate researchers include Edmund Halley, who published a map of the trade winds in 1686, after a voyage to the southern hemisphere. Benjamin Franklin, in the 18<sup>th</sup> century, was the first to map the course of the Gulf Stream for use in sending mail overseas from the United States to Europe. Francis Galton invented the term *anticyclone*. Helmut Landsberg led to statistical analysis being used in climatology, which led to its evolution into a physical science.

## Different approaches



Map of the average temperature over 30 years. Data sets formed from the long-term average of historical weather parameters are sometimes called a “climatology”.

Climatology is approached in a variety of ways. **Paleoclimatology** seeks to reconstruct past climates by examining records such as ice cores and tree rings (dendroclimatology). **Paleotempestology** uses these same records to help determine hurricane frequency over millennia. The study of contemporary climates incorporates meteorological data accumulated over many years, such as records of rainfall, temperature and atmospheric composition. Knowledge of the atmosphere and its dynamics is also embodied in models, either statistical or mathematical, which help by integrating different observations and testing how they fit together. Modeling is used for understanding past, present and

potential future climates. **Historical climatology** is the study of climate as related to human history and thus focuses only on the last few thousand years.

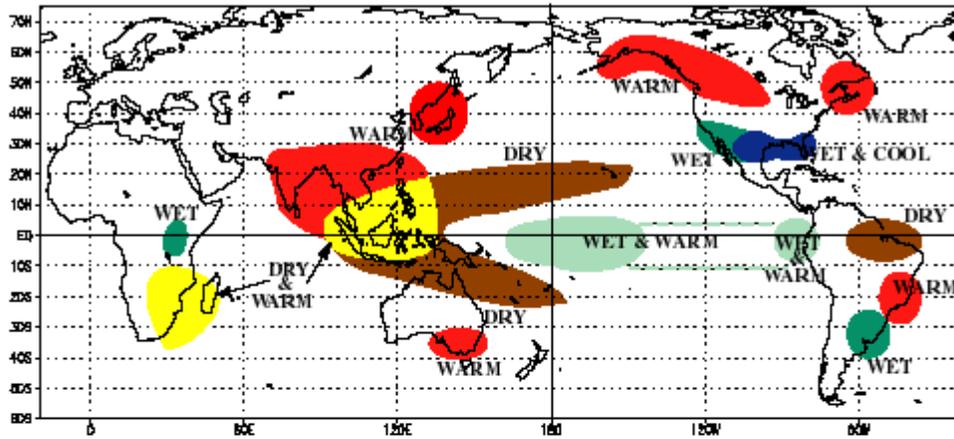
Climate research is made difficult by the large scale, long time periods, and complex processes which govern climate. Climate is governed by physical laws which can be expressed as differential equations. These equations are coupled and nonlinear, so that approximate solutions are obtained by using numerical methods to create global climate models. Climate is sometimes modeled as a stochastic process but this is generally accepted as an approximation to processes that are otherwise too complicated to analyze.

## **Indices**

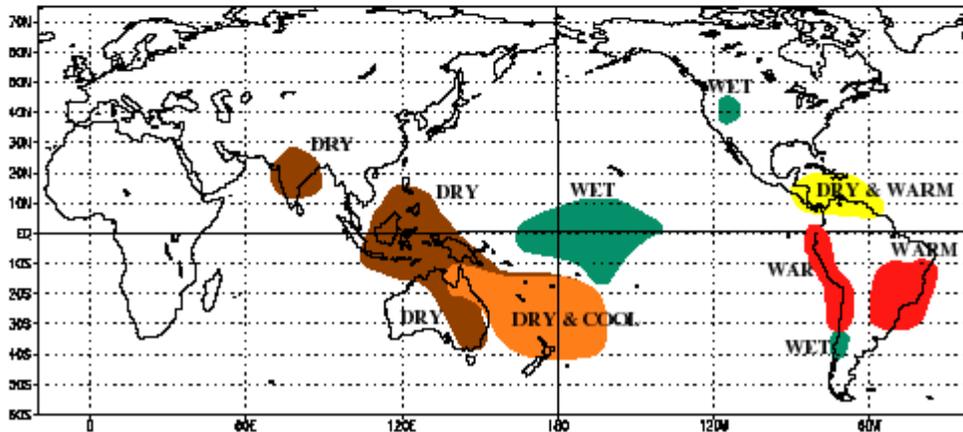
Scientists use climate indices based on several climate patterns (known as modes of variability) in their attempt to characterize and understand the various climate mechanisms that culminate in our daily weather. Much in the way the Dow Jones Industrial Average, which is based on the stock prices of 30 companies, is used to represent the fluctuations in the stock market as a whole, climate indices are used to represent the essential elements of climate. Climate indices are generally devised with the twin objectives of simplicity and completeness, and each index typically represents the status and timing of the climate factor it represents. By their very nature, indices are simple, and combine many details into a generalized, overall description of the atmosphere or ocean which can be used to characterize the factors which impact the global climate system.

# El Niño – Southern Oscillation

## WARM EPISODE RELATIONSHIPS DECEMBER - FEBRUARY

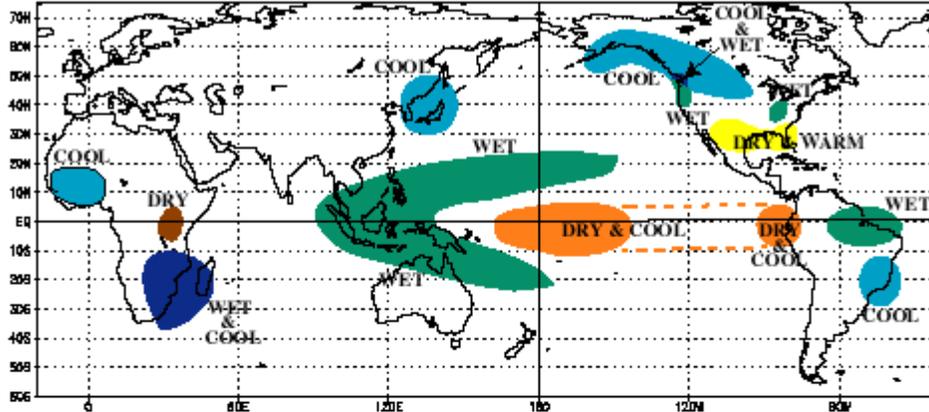


## WARM EPISODE RELATIONSHIPS JUNE - AUGUST

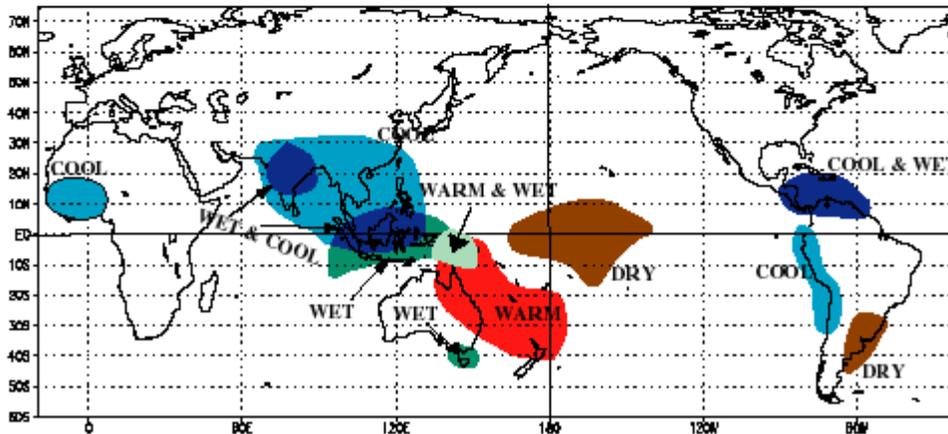


El Niño impacts

## COLD EPISODE RELATIONSHIPS DECEMBER - FEBRUARY



## COLD EPISODE RELATIONSHIPS JUNE - AUGUST



### La Niña impacts

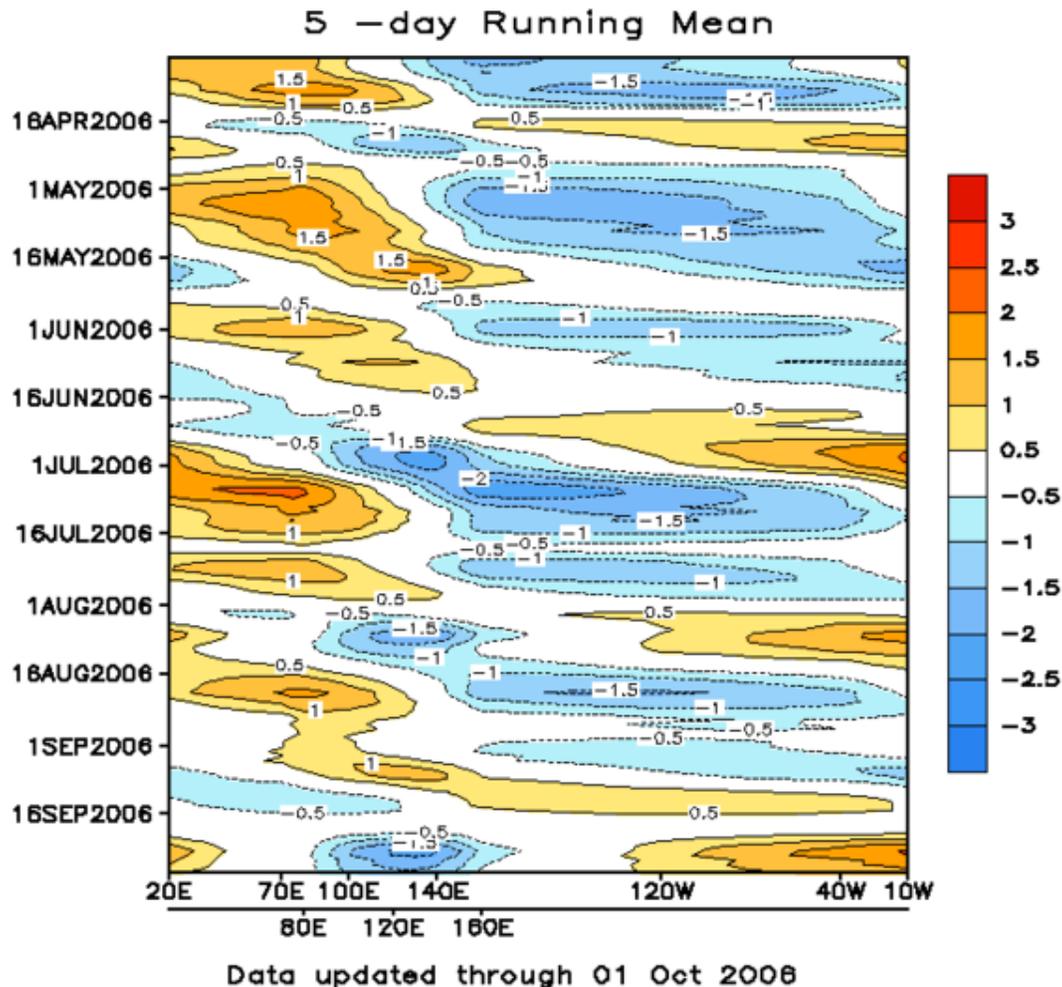
El Niño-Southern Oscillation (ENSO) is a global coupled ocean-atmosphere phenomenon. The Pacific ocean signatures, El Niño and La Niña are important temperature fluctuations in surface waters of the tropical Eastern Pacific Ocean. The name El Niño, from the Spanish for “the little boy”, refers to the Christ child, because the phenomenon is usually noticed around Christmas time in the Pacific Ocean off the west coast of South America. La Niña means “the little girl”. Their effect on climate in the subtropics and the tropics are profound. The atmospheric signature, the Southern Oscillation (SO) reflects the monthly or seasonal fluctuations in the air pressure difference between Tahiti and Darwin. The most recent occurrence of El Niño started in September 2006 and lasted until early 2007.

ENSO is a set of interacting parts of a single global system of coupled ocean-atmosphere climate fluctuations that come about as a consequence of oceanic and atmospheric circulation. ENSO is the most prominent known source of inter-annual variability in

weather and climate around the world. The cycle occurs every two to seven years, with El Niño lasting nine months to two years within the longer term cycle, though not all areas globally are affected. ENSO has signatures in the Pacific, Atlantic and Indian Oceans. El Niño causes weather patterns which causes it to rain in specific places but not in others, this is one of many causes for the drought.

In the Pacific, during major warm events, El Niño warming extends over much of the tropical Pacific and becomes clearly linked to the SO intensity. While ENSO events are basically in phase between the Pacific and Indian Oceans, ENSO events in the Atlantic Ocean lag behind those in the Pacific by 12 to 18 months. Many of the countries most affected by ENSO events are developing countries within tropical sections of continents with economies that are largely dependent upon their agricultural and fishery sectors as a major source of food supply, employment, and foreign exchange. New capabilities to predict the onset of ENSO events in the three oceans can have global socio-economic impacts. While ENSO is a global and natural part of the Earth's climate, whether its intensity or frequency may change as a result of global warming is an important concern. Low-frequency variability has been evidenced: the quasi-decadal oscillation (QDO). Inter-decadal (ID) modulation of ENSO (from PDO or IPO) might exist. This could explain the so-called protracted ENSO of the early 1990s.

## Madden-Julian Oscillation



Note how the MJO moves eastward with time.

The Madden-Julian Oscillation (MJO) is an equatorial traveling pattern of anomalous rainfall that is planetary in scale. It is characterized by an eastward progression of large regions of both enhanced and suppressed tropical rainfall, observed mainly over the Indian Ocean and Pacific Ocean. The anomalous rainfall is usually first evident over the western Indian Ocean, and remains evident as it propagates over the very warm ocean waters of the western and central tropical Pacific. This pattern of tropical rainfall then generally becomes very nondescript as it moves over the cooler ocean waters of the eastern Pacific but reappears over the tropical Atlantic and Indian Ocean. The wet phase of enhanced convection and precipitation is followed by a dry phase where convection is suppressed. Each cycle lasts approximately 30–60 days. The MJO is also known as the 30-60 day oscillation, 30-60 day wave, or intraseasonal oscillation.

## **North Atlantic Oscillation (NAO)**

Indices of the NAO are based on the difference of normalized sea level pressure (SLP) between Ponta Delgada, Azores and Stykkisholmur/Reykjavik, Iceland. The SLP anomalies at each station were normalized by division of each seasonal mean pressure by the long-term mean (1865-1984) standard deviation. Normalization is done to avoid the series of being dominated by the greater variability of the northern of the two stations. Positive values of the index indicate stronger-than-average westerlies over the middle latitudes.

## **Northern Annular Mode (NAM) or Arctic Oscillation (AO)**

The NAM, or AO, is defined as the first EOF of northern hemisphere winter SLP data from the tropics and subtropics. It explains 23% of the average winter (December-March) variance, and it is dominated by the NAO structure in the Atlantic. Although there are some subtle differences from the regional pattern over the Atlantic and Arctic, the main difference is larger amplitude anomalies over the North Pacific of the same sign as those over the Atlantic. This feature gives the NAM a more annular (or zonally symmetric) structure.

## **Northern Pacific (NP) Index**

The NP Index is the area-weighted sea level pressure over the region 30N-65N, 160E-140W.

## **Pacific Decadal Oscillation (PDO)**

The PDO is a pattern of Pacific climate variability that shifts phases on at least inter-decadal time scale, usually about 20 to 30 years. The PDO is detected as warm or cool surface waters in the Pacific Ocean, north of 20° N. During a “warm”, or “positive”, phase, the west Pacific becomes cool and part of the eastern ocean warms; during a “cool” or “negative” phase, the opposite pattern occurs. The mechanism by which the pattern lasts over several years has not been identified; one suggestion is that a thin layer of warm water during summer may shield deeper cold waters. A PDO signal has been reconstructed to 1661 through tree-ring chronologies in the Baja California area.

## **Interdecadal Pacific Oscillation (IPO)**

The Interdecadal Pacific Oscillation (IPO or ID) display similar sea surface temperature (SST) and sea level pressure patterns to the PDO, with a cycle of 15–30 years, but affects both the north and south Pacific. In the tropical Pacific, maximum SST anomalies are found away from the equator. This is quite different from the quasi-decadal oscillation (QDO) with a period of 8-to-12 years and maximum SST anomalies straddling the equator, thus resembling ENSO.

## Models

Climate models use quantitative methods to simulate the interactions of the atmosphere, oceans, land surface, and ice. They are used for a variety of purposes from study of the dynamics of the weather and climate system to projections of future climate. All climate models balance, or very nearly balance, incoming energy as short wave (including visible) electromagnetic radiation to the earth with outgoing energy as long wave (infrared) electromagnetic radiation from the earth. Any unbalance results in a change in the average temperature of the earth.

The most talked-about models of recent years have been those relating temperature to emissions of carbon dioxide. These models predict an upward trend in the surface temperature record, as well as a more rapid increase in temperature at higher altitudes.

Models can range from relatively simple to quite complex:

- A simple radiant heat transfer model that treats the earth as a single point and averages outgoing energy
- this can be expanded vertically (radiative-convective models), or horizontally
- finally, (coupled) atmosphere–ocean–sea ice **global climate models** discretise and solve the full equations for mass and energy transfer and radiant exchange.

## Differences with meteorology

In contrast to meteorology, which focuses on short term weather systems lasting up to a few weeks, climatology studies the frequency and trends of those systems. It studies the periodicity of weather events over years to millennia, as well as changes in long-term average weather patterns, in relation to atmospheric conditions. Climatologists, those who practice climatology, study both the nature of climates – local, regional or global – and the natural or human-induced factors that cause climates to change. Climatology considers the past and can help predict future climate change.

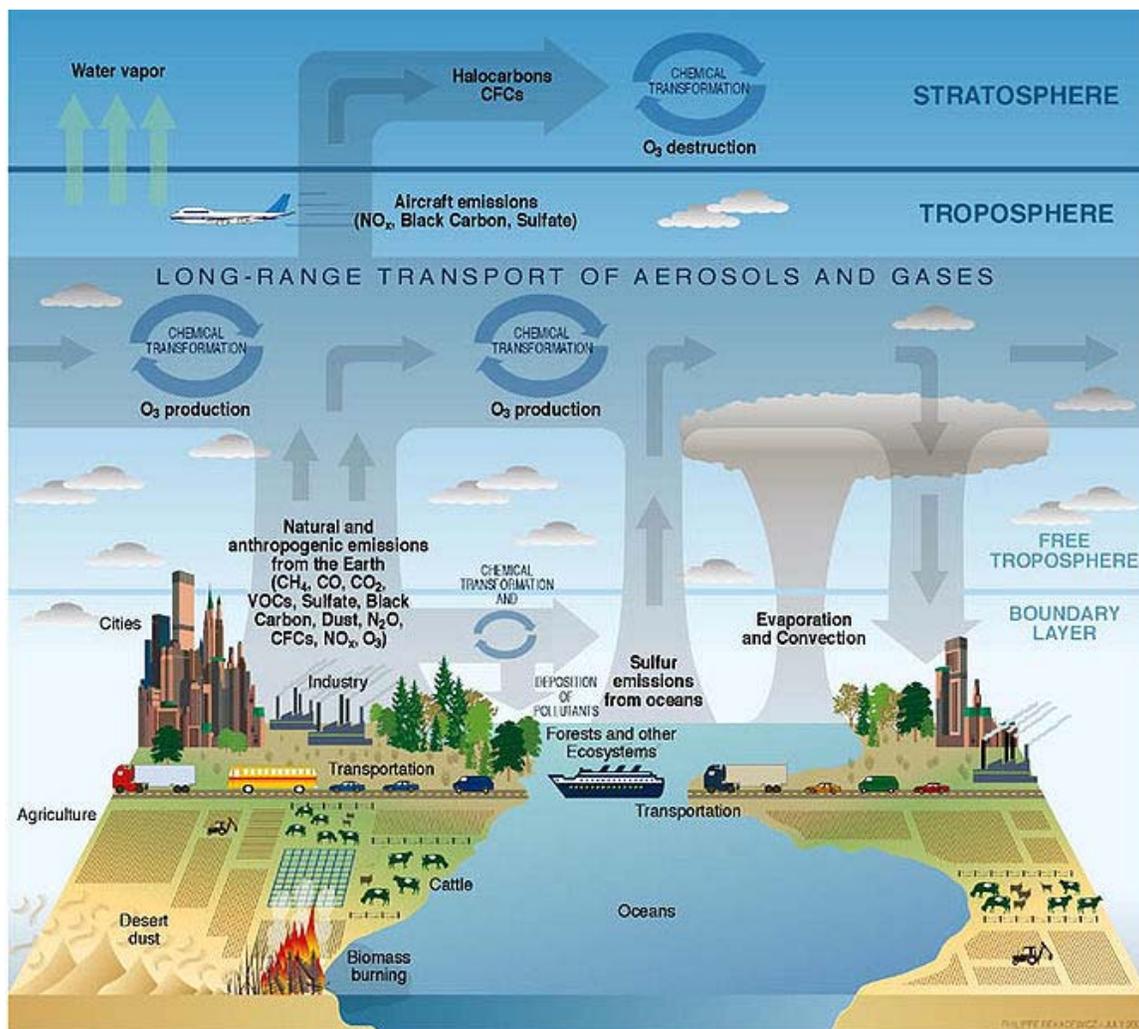
Phenomena of climatological interest include the atmospheric boundary layer, circulation patterns, heat transfer (radiative, convective and latent), interactions between the atmosphere and the oceans and land surface (particularly vegetation, land use and topography), and the chemical and physical composition of the atmosphere.

## Use in weather forecasting

A more complicated way of making a forecast, the analog technique requires remembering a previous weather event which is expected to be mimicked by an upcoming event. What makes it a difficult technique to use is that there is rarely a perfect analog for an event in the future. Some call this type of forecasting **pattern recognition**, which remains a useful method of observing rainfall over data voids such as oceans with knowledge of how satellite imagery relates to precipitation rates over land, as well as the

forecasting of precipitation amounts and distribution in the future. A variation on this theme is used in Medium Range forecasting, which is known as teleconnections, when you use systems in other locations to help pin down the location of another system within the surrounding regime. One method of using teleconnections are by using climate indices such as ENSO-related phenomena.

## Atmospheric chemistry



Schematic of chemical and transport processes related to atmospheric composition.

**Atmospheric chemistry** is a branch of atmospheric science in which the chemistry of the Earth's atmosphere and that of other planets is studied. It is a multidisciplinary field of research and draws on environmental chemistry, physics, meteorology, computer modeling, oceanography, geology and volcanology and other disciplines. Research is increasingly connected with other areas of study such as climatology.

The composition and chemistry of the atmosphere is of importance for several reasons, but primarily because of the interactions between the atmosphere and living organisms. The composition of the Earth's atmosphere has been changed by human activity and some of these changes are harmful to human health, crops and ecosystems. Examples of problems which have been addressed by atmospheric chemistry include acid rain, ozone depletion, photochemical smog and global warming. Atmospheric chemistry seeks to understand the causes of these problems, and by obtaining a theoretical understanding of them, allow possible solutions to be tested and the effects of changes in government policy evaluated.

## Atmospheric composition

### Average composition of dry atmosphere, by volume

Gas	per NASA
Nitrogen, N <sub>2</sub>	78.084%
Oxygen, O <sub>2</sub>	20.946%
Argon, Ar	0.934%
<b>Minor constituents (in ppm)</b>	
Carbon Dioxide, CO <sub>2</sub>	383
Neon, Ne	18.18
Helium, He	5.24
Methane, CH <sub>4</sub>	1.7
Krypton, Kr	1.14
Hydrogen, H <sub>2</sub>	0.55
<b>Water</b>	
Water vapour	Highly variable; typically makes up about 1%

Notes: the concentration of CO<sub>2</sub> and CH<sub>4</sub> vary by season and location. The mean molecular mass of air is 28.97 g/mol.

## History

The ancient Greeks regarded air as one of the four elements, but the first scientific studies of atmospheric composition began in the 18<sup>th</sup> century. Chemists such as Joseph Priestley, Antoine Lavoisier and Henry Cavendish made the first measurements of the composition of the atmosphere.

In the late 19<sup>th</sup> and early 20<sup>th</sup> centuries interest shifted towards trace constituents with very small concentrations. One particularly important discovery for atmospheric chemistry was the discovery of ozone by Christian Friedrich Schoenbein in 1840.

In the 20<sup>th</sup> century atmospheric science moved on from studying the composition of air to a consideration of how the concentrations of trace gases in the atmosphere have changed over time and the chemical processes which create and destroy compounds in the air. Two particularly important examples of this were the explanation of how the ozone layer is created and maintained by Sydney Chapman and Gordon Dobson, and the explanation of Photochemical smog by Arie Jan Haagen-Smit. Further studies on ozone issues led to the 1995 Nobel Prize in Chemistry award shared between Paul Crutzen, Mario Molina and Frank Sherwood Rowland.

In the 21<sup>st</sup> century the focus is now shifting again. Atmospheric chemistry is increasingly studied as one part of the Earth system. Instead of concentrating on atmospheric chemistry in isolation the focus is now on seeing it as one part of a single system with the rest of the atmosphere, biosphere and geosphere. An especially important driver for this is the links between chemistry and climate such as the effects of changing climate on the recovery of the ozone hole and vice versa but also interaction of the composition of the atmosphere with the oceans and terrestrial ecosystems.

## **Methodology**

Observations, lab measurements and modeling are the three central elements in atmospheric chemistry. Progress in atmospheric chemistry is often driven by the interactions between these components and they form an integrated whole. For example observations may tell us that more of a chemical compound exists than previously thought possible. This will stimulate new riticiz and laboratory studies which will increase our scientific understanding to a point where the observations can be explained.

### **Observation**

Observations of atmospheric chemistry are essential to our understanding. Routine observations of chemical composition tell us about changes in atmospheric composition over time. One important example of this is the Keeling Curve – a series of measurements from 1958 to today which show a steady rise in of the concentration of carbon dioxide. Observations of atmospheric chemistry are made in observatories such as that on Mauna Loa and on mobile platforms such as aircraft (e.g. the UK's Facility for Airborne Atmospheric Measurements), ships and balloons. Observations of atmospheric composition are increasingly made by satellites with important instruments such as GOME and MOPITT giving a global picture of air pollution and chemistry. Surface observations have the advantage that they provide long term records at high time resolution but are limited in the vertical and horizontal space they provide observations from. Some surface based instruments e.g. LIDAR can provide concentration profiles of chemical compounds and aerosol but are still restricted in the horizontal region they can cover. Many observations are available on line in Atmospheric Chemistry Observational Databases.

## Lab measurements

Measurements made in the laboratory are essential to our understanding of the sources and sinks of pollutants and naturally occurring compounds. Lab studies tell us which gases react with each other and how fast they react. Measurements of interest include reactions in the gas phase, on surfaces and in water. Also of high importance is photochemistry which quantifies how quickly molecules are split apart by sunlight and what the products are plus thermodynamic data such as Henry's law coefficients.

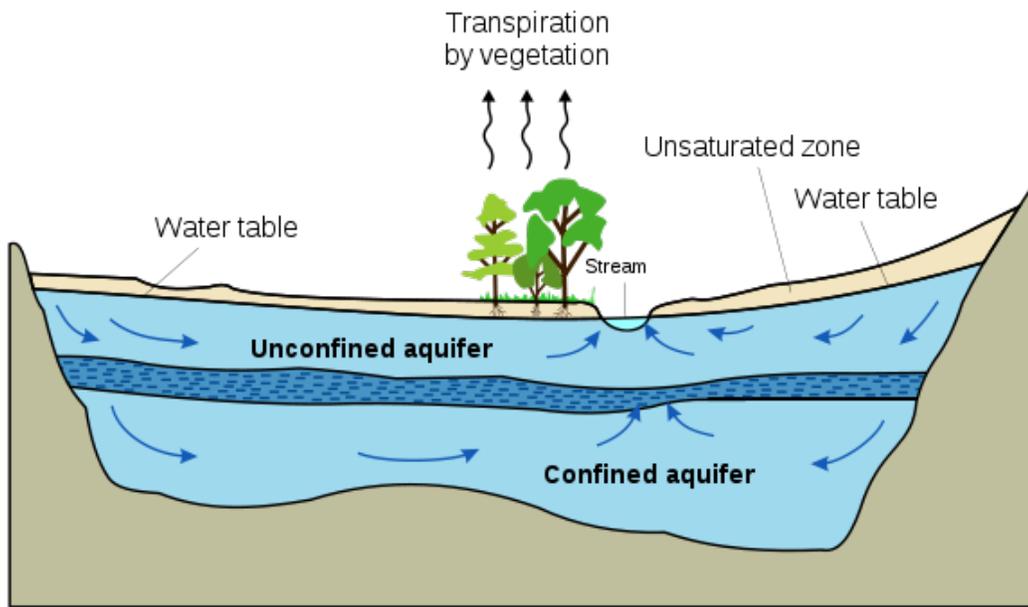
## Modeling

In order to synthesise and test theoretical understanding of atmospheric chemistry, computer models (such as chemical transport models) are used. Numerical models solve the differential equations governing the concentrations of chemicals in the atmosphere. They can be very simple or very complicated. One common trade off in numerical models is between the number of chemical compounds and chemical reactions versus the representation of transport and mixing in the atmosphere. For example, a box model might include hundreds or even thousands of chemical reactions but will only have a very crude representation of mixing in the atmosphere. In contrast, 3D models represent many of the physical processes of the atmosphere but due to constraints on computer resources will have far fewer chemical reactions and compounds. Models can be used to interpret observations, test understanding of chemical reactions and predict future concentrations of chemical compounds in the atmosphere. One important current trend is for atmospheric chemistry modules to become one part of earth system models in which the links between climate, atmospheric composition and the biosphere can be studied.

Some models are constructed by automatic code generators (e.g. Autochem or KPP). In this approach a set of constituents are chosen and the automatic code generator will then select the reactions involving those constituents from a set of reaction databases. Once the reactions have been chosen the ordinary differential equations (ODE) that describe their time evolution can be automatically constructed.

## Hydrogeology

**Hydrogeology** (*hydro-* meaning water, and *-geology* meaning the study of the Earth) is the area of geology that deals with the distribution and movement of groundwater in the soil and rocks of the Earth's crust, (commonly in aquifers). The term **geohydrology** is often used interchangeably. Some make the minor distinction between a hydrologist or engineer applying themselves to geology (geohydrology), and a geologist applying themselves to hydrology (hydrogeology).



-  **High hydraulic-conductivity aquifer**
-  **Low hydraulic-conductivity confining unit**
-  **Very low hydraulic-conductivity bedrock**
-  **Direction of ground-water flow**

Typical aquifer cross-section

## Introduction

Hydrogeology is an interdisciplinary subject; it can be difficult to account fully for the chemical, physical, biological and even legal interactions between soil, water, nature and society. The study of the interaction between groundwater movement and geology can be quite complex. Groundwater does not always flow in the subsurface down-hill following the surface topography; groundwater follows pressure gradients (flow from high pressure to low) often following fractures and conduits in circuitous paths. Taking into account the interplay of the different facets of a multi-component system often requires knowledge in several diverse fields at both the experimental and theoretical levels. The following is a more traditional introduction to the methods and nomenclature of saturated subsurface hydrology, or simply hydrogeology.

## **Hydrogeology in relation to other fields**

Hydrogeology, as stated above, is a branch of the earth sciences dealing with the flow of water through aquifers and other shallow porous media (typically less than 450 m or 1,500 ft below the land surface.) The very shallow flow of water in the subsurface (the upper 3 m or 10 ft) is pertinent to the fields of soil science, agriculture and civil engineering, as well as to hydrogeology. The general flow of fluids (water, hydrocarbons, geothermal fluids, etc.) in deeper formations is also a concern of geologists, geophysicists and petroleum geologists. Groundwater is a slow-moving, viscous fluid (with a Reynolds number less than unity); many of the empirically derived laws of groundwater flow can be alternately derived in fluid mechanics from the special case of Stokes flow (viscosity and pressure terms, but no inertial term).

The mathematical relationships used to describe the flow of water through porous media are the diffusion and Laplace equations, which have applications in many diverse fields. Steady groundwater flow (Laplace equation) has been simulated using electrical, elastic and heat conduction analogies. Transient groundwater flow is analogous to the diffusion of heat in a solid, therefore some solutions to hydrological problems have been adapted from heat transfer literature.

Traditionally, the movement of groundwater has been studied separately from surface water, climatology, and even the chemical and microbiological aspects of hydrogeology (the processes are uncoupled). As the field of hydrogeology matures, the strong interactions between groundwater, surface water, water chemistry, soil moisture and even climate are becoming more clear.

## **Definitions and material properties**

One of the main tasks a hydrogeologist typically performs is the prediction of future behavior of an aquifer system, based on analysis of past and present observations. Some hypothetical, but characteristic questions asked would be:

- Can the aquifer support another subdivision?
- Will the river dry up if the farmer doubles his irrigation?
- Did the chemicals from the dry cleaning facility travel through the aquifer to my well and make me sick?
- Will the plume of effluent leaving my neighbor's septic system flow to my drinking water well?

Most of these questions can be addressed through simulation of the hydrologic system (using numerical models or analytic equations). Accurate simulation of the aquifer system requires knowledge of the aquifer properties and boundary conditions. Therefore a common task of the hydrogeologist is determining aquifer properties using aquifer tests.

In order to further characterize aquifers and aquitards some primary and derived physical properties are introduced below. Aquifers are broadly classified as being either confined or unconfined (water table aquifers), and either saturated or unsaturated; the type of aquifer affects what properties control the flow of water in that medium (e.g., the release of water from storage for confined aquifers is related to the storativity, while it is related to the specific yield for unconfined aquifers).

## Hydraulic head

Changes in hydraulic head ( $h$ ) are the driving force which causes water to move from one place to another. It is composed of pressure head ( $\psi$ ) and elevation head ( $z$ ). The head gradient is the change in hydraulic head per length of flowpath, and appears in Darcy's law as being proportional to the discharge.

Hydraulic head is a directly measurable property that can take on any value (because of the arbitrary datum involved in the  $z$  term);  $\psi$  can be measured with a pressure transducer (this value can be negative, e.g., suction, but is positive in saturated aquifers), and  $z$  can be measured relative to a surveyed datum (typically the top of the well casing). Commonly, in wells tapping unconfined aquifers the water level in a well is used as a proxy for hydraulic head, assuming there is no vertical gradient of pressure. Often only *changes* in hydraulic head through time are needed, so the constant elevation head term can be left out ( $\Delta h = \Delta \psi$ ).

A record of hydraulic head through time at a well is a hydrograph or, the changes in hydraulic head recorded during the pumping of a well in a test are called drawdown.

## Porosity

Porosity ( $n$ ) is a directly measurable aquifer property; it is a fraction between 0 and 1 indicating the amount of pore space between unconsolidated soil particles or within a fractured rock. Typically, the majority of groundwater (and anything dissolved in it) moves through the porosity available to flow (sometimes called effective porosity).

**Permeability** is an expression of the connectedness of the pores. For instance, an unfractured rock unit may have a high *porosity* (it has lots of *holes* between its constituent grains), but a low *permeability* (none of the pores are connected). An example of this phenomenon is pumice, which, when in its unfractured state, can make a poor aquifer.

Porosity does not directly affect the distribution of hydraulic head in an aquifer, but it has a very strong effect on the migration of dissolved contaminants, since it affects groundwater flow velocities through an inversely proportional relationship.

## Water content

Water content ( $\theta$ ) is also a directly measurable property; it is the fraction of the total rock which is filled with liquid water. This is also a fraction between 0 and 1, but it must also be less than or equal to the total porosity.

The water content is very important in vadose zone hydrology, where the hydraulic conductivity is a strongly nonlinear function of water content; this complicates the solution of the unsaturated groundwater flow equation.

## Hydraulic conductivity

Hydraulic conductivity ( $K$ ) and transmissivity ( $T$ ) are indirect aquifer properties (they cannot be measured directly).  $T$  is the  $K$  integrated over the vertical thickness ( $b$ ) of the aquifer ( $T=Kb$  when  $K$  is constant over the entire thickness). These properties are measures of an aquifer's ability to transmit water. Intrinsic permeability ( $\kappa$ ) is a secondary medium property which does not depend on the viscosity and density of the fluid ( $K$  and  $T$  are specific to water); it is used more in the petroleum industry.

## Specific storage and specific yield

Specific storage ( $S_s$ ) and its depth-integrated equivalent, storativity ( $S=S_s b$ ), are indirect aquifer properties (they cannot be measured directly); they indicate the amount of groundwater released from storage due to a unit depressurization of a confined aquifer. They are fractions between 0 and 1.

Specific yield ( $S_y$ ) is also a ratio between 0 and 1 ( $S_y \leq$  porosity) and indicates the amount of water released due to drainage from lowering the water table in an unconfined aquifer. The value for specific yield is less than the value for porosity because some water will remain in the medium even after drainage due to molecular forces. Often the porosity or effective porosity is used as an upper bound to the specific yield. Typically  $S_y$  is orders of magnitude larger than  $S_s$ .

## Contaminant transport properties

Often we are interested in how the moving groundwater will move dissolved contaminants around (the sub-field of contaminant hydrogeology). The contaminants can be man-made (e.g., petroleum products, nitrate or Chromium) or naturally occurring (e.g., arsenic, salinity). Besides needing to understand where the groundwater is flowing, based on the other hydrologic properties discussed above, there are additional aquifer properties which affect how dissolved contaminants move with groundwater.

## Hydrodynamic dispersion

Hydrodynamic dispersivity ( $\alpha_L$ ,  $\alpha_T$ ) is an empirical factor which quantifies how much contaminants stray away from the path of the groundwater which is carrying it. Some of

the contaminants will be “behind” or “ahead” the mean groundwater, giving rise to a longitudinal dispersivity ( $\alpha_L$ ), and some will be “to the sides of” the pure advective groundwater flow, leading to a transverse dispersivity ( $\alpha_T$ ). Dispersion in groundwater is due to the fact that each water “particle”, passing beyond a soil particle, must choose where to go, whether left or right or up or down, so that the water “particles” (and their solute) are gradually spread in all directions around the mean path. This is the “microscopic” mechanism, on the scale of soil particles. More important, on long distances, can be the macroscopic inhomogeneities of the aquifer, which can have regions of larger or smaller permeability, so that some water can find a preferential path in one direction, some other in a different direction, so that the contaminant can be spread in a completely irregular way, like in a (three-dimensional) delta of a river.

Dispersivity is actually a factor which represents our *lack of information* about the system we are simulating. There are many small details about the aquifer which are being averaged when using a macroscopic approach (e.g., tiny beds of gravel and clay in sand aquifers), they manifest themselves as an *apparent* dispersivity. Because of this,  $\alpha$  is often claimed to be dependent on the length scale of the problem — the dispersivity found for transport through 1 m<sup>3</sup> of aquifer is different than that for transport through 1 cm<sup>3</sup> of the same aquifer material.

### **Molecular Diffusion**

Diffusion is a fundamental physical phenomenon, by which Einstein explained Brownian motion, which describes the random thermal movement of molecules and small particles in gases and liquids. It is an important phenomenon for small distances (it is essential for the achievement of  $\square$ riticized $\square$ ic equilibria), but, as the time necessary to cover a distance by diffusion is proportional to the square of the distance itself, it is ineffective for spreading a solute over macroscopic distances. The diffusion coefficient,  $D$ , is typically quite small, and its effect can often be considered negligible (unless groundwater flow velocities are extremely low, as they are in clay aquitards).

It is important not to confuse diffusion with dispersion, as the former is a physical phenomenon and the latter is an empirical factor which is cast into a similar form as diffusion, because we already know how to solve that problem.

### **Retardation by adsorption**

The retardation factor is another very important feature that make the motion of the contaminant to deviate from the average groundwater motion. It is analogous to the retardation factor of chromatography. Unlike diffusion and dispersion, which simply spread the contaminant, the retardation factor changes its *global average velocity*, so that it can be much slower than that of water. This is due to a chemico-physical effect: the adsorption to the soil, which holds the contaminant back and does not allow it to progress until the quantity corresponding to the chemical adsorption equilibrium has been adsorbed. This effect is particularly important for less soluble contaminants, which thus can move even hundreds or thousands times slower than water. The effect of this

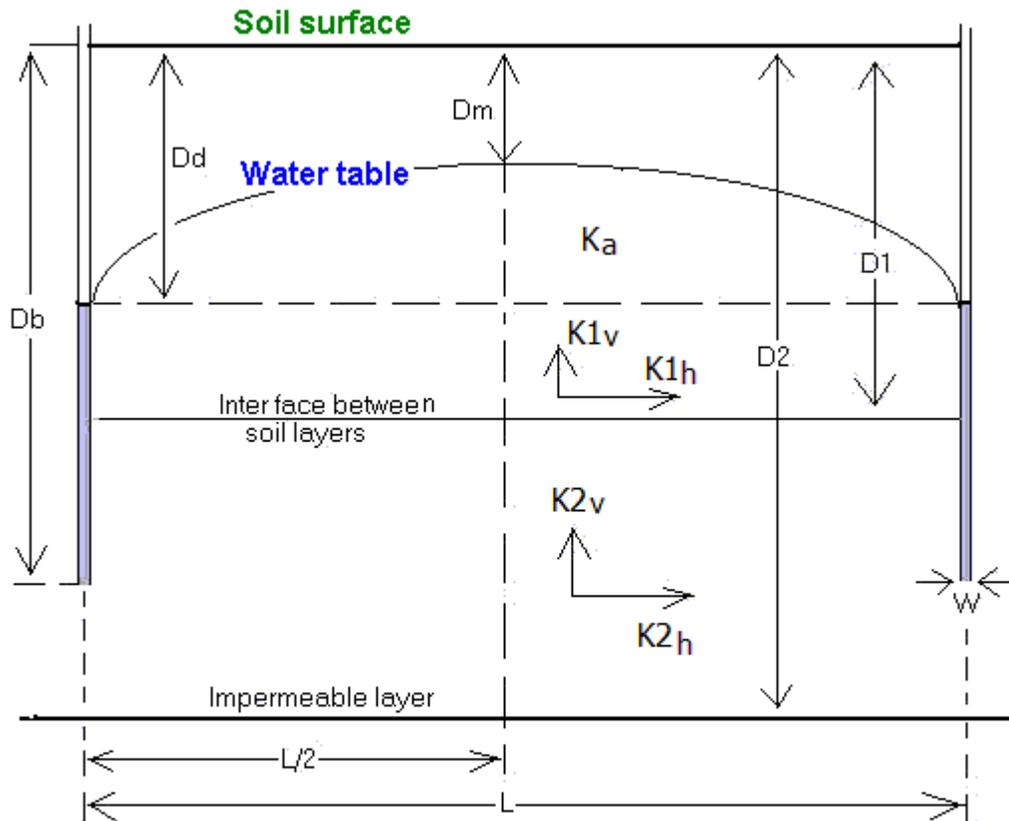
phenomenon is that only more soluble species can cover long distances. The retardation factor depends on the chemical nature of both the contaminant and the aquifer.

## Governing equations

### Darcy's Law

Darcy's law is a Constitutive equation (empirically derived by Henri Darcy, in 1856) that states the amount of groundwater discharging through a given portion of aquifer is proportional to the cross-sectional area of flow, the hydraulic head gradient, and the hydraulic conductivity.

### Groundwater flow equation



### Geometry well drainage system

$D$  = Depth     $K$  = hydraulic conductivity     $L$  = well spacing     $W$  = well diameter

Geometry of a partially penetrating well drainage system in an anisotropic layered aquifer

The groundwater flow equation, in its most general form, describes the movement of groundwater in a porous medium (aquifers and aquitards). It is known in mathematics as

the diffusion equation, and has many analogs in other fields. Many solutions for groundwater flow problems were borrowed or adapted from existing heat transfer solutions.

It is often derived from a physical basis using Darcy's law and a conservation of mass for a small control volume. The equation is often used to predict flow to wells, which have radial symmetry, so the flow equation is commonly solved in polar or cylindrical coordinates.

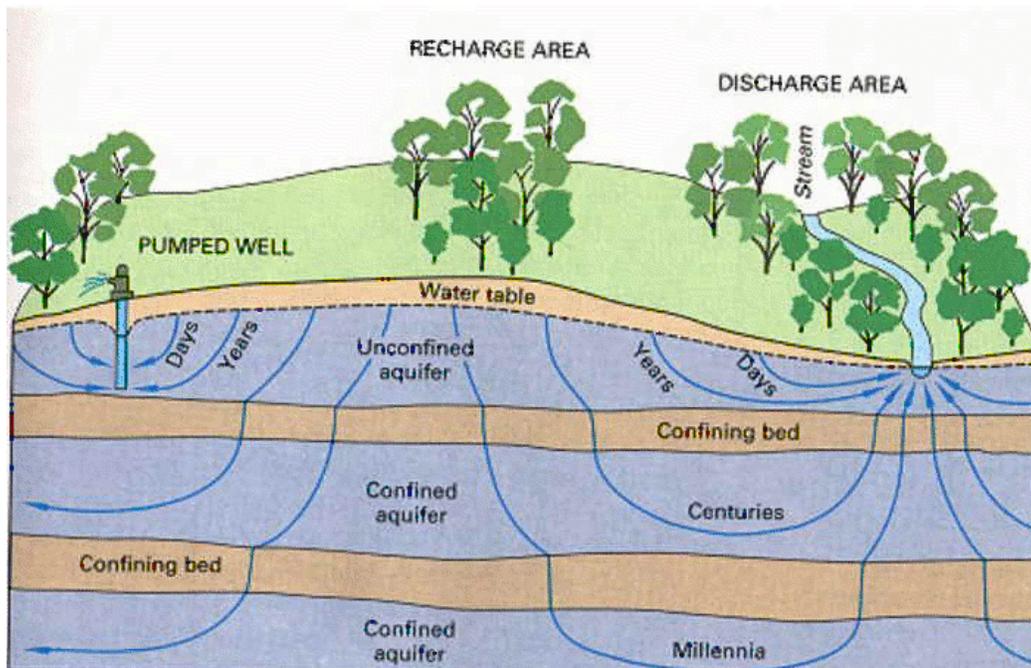
The Theis equation is one of the most commonly used and fundamental solutions to the groundwater flow equation; it can be used to predict the transient evolution of head due to the effects of pumping one or a number of pumping wells.

The Thiem equation is a solution to the steady state groundwater flow equation (Laplace's Equation) for flow to a well. Unless there are large sources of water nearby (a river or lake), true steady-state is rarely achieved in reality.

Both above equations are used in aquifer tests (pump tests).

The Hooghoudt equation is a groundwater flow equation applied to subsurface drainage by pipes, tile drains or ditches. An alternative subsurface drainage method is drainage by wells for which groundwater flow equations are also available.

## Calculation of groundwater flow



Relative groundwater travel times.

To use the groundwater flow equation to estimate the distribution of hydraulic heads, or the direction and rate of groundwater flow, this partial differential equation (PDE) must be solved. The most common means of analytically solving the diffusion equation in the hydrogeology literature are:

- Laplace, Hankel and Fourier transforms (to reduce the number of dimensions of the PDE),
- similarity transform (also called the Boltzmann transform) is commonly how the Theis solution is derived,
- separation of variables, which is more useful for non-Cartesian coordinates, and
- Green's functions, which is another common method for deriving the Theis solution — from the fundamental solution to the diffusion equation in free space.

No matter which method we use to solve the groundwater flow equation, we need both initial conditions (heads at time  $(t) = 0$ ) and boundary conditions (representing either the physical boundaries of the domain, or an approximation of the domain beyond that point). Often the initial conditions are supplied to a transient simulation, by a corresponding steady-state simulation (where the time derivative in the groundwater flow equation is set equal to 0).

There are two broad categories of how the (PDE) would be solved; either analytical methods, numerical methods, or something possibly in between. Typically, analytic methods solve the groundwater flow equation under a simplified set of conditions *exactly*, while numerical methods solve it under more general conditions to an *approximation*.

## **Analytic methods**

Analytic methods typically use the structure of mathematics to arrive at a simple, elegant solution, but the required derivation for all but the simplest domain geometries can be quite complex (involving non-standard coordinates, conformal mapping, etc.). Analytic solutions typically are also simply an equation that can give a quick answer based on a few basic parameters. The Theis equation is a very simple (yet still very useful) analytic solution to the groundwater flow equation, typically used to analyze the results of an aquifer test or slug test.

## **Numerical methods**

The topic of numerical methods is quite large, obviously being of use to most fields of engineering and science in general. Numerical methods have been around much longer than computers have (In the 1920s Richardson developed some of the finite difference schemes still in use today, but they were calculated by hand, using paper and pencil, by human “calculators”), but they have become very important through the availability of fast and cheap personal computers. A quick survey of the main numerical methods used in hydrogeology, and some of the most basic principles is shown below and further discussed in “Groundwater model”.

There are two broad categories of numerical methods: gridded or discretized methods and non-gridded or mesh-free methods. In the common finite difference method and finite element method (FEM) the domain is completely gridded (“cut” into a grid or mesh of small elements). The analytic element method (AEM) and the boundary integral equation method (BIEM — sometimes also called BEM, or Boundary Element Method) are only discretized at boundaries or along flow elements (line sinks, area sources, etc.), the majority of the domain is mesh-free.

### **General properties of gridded methods**

Gridded Methods like finite difference and finite element methods solve the groundwater flow equation by breaking the problem area (domain) into many small elements (squares, rectangles, triangles, blocks, tetrahedra, etc.) and solving the flow equation for each element (all material properties are assumed constant or possibly linearly variable within an element), then linking together all the elements using conservation of mass across the boundaries between the elements (similar to the divergence theorem). This results in a system which overall approximates the groundwater flow equation, but exactly matches the boundary conditions (the head or flux is specified in the elements which intersect the boundaries).

Finite differences are a way of representing continuous differential operators using discrete intervals ( $\Delta x$  and  $\Delta t$ ), and the finite difference methods are based on these (they are derived from a Taylor series). For example the first-order time derivative is often approximated using the following forward finite difference, where the subscripts indicate a discrete time location,

$$\frac{\partial h}{\partial t} = h'(t_i) \approx \frac{h_i - h_{i-1}}{\Delta t}.$$

The forward finite difference approximation is unconditionally stable, but leads to an implicit set of equations (that must be solved using matrix methods, e.g. LU or Cholesky decomposition). The similar backwards difference is only conditionally stable, but it is explicit and can be used to “march” forward in the time direction, solving one grid node at a time (or possibly in parallel, since one node depends only on its immediate neighbors). Rather than the finite difference method, sometimes the Galerkin FEM approximation is used in space (this is different from the type of FEM often used in structural engineering) with finite differences still used in time.

### **Application of finite difference models**

MODFLOW is a well-known example of a general finite difference groundwater flow model. It is developed by the US Geological Survey as a modular and extensible simulation tool for modeling groundwater flow. It is free software developed, documented and distributed by the USGS. Many commercial products have grown up around it, providing graphical user interfaces to its input file based interface, and typically incorporating pre- and post-processing of user data. Many other models have

been developed to work with MODFLOW input and output, making linked models which simulate several hydrologic processes possible (flow and transport models, surface water and groundwater models and chemical reaction models), because of the simple, well documented nature of MODFLOW.

### **Application of finite element models**

Finite Element programs are more flexible in design (triangular elements vs. the block elements most finite difference models use) and there are some programs available (SUTRA, a 2D or 3D density-dependent flow model by the USGS; Hydrus, a commercial unsaturated flow model; FEFLOW, a commercial modeling environment for subsurface flow, solute and heat transport processes; and COMSOL Multiphysics (FEMLAB) a commercial general modeling environment), but unless they are gaining in importance they are still not as popular in with practicing hydrogeologists as MODFLOW is. Finite element models are more popular in university and laboratory environments, where specialized models solve non-standard forms of the flow equation (unsaturated flow, density dependent flow, coupled heat and groundwater flow, etc.)

### **Application of finite volume models**

The finite volume method is a method for representing and evaluating partial differential equations as algebraic equations [LeVeque, 2002; Toro, 1999]. Similar to the finite difference method, values are calculated at discrete places on a meshed geometry. "Finite volume" refers to the small volume surrounding each node point on a mesh. In the finite volume method, volume integrals in a partial differential equation that contain a divergence term are converted to surface integrals, using the divergence theorem. These terms are then evaluated as fluxes at the surfaces of each finite volume. Because the flux entering a given volume is identical to that leaving the adjacent volume, these methods are conservative. Another advantage of the finite volume method is that it is easily formulated to allow for unstructured meshes. The method is used in many computational fluid dynamics packages.

**PORFLOW** software package is a comprehensive mathematical model for simulation of Ground Water Flow and Nuclear Waste Management developed by Analytic & Computational Research, Inc., ACRi]ACRi

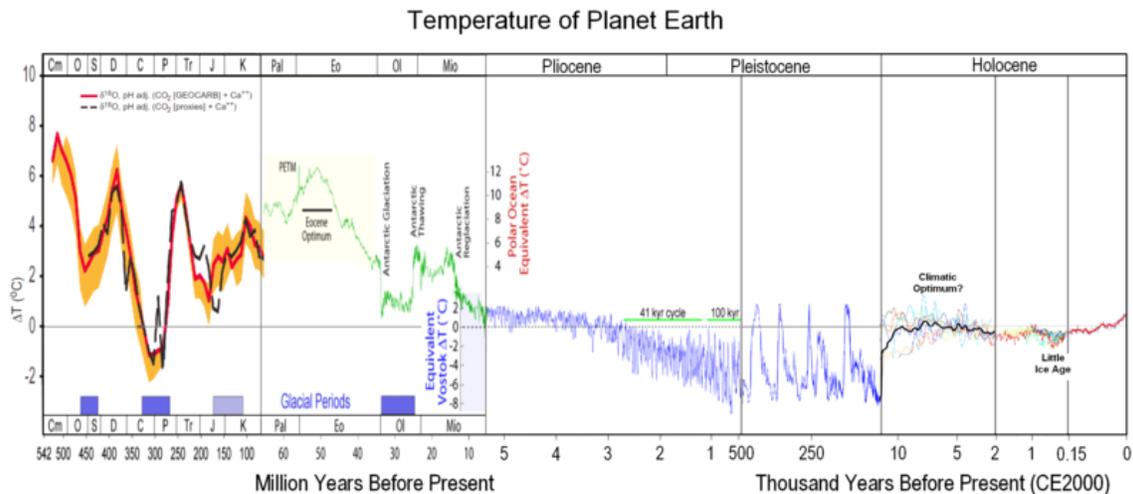
The **FEHM** software package is available free from Los Alamos National Laboratory and can be accessed at the FEHM Website. This versatile porous flow simulator includes capabilities to model multiphase, thermal, stress, and multicomponent reactive chemistry. Current work using this code includes simulation of methane hydrate formation, CO<sub>2</sub> sequestration, oil shale extraction, migration of both nuclear and chemical contaminants, environmental isotope migration in the unsaturated zone, and karst formation.

## **Other methods**

These include mesh-free methods like the Analytic Element Method (AEM) and the Boundary Element Method (BEM), which are closer to analytic solutions, but they do approximate the groundwater flow equation in some way. The BEM and AEM exactly solve the groundwater flow equation (perfect mass balance), while approximating the boundary conditions. These methods are more exact and can be much more elegant solutions (like analytic methods are), but have not seen as widespread use outside academic and research groups yet.

## Chapter 3

# Paleoclimatology



**Paleoclimatology** (also **Palaeoclimatology**) is the study of climate change taken on the scale of the entire history of Earth. It uses a variety of proxy records from ice sheets, tree rings, sediment, corals, shells and rocks to determine the past state of the climate system on Earth.

Paleoclimatology has wider implications for climate change today. Scientists often consider past changes in environment and biodiversity to reflect on the current situation, and specifically the impact of climate on mass extinctions and biotic recovery.

## Reconstructing ancient climates

Paleoclimatologists employ a wide variety of techniques to deduce ancient climates.

### Ice

Mountain Glaciers and the polar ice caps/ice sheets are a widely employed source of data in paleoclimatology. Recent ice coring projects in the ice caps of Greenland and Antarctica have yielded data going back several hundred thousand years—over 800,000 years in the case of the EPICA project.

- Air trapped within fallen snow becomes encased in tiny bubbles as the snow is compressed into ice in the glacier under the weight of later years' snow. This trapped air has proven a tremendously valuable source for direct measurement of the composition of air from the time the ice was formed.
- Layering can be observed due to seasonal pauses in ice accumulation and can be used to establish chronology; associating specific depths of the core with ranges of time.
- Changes in the layering thickness can be used to determine changes in precipitation or temperature.
- Oxygen-18 quantity changes ( $\delta^{18}\text{O}$ ) in ice layers represent changes in average ocean surface temperature. Water molecules containing the heavier O-18 evaporate at a higher temperature than water molecules containing the normal Oxygen-16 isotope. The ratio of O-18 to O-16 will be higher as temperature increases and less as temperature decreases. Various cycles in those isotope ratios have been detected.
- Pollen has been observed in the ice cores and can be used to understand which plants were present as the layer formed. Pollen is produced in abundance and its distribution is typically well understood. A pollen count for a specific layer can be produced by observing the total amount of pollen categorized by type (shape) in a controlled sample of that layer. Changes in plant frequency over time can be plotted through statistical analysis of pollen counts in the core. Knowing which plants were present leads to an understanding of precipitation and temperature, and types of fauna present. Palynology includes the study of pollen for these purposes.
- Volcanic ash is contained in some layers, and can be used to establish the time of the layer's formation. Each volcanic event distributed ash with a unique set of properties (shape and color of particles, chemical signature). Establishing the ash's source will establish a range of time to associate with layer of ice.

### Dendroclimatology

Climatic information can be obtained through an understanding of changes in tree growth. Generally, trees respond to changes in climatic variables by speeding up or slowing down growth, which in turn is generally reflected a greater or lesser thickness in growth rings. Different species, however, respond to changes in climatic variables in different ways. A tree-ring record is established by compiling information from many living trees in a specific area. Older intact wood that has escaped decay can extend the time covered by the record by matching the ring depth changes to contemporary specimens. Using this method some areas have tree-ring records dating back a few thousand years. Older wood not connected to a contemporary record can be dated generally with radiocarbon techniques. A tree-ring record can be used to produce information regarding precipitation, temperature, hydrology, and fire corresponding to a particular area.

On a longer time scale, geologists must refer to the sedimentary record for data.

## Sedimentary content

- Sediments, sometimes lithified to form rock, may contain remnants of preserved vegetation, animals, plankton or pollen, which may be characteristic of certain climatic zones.
- Biomarker molecules such as the alkenones may yield information about their temperature of formation.
- Chemical signatures, particularly Mg/Ca ratio of calcite in Foraminifera tests, can be used to reconstruct past temperature.
- Isotopic ratios can provide further information. Specifically, the  $\delta^{18}\text{O}$  record responds to changes in temperature and ice volume, and the  $\delta^{13}\text{C}$  record reflects a range of factors, which are often difficult to disentangle.

## Sedimentary facies

On a longer time scale, the rock record may show signs of sea level rise and fall; further, features such as “criticized” sand dunes can be identified. Scientists can get a grasp of long term climate by studying sedimentary rock going back billions of years. The division of earth history into separate periods is largely based on visible changes in sedimentary rock layers that demarcate major changes in conditions. Often these include major shifts in climate.

## Corals

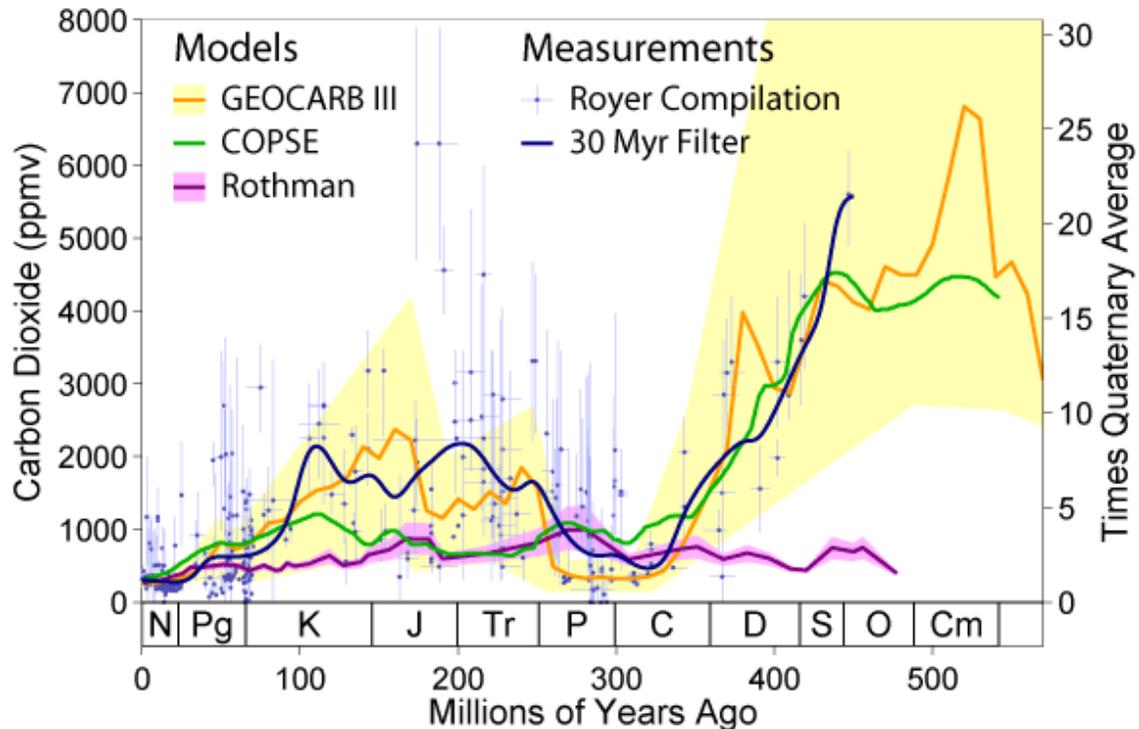
Coral “rings” are similar to tree rings, except they respond to different things, such as the water temperature and wave action. From this source, certain equipment can be used to derive the sea surface temperature and water salinity from the past few centuries. The  $\delta^{18}\text{O}$  of criticize red algae provides a useful proxy of sea surface temperature at high latitudes, where many traditional techniques are limited.

## Limitations

All records decrease in utility back in time. The oldest ice core taken was from the Antarctic and dates to 800,000 years old. An international effort is currently being made in the same location to core to 1.2 million years ago. The deep marine record, the source of most isotopic data, only exists on oceanic plates, which are eventually subducted — the oldest remaining material is 200 million years old. Older sediments are also more prone to corruption by diagenesis. Resolution and confidence in the data decrease over time.

# History of the atmosphere

## Phanerozoic Carbon Dioxide



Oxygen content of the atmosphere over the last billion years

### Earliest atmosphere

The outgassings of the Earth was stripped away by solar winds early in the history of the planet until a steady state was established, the first atmosphere. Based on today's volcanic evidence, this atmosphere would have contained 80% water vapor, 10% carbon dioxide, 5 to 7% hydrogen sulfide, and smaller amounts of nitrogen, carbon monoxide, hydrogen, methane and inert gases.

A major rainfall led to the buildup of a vast ocean, enriching the other agents, first carbon dioxide and later nitrogen and inert gases. A major part of carbon dioxide exhalations were soon dissolved in water and built up carbonate sediments.

### Second atmosphere

As early as 3.8 billion years ago, water related sediments have been found. About 3.4 billion years ago, nitrogen was the major part of the then stable second atmosphere. An influence of life has to be taken into account rather soon since hints on early life forms are to be found as early as 3.5 billion years ago. The fact that this is not in line with the

— compared to today 30% lower — solar radiance of the early sun has been described as the faint young Sun paradox.

The geological record, however, shows a continually relatively warm surface during the complete early temperature record of the earth with the exception of one cold glacial phase about 2.4 billion years ago. In the late Archean Era an oxygen containing atmosphere began to develop from photosynthesizing algae. The early basic carbon isotopy is very much in line with what is found today. As Jan Veizer assumed that not only did we have life as far back as we had rocks, but there was as much life then as today and the fundamental features of the carbon cycle were established as early as 4 billion years ago.

### **Third atmosphere**

The accretion of continents about 3.5 billion years ago added plate tectonics, constantly rearranging the continents and also shaping long-term climate evolution by allowing the transfer of carbon dioxide to large land-based carbonate storages. Free oxygen did not exist until about 1.7 billion years ago and this can be seen with the development of the red beds and the end of the banded iron formations. This signifies a shift from a reducing atmosphere to an oxidizing atmosphere. O<sub>2</sub> showed major ups and downs until reaching a steady state of more than 15%. The following time span was the Phanerozoic, during which oxygen-breathing metazoan life forms began to appear.

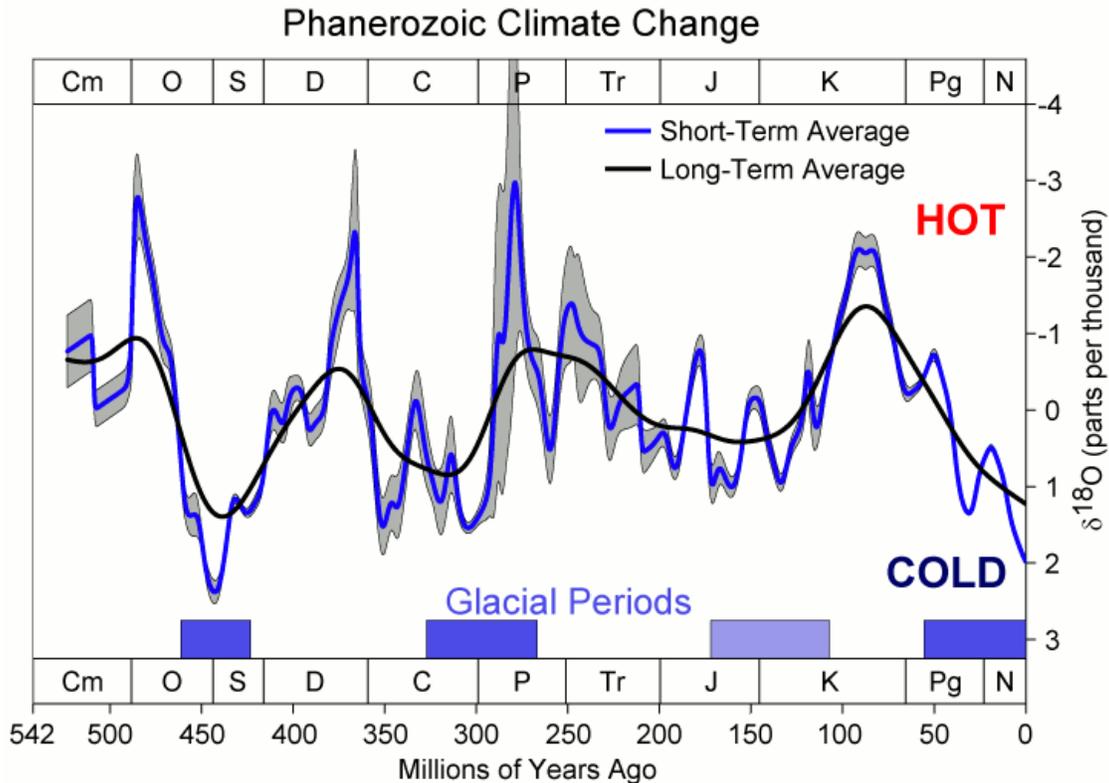
## **Climate during geological ages**

### **Precambrian climate**

In the first three quarters of the Earth's history, only one major glaciation is to be found in the geological record. Since about 950 million years ago, the Earth's climate has varied regularly between large-scale or just polar cap wide glaciation and extensively tropical climates. The time scale for this variation is roughly 140 million years and may be related to Earth's motion into and out of galactic spiral arms and compared to the previous time, significantly reduced solar wind.

The climate of the late Precambrian showed some major glaciation events spreading over much of the earth. At this time the continents were bunched up in the Rodinia supercontinent. Massive deposits of tillites are found and anomalous isotopic signatures are found, which gave rise to the Snowball Earth hypothesis. As the Proterozoic Eon drew to a close, the Earth started to warm up. By the dawn of the Cambrian and the Phanerozoic, life forms were abundant in the Cambrian explosion with average global temperatures of about 22 °C.

## Phanerozoic climate

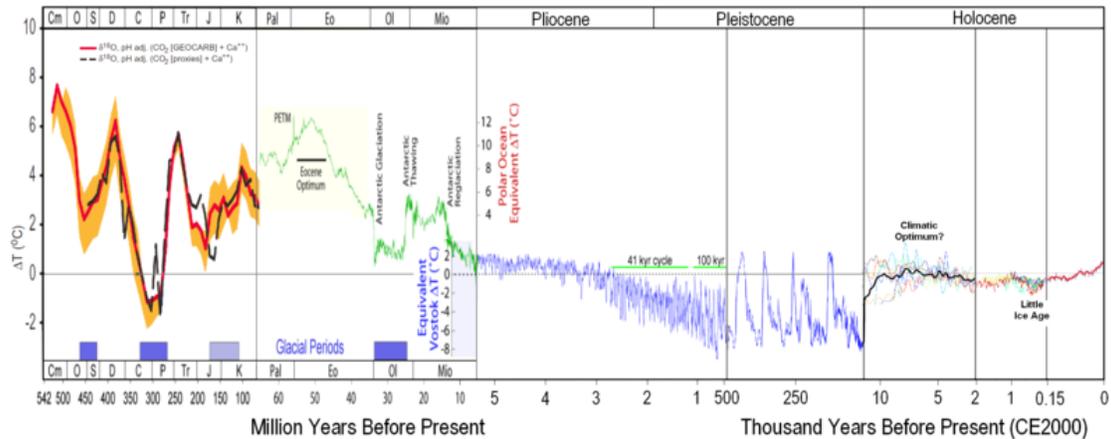


500 million years of climate change

Major drivers for the preindustrial ages have been variations of the sun, volcanic ashes and exhalations, relative movements of the earth towards the sun and tectonically induced effects as for major sea currents, watersheds and ocean oscillations. In the early Phanerozoic, increased atmospheric carbon dioxide concentrations have been linked to driving or amplifying increased global temperatures. Royer et al. 2004 found a climate sensitivity for the rest of the Phanerozoic which was calculated to be similar to today's modern range of values.

The difference in global mean temperatures between a fully glacial Earth and an ice free Earth is estimated at approximately 10 °C, though far larger changes would be observed at high latitudes and smaller ones at low latitudes. One requirement for the development of large scale ice sheets seems to be the arrangement of continental land masses at or near the poles. The constant rearrangement of continents by plate tectonics can also shape long-term climate evolution. However, the presence or absence of land masses at the poles is not sufficient to guarantee glaciations or exclude polar ice caps. Evidence exists of past warm periods in Earth's climate when polar land masses similar to Antarctica were home to deciduous forests rather than ice sheets.

## Temperature of Planet Earth



### Phanerozoic till today's temperature record

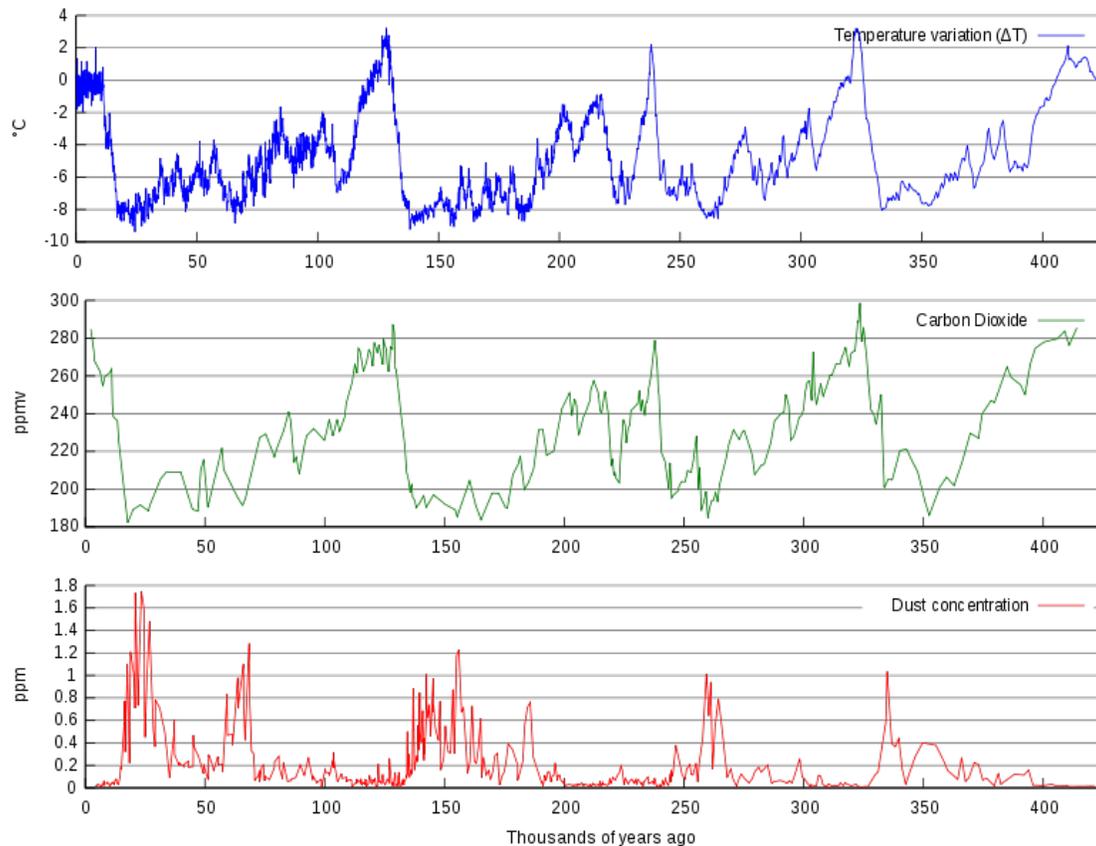
The relatively warm local minimum between Jurassic and Cretaceous goes along with widespread tectonic activity, e.g. the breakup of supercontinents.

Superimposed on the long-term evolution between hot and cold climates have been many short-term fluctuations in climate similar to, and sometimes more severe than, the varying glacial and interglacial states of the present ice age. Some of the most severe fluctuations, such as the Paleocene-Eocene Thermal Maximum, may be related to rapid climate changes due to sudden collapses of natural methane clathrate reservoirs in the oceans.

A similar, single event of induced severe climate change after a meteorite impact has been proposed as reason for the Cretaceous-Tertiary extinction event. Other major thresholds are the Permian-Triassic, and Ordovician-Silurian extinction events with various reasons suggested.

### Quaternary sub-era

The Quaternary sub-era includes the current climate. There has been a cycle of ice ages for the past 2.2–2.1 million years (starting before the Quaternary in the late Neogene Period).



Ice core data for the past 400,000 years. Note length of glacial cycles averages ~100,000 years. Blue curve is temperature, green curve is CO<sub>2</sub>, and red curve is windblown glacial dust (loess). Today's date is on the left side of the graph.

Note in the graphic on the right the strong 120,000-year periodicity of the cycles, and the striking asymmetry of the curves. This asymmetry is believed to result from complex interactions of feedback mechanisms. It has been observed that ice ages deepen by progressive steps, but the recovery to interglacial conditions occurs in one big step.

## Controlling Factors

### Short term (10<sup>4</sup> to 10<sup>6</sup> years)

Geologically short-term (<120,000 year) temperatures are believed to be driven by orbital factors amplified by changes in greenhouse gases. The arrangements of land masses on the Earth's surface are believed to influence the effectiveness of these orbital forcing effects.

## Medium term ( $10^6$ to $10^8$ years)

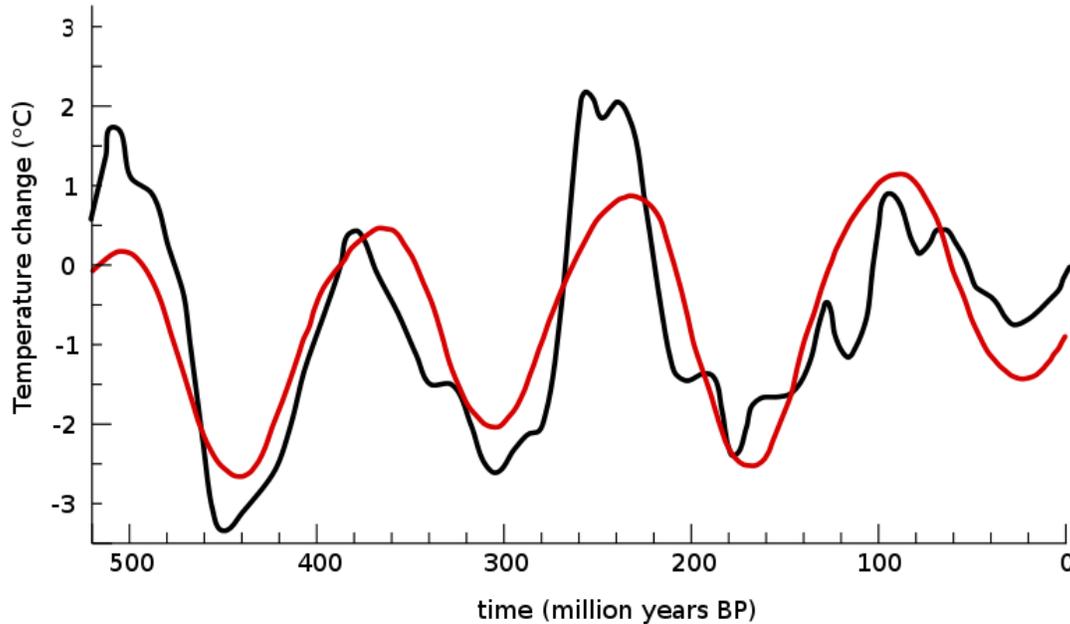
Continental drift affects the thermohaline circulation, which transfers heat between the equatorial regions and the poles, as does the extent of polar ice coverage.

The timing of ice ages throughout geologic history is in part controlled by the position of the continental plates on the surface of the Earth. When landmasses are concentrated near the polar regions, there is an increased chance for snow and ice to accumulate. Small changes in solar energy can tip the balance between summers in which the winter snow mass completely melts and summers in which the winter snow persists until the following winter.

Comparisons of plate tectonic continent reconstructions and paleoclimatic studies show that the Milankovitch cycles have the greatest effect during geologic eras when landmasses have been concentrated in polar regions, as is the case today. Today, Greenland, Antarctica, and the northern portions of Europe, Asia, and North America are situated such that a minor change in solar energy will tip the balance between year-round snow/ice preservation and complete summer melting. The presence of snow and ice is a well-understood positive feedback mechanism for climate. The Earth today is considered to be prone to ice age glaciations.

Another proposed factor in long term temperature change is the **Uplift-Weathering Hypothesis**, first put forward by T. C. Chamberlin in 1899 and later independently proposed in 1988 by Maureen Raymo and colleagues, where upthrusting mountain ranges expose minerals to weathering resulting in their chemical conversion to carbonates thereby removing  $\text{CO}_2$  from the atmosphere and cooling the earth. Others have proposed similar effects due to changes in average water table levels and consequent changes in sub-surface biological activity and PH levels.

## Long term ( $10^8$ to $10^9$ years)



### **Correlation between variations in cosmic ray flux (red) and change in sea temperature (black).** Data as presented by Shaviv & Veizer.

It has been proposed that long term galactic motions of the sun have a major influence on the Earth's climate. There are two principal motions, the first and most significant is the orbit of the sun around the galactic centre with a period of the order of 240 million years. Since this period is different from the rotation period of the galactic spiral arms, the sun, and the earth with it, will periodically pass through the arms (estimates of the period are uncertain and vary from 143 million years to 176 million years). The second is an oscillatory bobbing motion, similar to a floating buoy, which will periodically take the sun through the galactic disc. The period of this bobbing motion is 67 million years, so a pass through the galactic plane will occur every 33 million years. The causal link between these galactic motions and climate is unclear but one (controversial) postulate is the effect that entering a denser region of the galaxy will have on increasing the cosmic ray flux (CRF). This theory has been criticized, both for overstating the correlation with CRF and for failing to propose a believable mechanism that would allow CRF to drive temperature. The claims by Henrik Svensmark that CRF also strongly affects short term climate changes is even more controversial and has been challenged by many.

It has also been suggested that there is some correlation between these galactic cycles and geological periods. The reason for this is postulated to be that the earth experiences many more impact events while passing through high density regions of the galaxy. Both the climate changes and sudden impacts may cause, or contribute to, extinction events.

## **Very long term ( $10^9$ years or more)**

Jan Veizer and Nir Shaviv have proposed the interaction of cosmic rays, solar wind and the various magnetic fields to explain the long term evolution of earths climate.

According to Shaviv, the early sun had emitted a stronger solar wind with a protective effect against cosmic rays. In that early age, a moderate greenhouse effect comparable to today's would have been sufficient to explain an ice free earth and the faint young sun paradox. The solar minimum around 2.4 billion years ago is consistent with an established cosmic ray flux modulation by a variable star formation rate in the Milky Way and there is also a hint of an extinction event at this time. Within the last billion years the solar wind has significantly diminished. It is only within this more recent time that passages of the heliosphere through the spiral arms of the galaxy have been able to gain a strong and regularly modulating influence as described above.

Over the very long term the energy output of the sun has gradually increased, on the order of 5% per billion ( $10^9$ ) years, and will continue to do so until it reaches the end of its current phase of stellar evolution.

## Chapter 4

# Biosphere & its Components

## Biogeography

**Biogeography** is the study of the distribution of species spatially and temporally. Over areal ecological changes, it is also tied to the concepts of species and their past, or present living 'refugium', their survival locales, or their interim living sites. It aims to reveal where organisms live, and at what abundance. As writer David Quammen put it, "...biogeography does more than ask *Which species?* and *Where*. It also asks *Why?* and, what is sometimes more crucial, *Why not?*." The patterns of species distribution across geographical areas can usually be explained through a combination of historical factors such as speciation, extinction, continental drift, glaciation (and associated variations in sea level, river routes, and so on), and river capture, in combination with the area and isolation of landmasses (geographic constraints) and available energy supplies.

Modern biogeography often employs the use of Geographic Information Systems (GIS), to understand the factors affecting organism distribution, and to predict future trends in organism distribution. Often mathematical models and GIS are employed to solve ecological problems that have a spatial aspect to them.

## History

The scientific theory of biogeography grows out of the work of Hewett Cottrell Watson (1804–1881), Alfred Russel Wallace (1823–1913) and other early evolutionary scientists. Wallace studied the distribution of flora and fauna in the Malay Archipelago in the 19th century. With the exception of Wallace and a few others, prior to the publication of *The Theory of Island Biogeography* by Robert MacArthur and E.O. Wilson in 1967 the field of biogeography was seen as a primarily historical one and as such the field was seen as a purely descriptive one.

MacArthur and Wilson changed this perception, and showed that the species richness of an area could be predicted in terms of such factors as habitat area, immigration rate and extinction rate. This gave rise to an interest in island biogeography. The application of island biogeography theory to habitat fragments spurred the development of the fields of conservation biology and landscape ecology.

Classic biogeography has been expanded by the development of molecular systematics, creating a new discipline known as phylogeography. This development allowed scientists to test theories about the origin and dispersal of populations, such as island endemics. For example, while classic biogeographers were able to speculate about the origins of species in the Hawaiian Islands, phylogeography allows them to test theories of relatedness between these populations and putative source populations in Asia and North America.

## **Paleobiogeography**

Paleobiogeography goes one step further to include paleogeographic data and considerations of plate tectonics. Using molecular analyses and corroborated by fossils, it has been possible to demonstrate that perching birds evolved first in the region of Australia or the adjacent Antarctic (which at that time lay somewhat further north and had a temperate climate). From there, they spread to the other Gondwanan continents and Southeast Asia - the part of Laurasia then closest to their origin of dispersal – in the late Paleogene, before achieving a global distribution in the early Neogene. Not knowing the fact that at the time of dispersal, the Indian Ocean was much narrower than it is today, and that South America was closer to the Antarctic, one would be hard pressed to explain the presence of many "ancient" lineages of perching birds in Africa, as well as the mainly South American distribution of the suboscines.

## **Classification**

Biogeography is a synthetic science, related to geography, biology, soil science, geology, climatology, ecology and evolution.

Some fundamental concepts in biogeography include:

- evolution (change in genetic composition of a population)
- extinction (disappearance of a species)
- dispersal (movement of populations away from their point of origin, related to migration)
- geodispersal the erosion of barriers to biotic dispersal and gene flow, that permit range expansion and the merging of previously isolated biotas.
- range and distribution
- endemic areas
- vicariance the formation of barriers to biotic dispersal and gene flow, that tend to subdivide species and biotas, leading to speciation and extinction.

# Geomicrobiology

**Geomicrobiology** is the result of the combination of geology and microbiology. The field of geomicrobiology concerns the role of microbe and microbial processes in geological and geochemical processes and vice-versa. The field is especially important when dealing with microorganisms in aquifers and public drinking water supplies.

Another area of investigation in geomicrobiology is the study of extremophile organisms, the microorganisms that thrive in environments normally considered hostile. Such environments may include extremely hot (hot springs or mid-ocean ridge black smoker) environments, extremely saline environments, or even space environments such as Martian soil or comets.

Recent observations and research in hyper-saline lagoon environments in Brazil and Australia have shown that anaerobic sulfate-reducing bacteria may be directly involved in the formation of dolomite. This suggests the alteration and replacement of limestone sediments by dolomitization in ancient rocks was possibly aided by ancestors to these anaerobic bacteria.

Some bacteria use metal ions as their energy source. They convert (or chemically reduce) the dissolved metal ions from one electrical state to another. This reduction releases energy for the bacteria's use, and, as a side product, serves to concentrate the metals into what ultimately become *ore deposits*. Certain iron, uranium and even gold ores are thought to have formed as the result of microbe action.

Microbes are being studied and used to degrade organic and even nuclear waste pollution and assist in environmental cleanup. (It should be stressed however that microorganisms can not decrease the total radioactivity of nuclear waste.)

An application of geomicrobiology is bioleaching, the use of microbes to extract metals from mine waste.

# Geoarchaeology





ge archaeologist at work on column sample

**Geoarchaeology** is a multi disciplinary approach rather than a discipline and uses the techniques and subject matter of geography, geology and other Earth sciences to examine topics which inform archaeological knowledge and thought.

Geoarchaeologists study the natural physical processes that affect archaeological sites such as geomorphology, the formation of sites through geological processes and the effects on buried sites and artifacts post-deposition.

Geoarchaeologists' work frequently involves studying soil and sediments as well as other geographical concepts to contribute an archaeological study.

Geoarchaeology is a recent field of research that uses the computer cartography, the Geographic Information System (GIS) and the Digital Elevation Models (D.E.M.) in combination with disciplines from Human and Social Sciences and Earth Sciences.

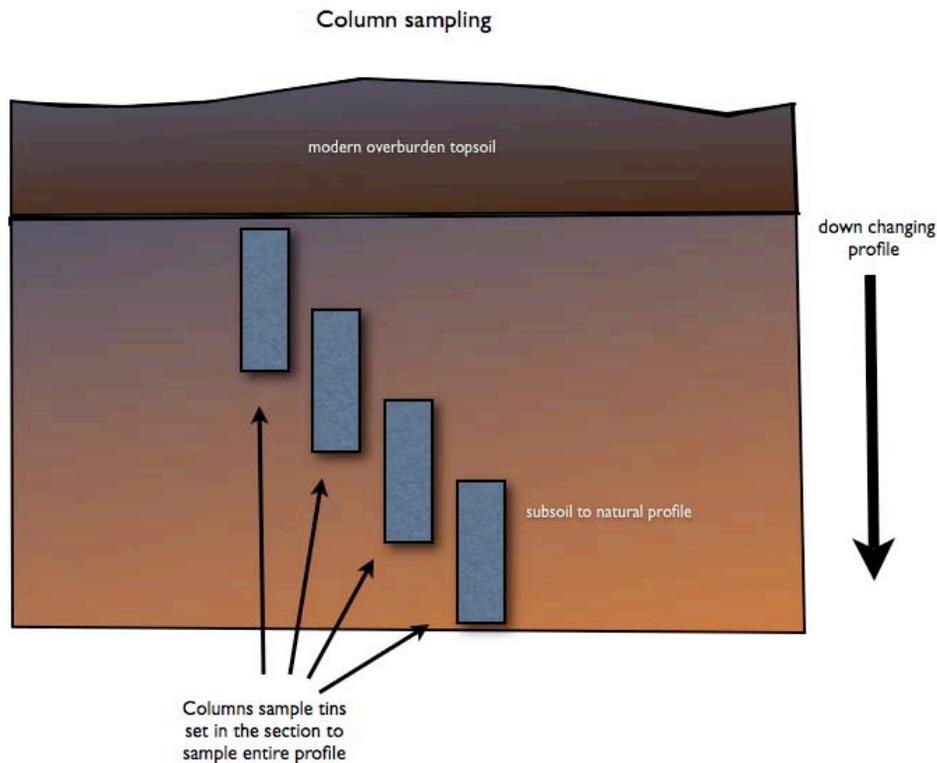
## **Techniques used in geoarchaeology**

### **Column sampling**

Column sampling is a technique of collecting samples from a section for analyzing and detecting the buried processes down the profile of the section. Narrow metal tins are bashed into the section in a series to collect the complete profile for study. If more than one tin is needed they are arranged offset and overlapping to one side so the complete profile can be rebuilt offsite in laboratory conditions.

### **Loss on ignition testing**

Loss on ignition testing for soil organic content.- a technique of measuring organic content in soil samples. Samples taken from a known place in the profile collected by column sampling are weighed then placed in a fierce oven which burns off the organic content. The resulting cooked sample is weighed again and the resulting loss in weight is an indicator of organic content in the profile at a certain depth. These readings are often used to detect buried soil horizons. A buried soil's horizons may not be visible in section and this horizon is an indicator of possible occupation levels. Ancient land surfaces especially from the prehistoric era can be difficult to discern so this technique is useful for evaluating an areas potential for prehistoric surfaces and archaeological evidence. Comparative measurements down the profile are made and a sudden rise in organic content at some point in the profile combined with other indicators is strong evidence for buried surfaces.



offset column sampling of the soil profile

## Magnetic susceptibility analysis

The magnetic susceptibility of a material is a measure of its ability to become magnetised by an external magnetic field (Dearing, 1999). The magnetic susceptibility of a soil reflects the presence of magnetic iron-oxide minerals such as maghaematite; just because a soil contains a lot of iron does not mean that it will have high magnetic susceptibility. Magnetic forms of iron can be formed by burning and microbial activity such as occurs in top soils and some anaerobic deposits. Magnetic iron compounds can also be found in igneous and metamorphic rocks.

The relationship between iron and burning means that magnetic susceptibility is often used for:

- Site prospection, to identify areas of archaeological potential prior to excavation.
- Identifying hearth areas and the presence of burning residues in deposits (Tite and Mullins, 1971).
- Explaining whether areas of reddening are due to burning or other natural processes such as gleying (waterlogging).

The relationship between soil formation and magnetic susceptibility means that it can also be used to:

- Identify buried soils in depositional sequences.
- Identify redeposited soil materials in peat, lake sediments etc.

## **Phosphate and orthophosphate content with spectrophotometry**

Phosphate in man-made soils derives from people, their animals, rubbish and bones. 100 people excrete about 62kg of phosphate annually, with about the same from their rubbish. Their animals excrete even more. A human body contains about 650g of PO<sub>4</sub>, (500g-80% in the skeleton), which results in elevated levels in burial sites. Most is quickly immobilised on the clay of the soil and 'fixed', where it can persist for thousands of years. For a 1 ha site this corresponds to about 150kg PO<sub>4</sub> ha<sup>-1</sup>yr<sup>-1</sup> about 0.5% to 10% of that already present in most soils. Therefore it doesn't take long for human occupation to make orders of magnitude differences to the phosphate concentration in soil. Phosphorus exist in different 'pools' in the soil 1) organic (available), 2) occluded (adsorbed), 3) bound (chemically bound). Each of these pools can be extracted using progressively more aggressive chemicals. Some workers (Eidt especially), think that the ratios between these pools can give information about past land use, and perhaps even dating. Whatever the method of getting the phosphorus from the soil into solution, the method of detecting it is usually the same. This uses the 'molybdate blue' reaction, where the depth of the colour is proportional to phosphorus concentration. In the lab, this is measured using a colorimeter, where light shining through a standard cell produces an electrical current proportional to the light attenuation. In the field, the same reaction is used on detector sticks, which are compared to a colour chart.

Phosphate concentrations can be plotted on archaeological plans to show former activity areas, and is also used to prospect for sites in the wider landscape.

## **Particle size analysis**

The particle size distribution of a soil sample may indicate the conditions under which the strata or sediment were deposited. Particle sizes are generally separated by means of dry or wet sieving (coarse samples such as till, gravel and sands, sometimes coarser silts) or by measuring the changes of the density of a dispersed solution (in sodiumpyrophosphate, for example)) of the sample (finer silts, clays). A rotating clock-glass with a very fine-grained dispersed sample under a heat lamp is useful in separating particles.

The results are plotted on curves which can be analyzed with statistical methods for particle distribution and other parameters.

The fractions received can be further investigated for cultural indicators, macro- and microfossils and other interesting features, so particle size analysis is in fact the first thing to do when handling these samples.

## **Trace Element Geochemistry**

Trace element geochemistry is the study of the abundances of elements in geological materials that do not occur in a large quantity in these materials. Because these trace elements concentration are determined by a large number of particular situations under which a certain geological material is formed, they are usually unique between two locations which contain the same type of rock or other geological material.

Geoarchaeologists use this uniqueness in trace element geochemistry to trace ancient patterns of resource-acquisition and trade. For example, researchers can look at the trace element composition of obsidian artifacts in order to "fingerprint" those artifacts. They can then study the trace element composition of obsidian outcrops in order to determine the original source of the raw material used to make the artifact.

## **Clay mineralogy analysis**

Geoarchaeologists study the mineralogical characteristics of pots through macroscopic and microscopic analyses. They can use these characteristics to understand the various manufacturing techniques used to make the pots, and through this, to know which production centers likely made these pots. They can also use the mineralogy to trace the raw materials used to make the pots to specific clay deposits.

## **Ostracode analysis**

Naturally occurring Ostracods in freshwater bodies are impacted by changes in salinity and pH due to human activities. Analysis of Ostracod shells in sediment columns show the changes brought about by farming and habitation activities. This record can be correlated with age dating techniques to help identify changes in human habitation patterns and population migrations.

## **Archaeological geology**

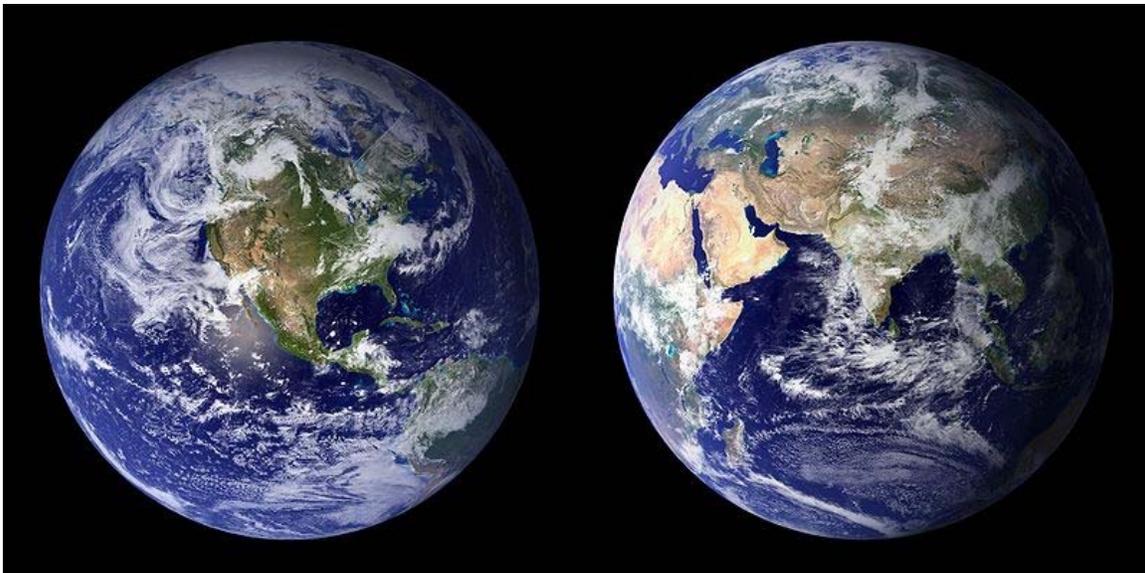
**Archaeological geology** (term coined by Werner Kasig, 1980), is a sub-field of geology which emphasises the value of earth constituents for human life.

Over the last decades, archaeologists and historians have faced the necessity to reconstruct ancient settlement history not only through the study of the material excavated, but also with the use of palaeo-environmental parameters.

## Chapter 5

# Earth Systems & Sciences

## Environmental science



Blue Marble composite images generated by NASA in 2001 (left) and 2002 (right).

**Environmental science** is an interdisciplinary academic field that integrates physical and biological sciences, (including physics, chemistry, biology, soil science, geology, and geography) to the study of the environment, and the solution of environmental problems. Environmental science provides an integrated, quantitative, and interdisciplinary approach to the study of environmental systems.

Related areas of study include environmental studies and environmental engineering. Environmental studies incorporates more of the social sciences for understanding human relationships, perceptions and policies towards the environment. Environmental engineering focuses on design and technology for improving environmental quality.

Environmental scientists work on subjects like the understanding of earth processes, evaluating alternative energy systems, pollution control and mitigation, natural resource

management, and the effects of global climate change. Environmental issues almost always include an interaction of physical, chemical, and biological processes. Environmental scientists bring a systems approach to the analysis of environmental problems. Key elements of an effective environmental scientist include the ability to relate space, and time relationships as well as quantitative analysis.

Environmental science came alive as a substantive, active field of scientific investigation in the 1960s and 1970s driven by (a) the need for a multi-disciplinary approach to analyze complex environmental problems, (b) the arrival of substantive environmental laws requiring specific environmental protocols of investigation and (c) the growing public awareness of a need for action in addressing environmental problems. Events that spurred this development included the publication of Rachael Carson's landmark environmental book *Silent Spring* along with major environmental issues becoming very public, such as the 1969 Santa Barbara oil spill, and the Cuyahoga River of Cleveland, Ohio, "catching fire" (also in 1969), and helped increase the visibility of environmental issues and create this new field of study.

## Components

**Atmospheric sciences** focuses on the Earth's atmosphere, with an emphasis upon its interrelation to other systems. Atmospheric sciences can include studies of meteorology, greenhouse gas phenomena, atmospheric dispersion modeling of airborne contaminants, sound propagation phenomena related to noise pollution, and even light pollution.

Taking the example of the global warming phenomena, physicists create computer models of atmospheric circulation and infra-red radiation transmission, chemists examine the inventory of atmospheric chemicals and their reactions, biologists analyze the plant and animal contributions to carbon dioxide fluxes, and specialists such as meteorologists and oceanographers add additional breadth in understanding the atmospheric dynamics.

**Ecology.** An interdisciplinary analysis of an ecological system which is being impacted by one or more stressors might include several related environmental science fields. For example, one might examine an estuarine setting where a proposed industrial development could impact certain species by water and air pollution. For this study, biologists would describe the flora and fauna, chemists would analyze the transport of water pollutants to the marsh, physicists would calculate air pollution emissions and geologists would assist in understanding the marsh soils and bay muds.

**Environmental chemistry** is the study of chemical alterations in the environment. Principal areas of study include soil contamination and water pollution. The topics of analysis include chemical degradation in the environment, multi-phase transport of chemicals (for example, evaporation of a solvent containing lake to yield solvent as an air pollutant), and chemical effects upon biota.

As an example study, consider the case of a leaking solvent tank which has entered the habitat soil of an endangered species of amphibian. As a method to resolve or understand

the extent of soil contamination and subsurface transport of solvent, a computer model would be implemented. Chemists would then characterize the molecular bonding of the solvent to the specific soil type, and biologists would study the impacts upon soil arthropods, plants, and ultimately pond-dwelling organisms that are the food of the endangered amphibian.

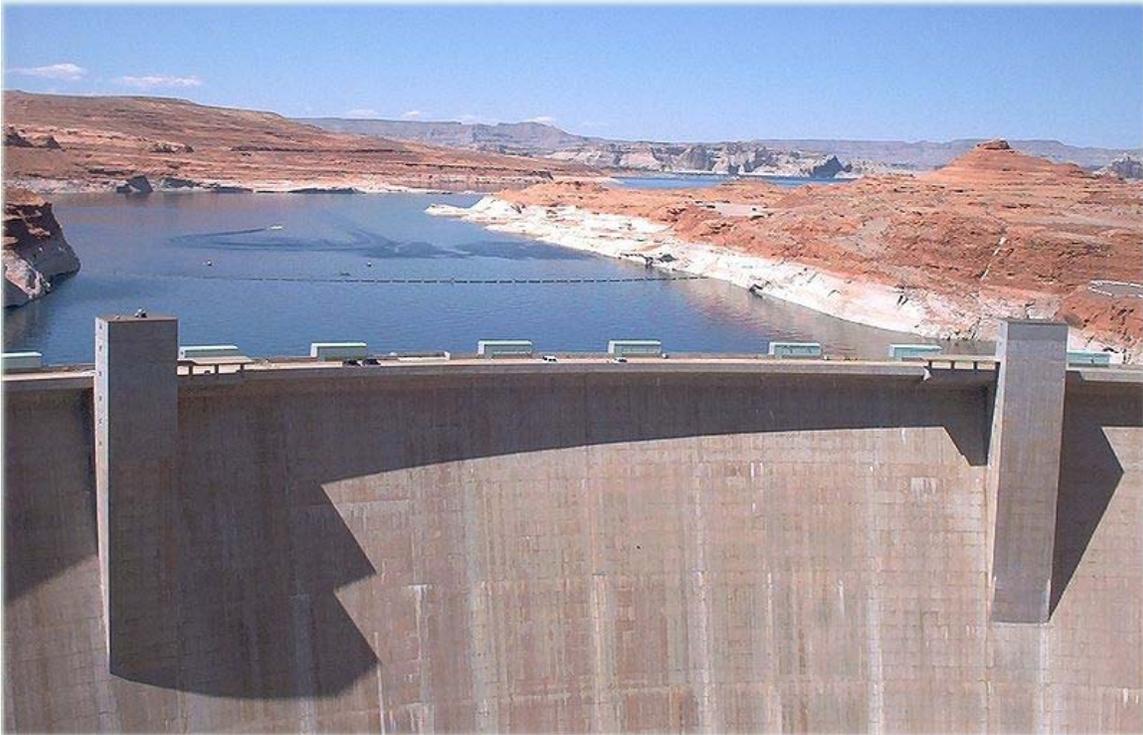
**Geosciences** include environmental geology, environmental soil science, volcanic phenomena and evolution of the Earth's crust. In some classification systems this can also include hydrology, including oceanography.

As an example study of soils erosion, calculations would be made of surface runoff by soil scientists. Fluvial geomorphologists would assist in examining sediment transport in overland flow. Physicists would contribute by assessing the changes in light transmission in the receiving waters. Biologists would analyze subsequent impacts to aquatic flora and fauna from increases in water turbidity.



Open-pit coal mining at Garzweiler, Germany

## Regulations driving the studies



Environmental science examines the effects of humans on nature (Glen Canyon Dam in the U.S.)

In the U.S. the National Environmental Policy Act (NEPA) of 1969 set forth requirements for analysis of major projects in terms of specific environmental criteria. Numerous state laws have echoed these mandates, applying the principles to local-scale actions. The upshot has been an explosion of documentation and study of environmental consequences before the fact of development actions.

One can examine the specifics of environmental science by reading examples of Environmental Impact Statements prepared under NEPA such as: *Wastewater treatment expansion options discharging into the San Diego/Tijuana Estuary*, *Expansion of the San Francisco International Airport*, *Development of the Houston, Metro Transportation system*, *Expansion of the metropolitan Boston MBTA transit system*, and *Construction of Interstate 66 through Arlington, Virginia*.

In England and Wales the Environment Agency (EA), formed in 1996, is a public body for protecting and improving the environment and enforces the regulations listed on the communities and local government site. (formerly the office of the deputy prime minister). The agency was set up under the Environment Act 1995 as an independent body and works closely with UK Government to enforce the regulations.

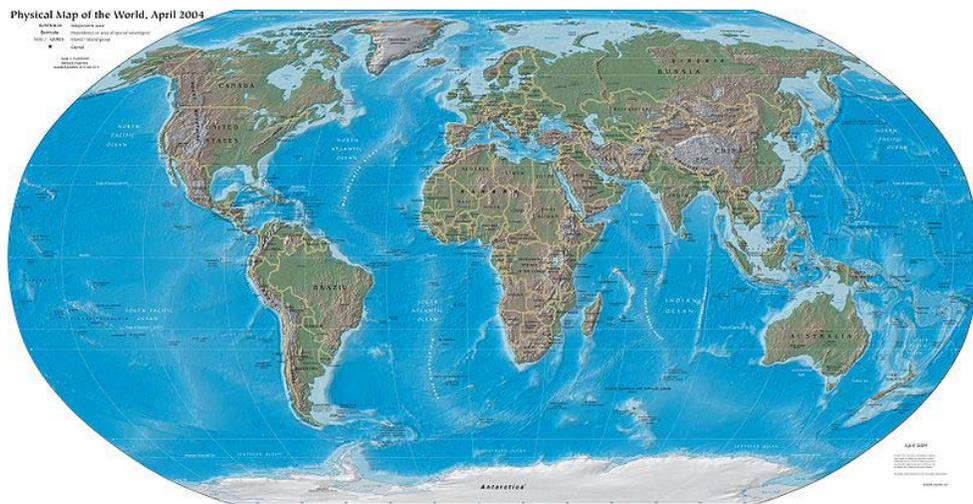
# Terminology

In common usage, "environmental science" and "ecology" are often used interchangeably, but technically, ecology refers only to the study of organisms and their interactions with each other and their environment. Ecology could be considered a subset of environmental science, which also could involve purely chemical or public health issues (for example) ecologists would be unlikely to study. In practice, there is considerable overlap between the work of ecologists and other environmental scientists.

The National Center for Education Statistics in the United States defines an academic program in environmental science as follows:

A program that focuses on the application of biological, chemical, and physical principles to the study of the physical environment and the solution of environmental problems, including subjects such as abating or controlling environmental pollution and degradation; the interaction between human society and the natural environment; and natural resources management. Includes instruction in biology, chemistry, physics, geosciences, climatology, statistics, and mathematical modeling.

# Geography



Map of the Earth

**Geography** is the study of the Earth and its lands, features, inhabitants, and phenomena. A literal translation would be "to describe or write about the Earth". The first person to use the word "geography" was Eratosthenes (276-194 B.C.). Four historical traditions in geographical research are the spatial analysis of natural and human phenomena (geography as a study of distribution), area studies (places and regions), study of man-

land relationship, and research in earth sciences. Nonetheless, modern geography is an all-encompassing discipline that foremost seeks to understand the Earth and all of its human and natural complexities—not merely where objects are, but how they have changed and come to be. Geography has been called 'the world discipline'. As "the bridge between the human and physical sciences," geography is divided into two main branches—human geography and physical geography.

## Introduction

Traditionally, geographers have been viewed the same way as cartographers and people who study place names and numbers. Although many geographers are trained in toponymy and cartology, this is not their main preoccupation. Geographers study the spatial and temporal distribution of phenomena, processes and features as well as the interaction of humans and their environment. As space and place affect a variety of topics such as economics, health, climate, plants and animals, geography is highly interdisciplinary.

“ ...mere names of places...are not geography...know by heart a whole gazetteer full of them would not, in itself, constitute anyone a geographer. Geography has higher aims than this: it seeks to classify phenomena (alike of the natural and of the political world, in so far as it treats of the latter), to compare, to generalize, to ascend from effects to causes, and, in doing so, to trace out the great laws of nature and to mark their influences upon man. This is 'a description of the world'—that is Geography. In a word Geography is a Science—a thing not of mere names but of argument and reason, of cause and effect. ”

— William Hughes, 1863

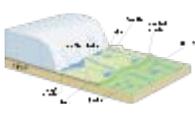
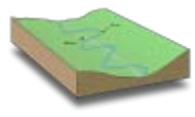
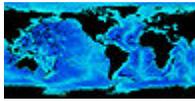
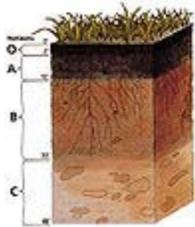
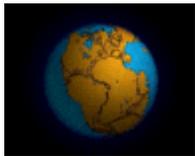
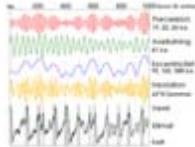
Geography as a discipline can be split broadly into two main subsidiary fields: human geography and physical geography. The former focuses largely on the built environment and how space is created, viewed and managed by humans as well as the influence humans have on the space they occupy. The latter examines the natural environment and how the climate, vegetation & life, soil, water, and landforms are produced and interact. As a result of the two subfields using different approaches a third field has emerged, which is environmental geography. Environmental geography combines physical and human geography and looks at the interactions between the environment and humans.

## Branches

### Physical geography

Physical geography (or physiography) focuses on geography as an Earth science. It aims to understand the physical lithosphere, hydrosphere, atmosphere, pedosphere, and global

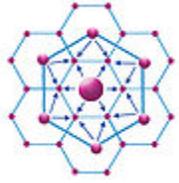
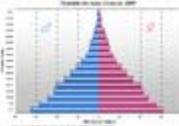
flora and fauna patterns (biosphere). Physical geography can be divided into the following broad categories:

			
Biogeography	Climatology & paleoclimatology	Coastal geography	Env. geog. & management
			
Geodesy	Geomorphology	Glaciology	Hydrology & Hydrography
			
Landscape ecology	Oceanography	Pedology	Palaeogeography
			
Quaternary science			

## Human geography

Human geography is a branch of geography that focuses on the study of patterns and processes that shape human interaction with various environments. It encompasses human, political, cultural, social, and economic aspects. While the major focus of human

geography is not the physical landscape of the Earth, it is hardly possible to discuss human geography without referring to the physical landscape on which human activities are being played out, and environmental geography is emerging as a link between the two. Human geography can be divided into many broad categories, such as:

			
Cultural geography	Development geography	Economic geography	Health geography
			
Historical & Time geog.	Political geog. & Geopolitics	Pop. geog. or Demography	Religion geography
			
Social geography	Transportation geography	Tourism geography	Urban geography

Various approaches to the study of human geography have also arisen through time and include:

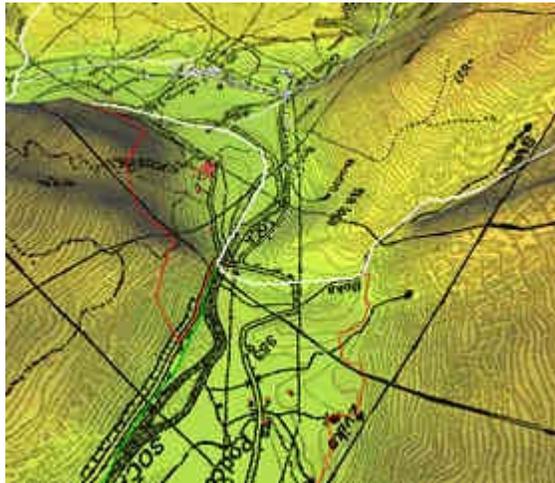
- Behavioral geography
- Feminist geography
- Culture theory
- Geosophy

## Environmental geography

Environmental geography is the branch of geography that describes the spatial aspects of interactions between humans and the natural world. It requires an understanding of the traditional aspects of physical and human geography, as well as the ways in which human societies conceptualize the environment.

Environmental geography has emerged as a bridge between human and physical geography as a result of the increasing specialisation of the two sub-fields. Furthermore, as human relationship with the environment has changed as a result of globalization and technological change a new approach was needed to understand the changing and dynamic relationship. Examples of areas of research in environmental geography include emergency management, environmental management, sustainability, and political ecology.

## Geomatics



Digital Elevation Model (DEM)

Geomatics is a branch of geography that has emerged since the quantitative revolution in geography in the mid 1950s. Geomatics involves the use of traditional spatial techniques used in cartography and topography and their application to computers. Geomatics has become a widespread field with many other disciplines using techniques such as GIS and remote sensing. Geomatics has also led to a revitalization of some geography departments especially in Northern America where the subject had a declining status during the 1950s.

Geomatics encompasses a large area of fields involved with spatial analysis, such as Cartography, Geographic information systems (GIS), Remote sensing, and Global positioning systems (GPS).

## Regional geography

Regional geography is a branch of geography that studies the regions of all sizes across the Earth. It has a prevailing descriptive character. The main aim is to understand or define the uniqueness or character of a particular region which consists of natural as well as human elements. Attention is paid also to regionalization which covers the proper techniques of space delimitation into regions.

Regional geography is also considered as a certain approach to study in geographical sciences (similar to quantitative or critical geographies).

## Related fields

- Urban planning, regional planning and spatial planning: use the science of geography to assist in determining how to develop (or not develop) the land to meet particular criteria, such as safety, beauty, economic opportunities, the preservation of the built or natural heritage, and so on. The planning of towns, cities, and rural areas may be seen as applied geography.
- Regional science: In the 1950s the regional science movement led by Walter Isard arose, to provide a more quantitative and analytical base to geographical questions, in contrast to the descriptive tendencies of traditional geography programs. Regional science comprises the body of knowledge in which the spatial dimension plays a fundamental role, such as regional economics, resource management, location theory, urban and regional planning, transport and communication, human geography, population distribution, landscape ecology, and environmental quality.
- Interplanetary Sciences: While the discipline of geography is normally concerned with the Earth, the term can also be informally used to describe the study of other worlds, such as the planets of the Solar System and even beyond. The study of systems larger than the earth itself usually forms part of Astronomy or Cosmology. The study of other planets is usually called planetary science. Alternative terms such as Areology (the study of Mars) have been proposed, but are not widely used.

## Techniques

As spatial interrelationships are key to this synoptic science, maps are a key tool. Classical cartography has been joined by a more modern approach to geographical analysis, computer-based geographic information systems (GIS).

In their study, geographers use four interrelated approaches:

- Systematic - Groups geographical knowledge into categories that can be explored globally.
- Regional - Examines systematic relationships between categories for a specific region or location on the planet.

- Descriptive - Simply specifies the locations of features and populations.
- Analytical - Asks *why* we find features and populations in a specific geographic area.

## **Cartography**

Cartography studies the representation of the Earth's surface with abstract symbols (map making). Although other subdisciplines of geography rely on maps for presenting their analyses, the actual making of maps is abstract enough to be regarded separately. Cartography has grown from a collection of drafting techniques into an actual science.

Cartographers must learn cognitive psychology and ergonomics to understand which symbols convey information about the Earth most effectively, and behavioral psychology to induce the readers of their maps to act on the information. They must learn geodesy and fairly advanced mathematics to understand how the shape of the Earth affects the distortion of map symbols projected onto a flat surface for viewing. It can be said, without much controversy, that cartography is the seed from which the larger field of geography grew. Most geographers will cite a childhood fascination with maps as an early sign they would end up in the field.

## **Geographic information systems**

Geographic information systems (GIS) deal with the storage of information about the Earth for automatic retrieval by a computer, in an accurate manner appropriate to the information's purpose. In addition to all of the other subdisciplines of geography, GIS specialists must understand computer science and database systems. GIS has revolutionized the field of cartography; nearly all mapmaking is now done with the assistance of some form of GIS software. GIS also refers to the science of using GIS software and GIS techniques to represent, analyze and predict spatial relationships. In this context, GIS stands for Geographic Information Science.

## **Remote sensing**

Remote sensing is the science of obtaining information about Earth features from measurements made at a distance. Remotely sensed data comes in many forms such as satellite imagery, aerial photography and data obtained from hand-held sensors. Geographers increasingly use remotely sensed data to obtain information about the Earth's land surface, ocean and atmosphere because it: a) supplies objective information at a variety of spatial scales (local to global), b) provides a synoptic view of the area of interest, c) allows access to distant and/or inaccessible sites, d) provides spectral information outside the visible portion of the electromagnetic spectrum, and e) facilitates studies of how features/areas change over time. Remotely sensed data may be analyzed either independently of, or in conjunction with, other digital data layers (e.g., in a Geographic Information System).

## Quantitative methods

Geostatistics deal with quantitative data analysis, specifically the application of statistical methodology to the exploration of geographic phenomena. Geostatistics is used extensively in a variety of fields including: hydrology, geology, petroleum exploration, weather analysis, urban planning, logistics, and epidemiology. The mathematical basis for geostatistics derives from cluster analysis, linear discriminant analysis and non-parametric statistical tests, and a variety of other subjects. Applications of geostatistics rely heavily on geographic information systems, particularly for the interpolation (estimate) of unmeasured points. Geographers are making notable contributions to the method of quantitative techniques.

## Qualitative methods

Geographic qualitative methods, or ethnographical; research techniques, are used by human geographers. In cultural geography there is a tradition of employing qualitative research techniques also used in anthropology and sociology. Participant observation and in-depth interviews provide human geographers with qualitative data.

## History

The oldest known world maps date back to ancient Babylon from the 9th century BC. The best known Babylonian world map, however, is the *Imago Mundi* of 600 BC. The map as reconstructed by Eckhard Unger shows Babylon on the Euphrates, surrounded by a circular landmass showing Assyria, Urartu and several cities, in turn surrounded by a "bitter river" (Oceanus), with seven islands arranged around it so as to form a seven-pointed star. The accompanying text mentions seven outer regions beyond the encircling ocean. The descriptions of five of them have survived. In contrast to the *Imago Mundi*, an earlier Babylonian world map dating back to the 9th century BC depicted Babylon as being further north from the center of the world, though it is not certain what that center was supposed to represent.

The ideas of Anaximander (c. 610 B.C.-c. 545 B.C.), considered by later Greek writers to be the true founder of geography, come to us through fragments quoted by his successors. Anaximander is credited with the invention of the gnomon, the simple yet efficient Greek instrument that allowed the early measurement of latitude. Thales, Anaximander is also credited with the prediction of eclipses. The foundations of geography can be traced to the ancient cultures, such as the ancient, medieval, and early modern Chinese. The Greeks, who were the first to explore geography as both art and science, achieved this through Cartography, Philosophy, and Literature, or through Mathematics. There is some debate about who was the first person to assert that the Earth is spherical in shape, with the credit going either to Parmenides or Pythagoras. Anaxagoras was able to demonstrate that the profile of the Earth was circular by explaining eclipses. However, he still believed that the Earth was a flat disk, as did many of his contemporaries. One of the first estimates of the radius of the Earth was made by Eratosthenes.

The first rigorous system of latitude and longitude lines is credited to Hipparchus. He employed a sexagesimal system that was derived from Babylonian mathematics. The parallels and meridians were sub-divided into 360°, with each degree further subdivided 60' (minutes). To measure the longitude at different location on Earth, he suggested using eclipses to determine the relative difference in time. The extensive mapping by the Romans as they explored new lands would later provide a high level of information for Ptolemy to construct detailed atlases. He extended the work of Hipparchus, using a grid system on his maps and adopting a length of 56.5 miles for a degree.

From the 3rd century onwards, Chinese methods of geographical study and writing of geographical literature became much more complex than what was found in Europe at the time (until the 13th century). Chinese geographers such as Liu An, Pei Xiu, Jia Dan, Shen Kuo, Fan Chengda, Zhou Daguan, and Xu Xiake wrote important treatises, yet by the 17th century, advanced ideas and methods of Western-style geography were adopted in China.

During the Middle Ages, the fall of the Roman empire led to a shift in the evolution of geography from Europe to the Islamic world. Muslim geographers such as Muhammad al-Idrisi produced detailed world maps (such as Tabula Rogeriana), while other geographers such as Yaqut al-Hamawi, Abu Rayhan Biruni, Ibn Battuta and Ibn Khaldun provided detailed accounts of their journeys and the geography of the regions they visited. Turkish geographer, Mahmud al-Kashgari drew a world map on a linguistic basis, and later so did Piri Reis (Piri Reis map). Further, Islamic scholars translated and interpreted the earlier works of the Romans and Greeks and established the House of Wisdom in Baghdad for this purpose. Abū Zayd al-Balkhī, originally from Balkh, founded the "Balkhī school" of terrestrial mapping in Baghdad. Suhrāb, a late tenth century Muslim geographer, accompanied a book of geographical coordinates with instructions for making a rectangular world map, with equirectangular projection or cylindrical equidistant projection. In the early 11th century, Avicenna hypothesized on the geological causes of mountains in *The Book of Healing* (1027).

Abu Rayhan Biruni (976-1048) first described a polar equi-azimuthal equidistant projection of the celestial sphere. He was regarded as the most skilled when it came to mapping cities and measuring the distances between them, which he did for many cities in the Middle East and Indian subcontinent. He often combined astronomical readings and mathematical equations, in order to develop methods of pin-pointing locations by recording degrees of latitude and longitude. He also developed similar techniques when it came to measuring the heights of mountains, depths of valleys, and expanse of the horizon. He also discussed human geography and the planetary habitability of the Earth. He hypothesized that roughly a quarter of the Earth's surface is habitable by humans. He also calculated the latitude of Kath, Khwarezm, using the maximum altitude of the Sun, and solved a complex geodesic equation in order to accurately compute the Earth's circumference, which were close to modern values of the Earth's circumference. His estimate of 6,339.9 km for the Earth radius was only 16.8 km less than the modern value of 6,356.7 km. In contrast to his predecessors who measured the Earth's circumference by sighting the Sun simultaneously from two different locations, al-Biruni developed a new

method of using trigonometric calculations based on the angle between a plain and mountain top which yielded more accurate measurements of the Earth's circumference and made it possible for it to be measured by a single person from a single location. He also published a study of map projections, *Cartography*, which included a method for projecting a hemisphere on a plane.



Self portrait of Alexander von Humboldt, one of the early pioneers of geography

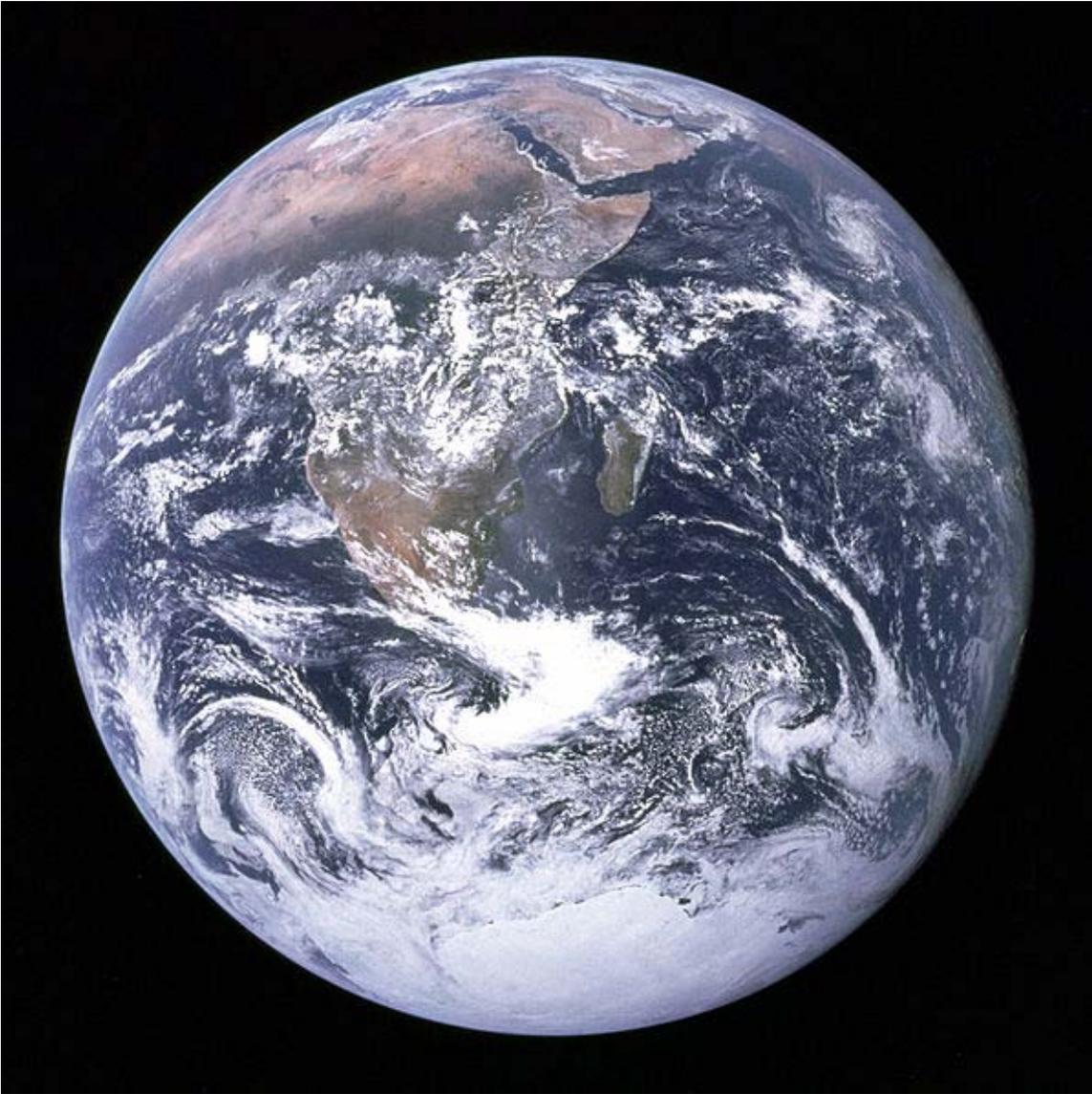
The European Age of Discovery during the 16th and 17th centuries, where many new lands were discovered and accounts by European explorers such as Christopher Columbus, Marco Polo and James Cook, revived a desire for both accurate geographic detail, and more solid theoretical foundations in Europe. The problem facing both explorers and geographers was finding the latitude and longitude of a geographic

location. The problem of latitude was solved long ago but that of longitude remained; agreeing on what zero meridian should be was only part of the problem. It was left to John Harrison to solve it by inventing the chronometer H-4, in 1760, and later in 1884 for the International Meridian Conference to adopt by convention the Greenwich meridian as zero meridian.

The 18th and 19th centuries were the times when geography became recognized as a discrete academic discipline and became part of a typical university curriculum in Europe (especially Paris and Berlin). The development of many geographic societies also occurred during the 19th century with the foundations of the Société de Géographie in 1821, the Royal Geographical Society in 1830, Russian Geographical Society in 1845, American Geographical Society in 1851, and the National Geographic Society in 1888. The influence of Immanuel Kant, Alexander von Humboldt, Carl Ritter and Paul Vidal de la Blache can be seen as a major turning point in geography from a philosophy to an academic subject.

Over the past two centuries the advancements in technology such as computers, have led to the development of geomatics and new practices such as participant observation and geostatistics being incorporated into geography's portfolio of tools. In the West during the 20th century, the discipline of geography went through four major phases: environmental determinism, regional geography, the quantitative revolution, and critical geography. The strong interdisciplinary links between geography and the sciences of geology and botany, as well as economics, sociology and demographics have also grown greatly especially as a result of Earth System Science that seeks to understand the world in a holistic view.

# Gaia hypothesis



The study of planetary habitability is partly based upon extrapolation from knowledge of the Earth's conditions, as the Earth is the only planet currently known to harbour life.

The **Gaia hypothesis**, **Gaia theory** or **Gaia principle** is an ecological hypothesis or theory proposing that the biosphere and the physical components of the Earth (atmosphere, cryosphere, hydrosphere and lithosphere) are closely integrated to form a complex interacting system that maintains the climatic and biogeochemical conditions on Earth in a preferred homeorhesis. Originally proposed by James Lovelock as the earth feedback hypothesis, it was named the Gaia Hypothesis after the Greek primordial goddess of the Earth, at the suggestion of William Golding, Nobel prizewinner in

literature and friend and neighbour of Lovelock. The hypothesis is frequently described as viewing the Earth as a single organism.

## History

The Gaia hypothesis was first scientifically formulated in the 1960s by the independent research scientist James Lovelock, as a consequence of his work for NASA on methods of detecting life on Mars. He initially published the *Gaia Hypothesis* in journal articles in the early 1970s followed by a popularizing 1979 book *Gaia: A new look at life on Earth*.

The hypothesis was initially, according to Lovelock, a way to explain the fact that combinations of chemicals including oxygen and methane persist in stable concentrations in the atmosphere of the Earth. Lovelock suggested detecting such combinations in other planets' atmospheres as a relatively reliable and cheap way to detect life, which many biologists opposed at the time and since. Later, other relationships such as sea creatures producing sulfur and iodine in approximately the same quantities as required by land creatures emerged and helped bolster the theory. Rather than invent many different theories to describe each such equilibrium, Lovelock dealt with them holistically, naming this self-regulating living system after the Greek goddess Gaia, using a suggestion from the novelist William Golding, who was living in the same village as Lovelock at the time (Bowerchalke, Wiltshire, UK). The Gaia Hypothesis has since been supported by a number of scientific experiments and provided a number of useful predictions, and hence is properly referred to as the Gaia theory.

Since 1971, the noted microbiologist Dr. Lynn Margulis has been Lovelock's most important collaborator in developing Gaian concepts.

Until 1975 the hypothesis was almost totally ignored. An article in the *New Scientist* of February 15, 1975, and a popular book length version of the theory, published in 1979 as *The Quest for Gaia*, began to attract scientific and critical attention to the hypothesis. The theory was then attacked by many mainstream biologists. Championed by certain environmentalists and climate scientists, it was vociferously rejected by many others, both within scientific circles and outside them.

## Lovelock's initial hypothesis

James Lovelock defined Gaia as:

*a complex entity involving the Earth's biosphere, atmosphere, oceans, and soil; the totality constituting a feedback or cybernetic system which seeks an optimal physical and chemical environment for life on this planet.*

His initial hypothesis was that the biomass modifies the conditions on the planet to make conditions on the planet more hospitable – the Gaia Hypothesis properly defined this "hospitality" as a full homeostasis. Lovelock's initial hypothesis, accused of being

teleological by his critics, was that the atmosphere is kept in homeostasis by and for the biosphere.

Lovelock suggested that life on Earth provides a cybernetic, homeostatic feedback system operated automatically and unconsciously by the biota, leading to broad stabilization of global temperature and chemical composition.

With his initial hypothesis, Lovelock claimed the existence of a global control system of surface temperature, atmosphere composition and ocean salinity. His arguments were:

- The global surface temperature of the Earth has remained constant, despite an increase in the energy provided by the Sun.
- Atmospheric composition remains constant, even though it should be unstable.
- Ocean salinity is constant.

Since life started on Earth, the energy provided by the Sun has increased by 25% to 30%; however, the surface temperature of the planet has remained remarkably constant when measured on a global scale. Furthermore, he argued, the atmospheric composition of the Earth is constant. The Earth's atmosphere currently consists of 79% nitrogen, 20.7% oxygen and 0.03% carbon dioxide. Oxygen is the second most reactive element after fluorine, and should combine with gases and minerals of the Earth's atmosphere and crust. Traces of methane (at an amount of 100,000 tonnes produced per annum) should not exist, as methane is combustible in an oxygen atmosphere. This composition should be unstable, and its stability can only have been maintained with removal or production by living organisms.

Ocean salinity has been constant at about 3.4% for a very long time. Salinity stability is important as most cells require a rather constant salinity and do not generally tolerate values above 5%. Ocean salinity constancy was a long-standing mystery, because river salts should have raised the ocean salinity much higher than observed. Recently it was suggested that salinity may also be strongly influenced by seawater circulation through hot basaltic rocks, and emerging as hot water vents on ocean spreading ridges. However, the composition of sea water is far from equilibrium, and it is difficult to explain this fact without the influence of organic processes.

The only significant natural source of atmospheric carbon dioxide (CO<sub>2</sub>) is volcanic activity, while the only significant removal is through the precipitation of carbonate rocks. In water, CO<sub>2</sub> is dissolved as a "carbonic acid", which may be combined with dissolved calcium to form solid calcium carbonate (limestone). Both precipitation and solution are influenced by the bacteria and plant roots in soils, where they improve gaseous circulation, or in coral reefs, where calcium carbonate is deposited as a solid on the sea floor. Calcium carbonate can also be washed from continents to the sea where it is used by living organisms to manufacture carbonaceous tests and shells. Once dead, the living organisms' shells fall to the bottom of the oceans where they generate deposits of chalk and limestone. Part of the organisms with carbonaceous shells are the coccolithophores (algae), which also have a role in the formation of clouds. When they

die, they release dimethyl sulfide gas (DMS),  $(\text{CH}_3)_2\text{S}$ , which is converted by atmospheric processes to sulfate particles on which water vapor condenses to make clouds.

Lovelock sees this as one of the complex processes that maintain conditions suitable for life. The volcanoes produce  $\text{CO}_2$  in the atmosphere,  $\text{CO}_2$  participates in rock weathering as carbonic acid, itself accelerated by temperature and soil life, the dissolved  $\text{CO}_2$  is then used by the algae and released on the ocean floor.  $\text{CO}_2$  excess can be compensated by an increase of coccolithophoride life, increasing the amount of  $\text{CO}_2$  locked in the ocean floor. Coccolithophorides increase the cloud cover, hence control the surface temperature, help cool the whole planet and favor precipitations necessary for terrestrial plants. For Lovelock and other Gaia scientists like Stephan Harding, coccolithophorides are one stage in a regulatory feedback loop. Lately the atmospheric  $\text{CO}_2$  concentration has increased and there is some evidence that concentrations of ocean algal blooms are also increasing.

## **Controversial concepts**

Lovelock used language that caused disagreement. For instance, many evolutionary biologists such as the late science historian Stephen Jay Gould and ethologist Richard Dawkins attacked his statement in the first paragraph of his book (1979), that "the quest for Gaia is an attempt to find the largest living creature on Earth."

Lynn Margulis, a microbiologist who collaborated with Lovelock in supporting the Gaia hypothesis, argued that "Darwin's grand vision was not wrong, only incomplete. In accentuating the direct competition between individuals for resources as the primary selection mechanism, Darwin (and especially his followers) created the impression that the environment was simply a static arena." In 1999, she wrote that the composition of the Earth's atmosphere, hydrosphere, and lithosphere are regulated around "set points" as in homeostasis, but those set points change with time.

She also wrote that there is no special tendency of biospheres to preserve their current inhabitants, and certainly not to make them comfortable. According to her, the Earth is a kind of community of trust that can exist at many discrete levels of integration. All multicellular organisms do not live or die all at once: not all cells in the body die instantaneously, nor are homeostatic "set points" constant through the life of an organism.

## **Critical analysis**

This theory is based on the idea that the biomass self-regulates the conditions on the planet to make its physical environment (in particular temperature and chemistry of the atmosphere) on the planet more hospitable to the species that constitute its "life." The Gaia Hypothesis properly defined this "hospitality" as a full homeostasis. A model that is often used to illustrate the original Gaia Hypothesis is the so-called Daisyworld simulation.

Whether this sort of system is present on Earth is still open to debate. Some relatively simple homeostatic mechanisms are generally accepted. For example, when atmospheric carbon dioxide levels rise, the biomass of photosynthetic organisms increases and thus removes more carbon dioxide from the atmosphere, but the extent to which these mechanisms stabilize and modify the Earth's overall climate are not yet known. Less clear is the reason why such traits should evolve in a system to produce such effects. Lovelock accepts a process of systemic Darwinian evolution for such mechanisms: creatures that improve their environment for their survival do better than those that damage their environment. However, many scientists do not believe such mechanisms exist.

## Criticism

After initially being largely ignored by most scientists, (from 1969 until 1977), thereafter for a period, the initial Gaia hypothesis was ridiculed by a number of scientists, such as Ford Doolittle, Dawkins and Gould. Lovelock has said that by naming his theory after a Greek goddess, championed by many non scientists, the Gaia hypothesis was derided as some kind of neo-Pagan New Age religion. Many scientists in particular also criticised the approach taken in his popular book "Gaia, a New look at Life on Earth" for being teleological; a belief that all things have a predetermined purpose. Responding to this assertion in 1990, Lovelock stated "Nowhere in our writings do we express the idea that planetary self-regulation is purposeful, or involves foresight or planning by the biota."

In 1981, W. Ford Doolittle, in the *CoEvolution Quarterly* article "Is Nature Motherly" argued that nothing in the genome of individual organisms could provide the feedback mechanisms Gaia theory proposed, and therefore the Gaia hypothesis was an unscientific theory of a maternal type without any explanatory mechanism. In Richard Dawkins' 1982 book, *The Extended Phenotype*, he argued that organisms could not act in concert as this would require foresight and planning from them. Like Doolittle he rejected the possibility that feedback loops could stabilize the system. Dawkins claimed "there was no way for evolution by natural selection to lead to altruism on a Global scale".

Stephen Jay Gould criticised Gaia as merely a metaphorical description of Earth processes. He wanted to know the actual mechanisms by which self-regulating homeostasis was regulated. Lovelock argues that no one mechanism is responsible, that the connections between the various known mechanisms may never be known, that this is accepted in other fields of biology and ecology as a matter of course, and that specific hostility is reserved for his own theory for political reasons.

Aside from clarifying his language and understanding of what is meant by a life form, Lovelock himself ascribes most of the criticism to a lack of understanding of non-linear mathematics by his critics, and a linearizing form of greedy reductionism in which all events have to be immediately ascribed to specific causes before the fact. He notes also that his theory suggests experiments in many different fields, but few of them in biology, which most of his critics are trained in. "I'm a general practitioner in a world where there's nothing but specialists... science in the last two centuries has tended to be ever-

dividing" and often rivalrous, especially for funding, which Lovelock describes as overly abundant and overly focused on institutions rather than original thought. He points out that Richard Feynman not only shared this opinion (coining the term cargo cult science) but also accepted a lack of general cause and effect explanation as an inevitable phase in a theory's development, and believed that some self-regulating phenomena may not be explainable at all mathematically.

## Theory

One of the criteria of the empirical definition of life is its ability to replicate and pass on their genetic information to succeeding generations. Consequently, an argument against the idea that Gaia is a "living" organism is the fact that the planet is unable to reproduce.

Lovelock, however, defines life as a self-preserving, self-similar system of feedback loops like Humberto Maturana's autopoiesis; as a self-similar system, life could be a cell as well as an organ embedded into a larger organism as well as an individual in a larger inter-dependent social context. The biggest context of interacting inter-dependent living entities is the Earth. The problematic empirical definition is getting "fuzzy on the edges": Why are highly specialized bacteria like E. coli that are unable to thrive outside their habitat considered "life", while mitochondria, which have evolved independently from the rest of the cell, are not?

Maturana and Lovelock changed this with the autopoiesis deductive definition, which to them explains the phenomenon of life better. Some aspects of the empirical definition, however, no longer apply. Reproduction becomes optional: bee swarms reproduce, while the biosphere has no need to. Lovelock himself states in the original Gaia book that even that is not true; given the possibilities, the biosphere may multiply in the future by colonizing other planets, as humankind may be the primer by which Gaia will reproduce. Humanity's exploration of space, its interest in colonizing and even terraforming other planets, lends some plausibility to the idea that Gaia might in effect be able to reproduce.

The astronomer Carl Sagan also remarked that from a cosmic viewpoint, the space probes since 1959 have the character of a planet preparing to go to seed. This might warrant interpretation as a rhetorical point, however, as it equivocates two differing meanings of "reproduction" otherwise.

## Daisyworld simulations

Lovelock responded to criticisms by developing the mathematical model Daisyworld with Andrew Watson to demonstrate that feedback mechanisms could evolve from the actions or activities of self-interested organisms, rather than through classic group selection mechanisms.

Daisyworld examines the energy budget of a planet populated by two different types of plants, black daisies and white daisies. The colour of the daisies influences the albedo of the planet such that black daisies absorb light and warm the planet, while white daisies

reflect light and cool the planet. Competition between the daisies (based on temperature-effects on growth rates) leads to a balance of populations that tends to favour a planetary temperature close to the optimum for daisy growth. Lovelock and Watson demonstrated the stability of Daisyworld by forcing the sun that it orbits to evolve along the main sequence, taking it from low to high solar constant. This perturbation of Daisyworld's receipt of solar radiation caused the balance of daisies to gradually shift from black to white but the planetary temperature was always regulated back to this optimum (except at the extreme ends of solar evolution). This situation is very different from the corresponding abiotic world, where temperature is unregulated and rises linearly with solar output. Later versions of Daisyworld introduced a range of grey daisies and populations of grazers and predators, and found that these further increased the stability of the homeostasis. More recently other research, modelling the real biochemical cycles of Earth, and using various "guilds" of life (e.g. photosynthesisers, decomposers, herbivores and primary and secondary carnivores) has also been shown to produce Daisyworld-like regulation and stability, which helps to explain planetary biological diversity.

This enables nutrient recycling within a regulatory framework derived by natural selection amongst species, where one being's harmful waste becomes low energy food for members of another guild. This research on the Redfield ratio of Nitrogen to Phosphorus shows that local biotic processes can regulate global systems.

## **First Gaia conference**

In 1985, the first public symposium on the Gaia Hypothesis—Is The Earth A Living Organism? -- was held at the University of Massachusetts August 1–6. The principal sponsor was the National Audubon Society Expedition Institute. Speakers included James Lovelock, George Wald, Mary Catherine Bateson, Lewis Thomas, John Todd, Donald Michael, Christopher Bird, Thomas Berry, Michael Cohen, and William Fields. Some 500 people attended and a concert by Paul Winter concluded the program. The symposium was produced by James A. Swan and Roberta Swan.

## **Second Gaia conference**

In 1988, to draw attention to the Gaia hypothesis, the climatologist Stephen Schneider organised a conference of the American Geophysical Union's first Chapman Conference on Gaia, held at San Diego in 1989, solely to discuss Gaia.

At the conference James Kirchner criticised the Gaia hypothesis for its imprecision. He claimed that Lovelock and Margulis had not presented one Gaia hypothesis, but four -

- CoEvolutionary Gaia — that life and the environment had evolved in a coupled way. Kirchner claimed that this was already accepted scientifically and was not new.

- Homeostatic Gaia — that life maintained the stability of the natural environment, and that this stability enabled life to continue to exist.
- Geophysical Gaia — that the Gaia theory generated interest in geophysical cycles and therefore led to interesting new research in terrestrial geophysical dynamics.
- Optimising Gaia — that Gaia shaped the planet in a way that made it an optimal environment for life as a whole. Kirchner claimed that this was not testable and therefore was not scientific.

Of Homeostatic Gaia, Kirchner recognised two alternatives. "Weak Gaia" asserted that life tends to make the environment stable for the flourishing of all life. "Strong Gaia" according to Kirchner, asserted that life tends to make the environment stable, *to enable* the flourishing of all life. Strong Gaia, Kirchner claimed, was untestable and therefore not scientific.

Referring to the Daisyworld Simulations, Kirchner responded that these results were predictable because of the intention of the programmers — Lovelock and Watson, who selected examples that produced the responses they desired.

Lawrence E. Joseph in his book *Gaia: The Growth of an Idea* argued that Kirchner's attack was principally against Lovelock's integrity as a scientist. Lovelock did not attack Kirchner's views for ten years, until his autobiography "Homage to Gaia", where he calls Kirchner's position *sophistry*. Lovelock and other Gaia-supporting scientists, however, did attempt to disprove the claim that the theory is not scientific because it is impossible to test it by controlled experiment. For example, against the charge that Gaia was teleological Lovelock and Andrew Watson offered the Daisyworld model (and its modifications, above) as evidence against most of these criticisms.

Lovelock was careful to present a version of the Gaia Hypothesis that had no claim that Gaia intentionally or consciously maintained the complex balance in her environment that life needed to survive. It would appear that the claim that Gaia acts "intentionally" was a metaphoric statement in his popular initial book and was not meant to be taken literally. This new statement of the Gaia hypothesis was more acceptable to the scientific community.

The accusations of teleologism were largely dropped after this conference.

## Range of views

Some have found James Kirchner's suggested spectrum, proposed at the First Gaia Chapman Conference, useful in suggesting that the original Gaia hypothesis could be split into a spectrum of hypotheses, ranging from the undeniable (Weak Gaia) to the radical (Strong Gaia).

## Weak Gaia

At one end of this spectrum is the undeniable statement that the organisms on the Earth have altered its composition. A stronger position is that the Earth's biosphere effectively acts as if it is a self-organizing system, which works in such a way as to keep its systems in some kind of "meta-equilibrium" that is broadly conducive to life. The history of evolution, ecology and climate show that the exact characteristics of this equilibrium intermittently have undergone rapid changes, which are believed to have caused extinctions and felled civilizations.

Weak Gaian hypotheses suggest that Gaia is co-evolutive. Co-evolution in this context has been thus defined: "Biota influence their abiotic environment, and that environment in turn influences the biota by Darwinian process." Lovelock (1995) gave evidence of this in his second book, showing the evolution from the world of the early thermo-acidophilic and methanogenic bacteria towards the oxygen enriched atmosphere today that supports more complex life.

The weakest form of the theory has been called "influential Gaia". It states that biota minimally influence certain aspects of the abiotic world, e.g. temperature and atmosphere.

The weak versions are more acceptable from an orthodox science perspective, as they assume non-homeostasis. They state the evolution of life and its environment may affect each other. An example is how the activity of photosynthetic bacteria during Precambrian times have completely modified the Earth atmosphere to turn it aerobic, and as such supporting evolution of life (in particular eukaryotic life). However, these theories do not claim the atmosphere modification has been done in coordination and through homeostasis. Also such critical theories have yet to explain how conditions on Earth have not been changed by the kinds of run-away positive feedbacks that have affected Mars and Venus.

Biologists and earth scientists usually view the factors that stabilize the characteristics of a period as an undirected emergent property or entelechy of the system; as each individual species pursues its own self-interest, for example, their combined actions tend to have counterbalancing effects on environmental change. Opponents of this view sometimes reference examples of lives' actions that have resulted in dramatic change rather than stable equilibrium, such as the conversion of the Earth's atmosphere from a reducing environment to an oxygen-rich one. However, proponents argue these atmospheric changes improved the environment's suitability for life.

Some go a step further and hypothesize that all lifeforms are part of one single living planetary being called *Gaia*. In this view, the atmosphere, the seas and the terrestrial crust would be results of interventions carried out by Gaia through the coevolving diversity of living organisms. While it is arguable that the Earth as a unit does not match the generally accepted biological criteria for life itself (*Gaia* has not yet reproduced, for instance; it still might *spread* to other planets through human space colonization and

terraforming), many scientists would be comfortable characterizing the earth as a single "system".

## **Strong Gaia**

A version called "Optimizing Gaia" asserts that biota manipulate their physical environment for the purpose of creating biologically favorable, or even optimal, conditions for themselves. "The Earth's atmosphere is more than merely anomalous; it appears to be a contrivance specifically constituted for a set of purposes". Further, "... it is unlikely that chance alone accounts for the fact that temperature, pH and the presence of compounds of nutrient elements have been, for immense periods, just those optimal for surface life. Rather, ... energy is expended by the biota to actively maintain these optima".

Another strong hypothesis is the one called "Omega Gaia". Teilhard de Chardin claimed that the Earth is evolving through stages of cosmogenesis, affecting the geosphere, biogenesis of the biosphere, and noogenesis of the noosphere, culminating in the *Omega Point*. Another form of the strong Gaia hypothesis is proposed by Guy Murchie who extends the quality of a holistic lifeform to galaxies. "After all, we are made of star dust. Life is inherent in nature." Murchie describes geologic phenomena such as sand dunes, glaciers, fires, etc. as living organisms, as well as the life of metals and crystals. "The question is not whether there is life outside our planet, but whether it is possible to have "nonlife".

There are speculative versions of the Gaia hypothesis, including versions that hold that the Earth is conscious or part of some universe-wide evolution such as expressed in the Selfish Biocosm hypothesis strain of a larger speculative Gaia philosophy. These extreme forms of the Gaia hypothesis, that the entire Earth is a single unified organism that is *consciously* manipulating the climate to make conditions more conducive to life, are metaphysical or mystical views for which no evidence exists, and that cannot be tested scientifically. The political branch of Gaia theory is the Gaia Movement, a collection of different organisations operating in different countries, but all sharing a concern for how humans might live more sustainably within the "living system".

## **Recent developments**

Gaia Theory has developed considerably and in recent years both Lovelock's and Margulis's understanding of Gaia have gained some increased support as a potentially viable, testable scientific hypothesis or theory.. Margulis dedicated the last of eight chapters in her book, *The Symbiotic Planet*, to Gaia. She resented the widespread personification of Gaia and stressed that Gaia is "not an organism", but "an emergent property of interaction among organisms". She defined Gaia as "the series of interacting ecosystems that compose a single huge ecosystem at the Earth's surface. Period." Yet still she argues, "the surface of the planet behaves as a physiological system in certain limited ways". Margulis seems to agree with Lovelock in that, in what comes to these physiological processes, the Earth's surface is "best regarded as alive". The book's most

memorable "slogan" was actually quipped by a student of Margulis': "Gaia is just symbiosis as seen from space". This neatly connects Gaia theory to Margulis' own theory of endosymbiosis.

### **Third Gaia conference**

By the time of the 2nd Chapman Conference on the Gaia Hypothesis, held at Valencia, Spain, on 23 June 2000, the situation had developed significantly in accordance with the developing science of Bio-geophysiology. Rather than a discussion of the Gaian teleological views, or "types" of Gaia Theory, the focus was upon the specific mechanisms by which basic short term homeostasis was maintained within a framework of significant evolutionary long term structural change.

The major questions were:

1. "How has the global biogeochemical/climate system called Gaia changed in time? What is its history? Can Gaia maintain stability of the system at one time scale but still undergo vectorial change at longer time scales? How can the geologic record be used to examine these questions?"
2. "What is the structure of Gaia? Are the feedbacks sufficiently strong to influence the evolution of climate? Are there parts of the system determined pragmatically by whatever disciplinary study is being undertaken at any given time or are there a set of parts that should be taken as most true for understanding Gaia as containing evolving organisms over time? What are the feedbacks among these different parts of the Gaian system, and what does the near closure of matter mean for the structure of Gaia as a global ecosystem and for the productivity of life?"
3. "How do models of Gaian processes and phenomena relate to reality and how do they help address and understand Gaia? How do results from Daisyworld transfer to the real world? What are the main candidates for "daisies"? Does it matter for Gaia theory whether we find daisies or not? How should we be searching for daisies, and should we intensify the search? How can Gaian mechanisms be investigated using process models or global models of the climate system that include the biota and allow for chemical cycling?"

In 1997, Tyler Volk argued that a Gaian system is almost inevitably produced as a result of an evolution towards far-from-equilibrium homeostatic states that maximise entropy production, and Kleidon (2004) agreed stating: "...homeostatic behavior can emerge from a state of MEP associated with the planetary albedo"; "...the resulting behavior of a biotic Earth at a state of MEP may well lead to near-homeostatic behavior of the Earth system on long time scales, as stated by the Gaia hypothesis." Staley (2002) has similarly proposed "...an alternative form of Gaia theory based on more traditional Darwinian principles... In [this] new approach, environmental regulation is a consequence of population dynamics, not Darwinian selection. The role of selection is to favor organisms that are best adapted to prevailing environmental conditions. However, the environment is not a static backdrop for evolution, but is heavily influenced by the presence of living

organisms. The resulting co-evolving dynamical process eventually leads to the convergence of equilibrium and optimal conditions."

### **Fourth Gaia conference**

A third international conference on the Gaia Theory, sponsored by the Northern Virginia Regional Park Authority and others, was held in October 2006 at the Arlington, VA campus of George Mason University. Lynn Margulis, Distinguished University Professor in the Department of Geosciences, University of Massachusetts-Amherst, and long-time advocate of the Gaia Theory, was a keynote speaker. Among many other speakers: Tyler Volk, Co-director of the Program in Earth and Environmental Science at New York University; Dr. Donald Aitken, Principal of Donald Aitken Associates; Dr. Thomas Lovejoy, President of the Heinz Center for Science, Economics and the Environment; Robert Correll, Senior Fellow, Atmospheric Policy Program, American Meteorological Society and noted environmental ethicist, J. Baird Callicott. James Lovelock, the theory's progenitor, prepared a video specifically for the event.

This conference approached Gaia Theory as both science and metaphor as a means of understanding how we might begin addressing 21st century issues such as climate change and ongoing environmental destruction.

### **Gaia hypothesis in ecology**

After much criticism, a modified Gaia hypothesis is now considered within ecological science basically consistent with the planet Earth being the ultimate object of ecological study. Ecologists generally consider the biosphere as an ecosystem and the Gaia hypothesis, though a simplification of that original proposed, to be consistent with a modern vision of global ecology, relaying the concepts of biosphere and biodiversity. The Gaia hypothesis has been called geophysiology or Earth System Science, which takes into account the interactions between biota, the oceans, the geosphere, and the atmosphere. To promote research and discussion in these fields an organisation, "Gaia Society for Research and Education in Earth System Science" was started.

An example of the change in acceptability of Gaia theories is the Amsterdam declaration of the scientific communities of four international global change research programmes — the International Geosphere-Biosphere Programme (IGBP), the International Human Dimensions Programme on Global Environmental Change (IHDP), the World Climate Research Programme (WCRP) and the international biodiversity programme DIVERSITAS — recognise that, in addition to the threat of significant climate change, there is growing concern over the ever-increasing human modification of other aspects of the global environment and the consequent implications for human well-being.

The programmes have stated the following:

"Research carried out over the past decade under the auspices of the four programmes to address these concerns has shown that:

1. The Earth System behaves as a single, self-regulating system with physical, chemical, biological, and human components. The interactions and feedbacks between the component parts are complex and exhibit multi-scale temporal and spatial variability. The understanding of the natural dynamics of the Earth System has advanced greatly in recent years and provides a sound basis for evaluating the effects and consequences of human-driven change.
2. Human activities are significantly influencing Earth's environment in many ways in addition to greenhouse gas emissions and climate change. Anthropogenic changes to Earth's land surface, oceans, coasts and atmosphere and to biological diversity, the water cycle and biogeochemical cycles are clearly identifiable beyond natural variability. They are equal to some of the great forces of nature in their extent and impact. Many are accelerating. Global change is real and is happening now.
3. Global change cannot be understood in terms of a simple cause-effect paradigm. Human-driven changes cause multiple effects that cascade through the Earth System in complex ways. These effects interact with each other and with local- and regional-scale changes in multidimensional patterns that are difficult to understand and even more difficult to predict.
4. Earth System dynamics are characterised by critical thresholds and abrupt changes. Human activities could inadvertently trigger such changes with severe consequences for Earth's environment and inhabitants. The Earth System has operated in different states over the last half million years, with abrupt transitions (a decade or less) sometimes occurring between them. Human activities have the potential to switch the Earth System to alternative modes of operation that may prove irreversible and less hospitable to humans and other life. The probability of a human-driven abrupt change in Earth's environment has yet to be quantified but is not negligible.
5. In terms of some key environmental parameters, the Earth System has moved well outside the range of the natural variability exhibited over the last half million years at least. The nature of changes now occurring simultaneously in the Earth System, their magnitudes and rates of change are unprecedented. The Earth is currently operating in a no-analogue state."

Sir Crispin Tickell in the 46th Annual Bennett Lecture for the 50th Anniversary of Geology at the University of Leicester in his recent talk "Earth Systems Science: Are We Pushing Gaia Too Hard?" stated "as a theory, Gaia is now winning."

He continued "The same goes for the earth systems science, which is now the concern of the Geological Society of London (with which the Gaia Society recently merged). Whatever the label, earth systems science, or Gaia, has now become a major subject of inquiry and research, and no longer has to justify itself."

These findings would seem to be fully in accord with the Gaia theory. Despite this endorsement, the late W. D. Hamilton, one of the founders of modern Darwinism, whilst conceding the empirical basis of the planetary homeostatic processes on which Gaia is based, states that it is a theory still awaiting its Copernicus. The homeostatic nature of the

global system has been recognized as a consequence of the 2<sup>nd</sup> law of thermodynamics. In their comprehensive book on the thermodynamics of life, Eric D. Schneider and Dorion Sagan argue that Gaia belongs to a class of complex thermodynamic systems, not just living ones, that are naturally purposeful; and that life optimizes rather than maximizes entropy production.

## ***The Revenge of Gaia***

In James Lovelock's 2006 book, *The Revenge of Gaia*, he argues that the lack of respect humans have had for Gaia, through the damage done to rainforests and the reduction in planetary biodiversity, is testing Gaia's capacity to minimize the effects of the addition of greenhouse gases in the atmosphere. This eliminates the planet's negative feedbacks and increases the likelihood of homeostatic positive feedback potential associated with runaway global warming. Similarly the warming of the oceans is extending the oceanic thermocline layer of tropical oceans into the Arctic and Antarctic waters, preventing the rise of oceanic nutrients into the surface waters and eliminating the algal blooms of phytoplankton on which oceanic foodchains depend. As phytoplankton and forests are the main ways in which Gaia draws down greenhouse gases, particularly carbon dioxide, taking it out of the atmosphere, the elimination of this environmental buffering will see, according to Lovelock, most of the earth becoming uninhabitable for humans and other life-forms by the middle of this century, with a massive extension of tropical deserts.

Given these conditions, Lovelock expects human civilization will be hard pressed to survive. He expects the change to be similar to the Paleocene-Eocene Thermal Maximum when atmospheric concentration of CO<sub>2</sub> was 450 ppm. At that point the Arctic Ocean was 23 °C and had crocodiles in it, with the rest of the world mostly scrub and desert. He says of sustainable development and renewable energy that it came "200 years too late" and that more effort should go into adaptation, including more use of nuclear fission as a viable energy source in the future (unclear reference - clarification needed). He likens the Kyoto Protocol to the Munich conferences that failed to prevent World War II, including the likelihood that the disaster will cause people to come together in common cause. "We have been through no less than seven of these events as humans...comparable in extent to the change" likely to be wrought by global warming.

He claims that Gaia's self-regulation will likely prevent any extraordinary runaway effects that wipe out life itself, but that humans will survive and be "culled and, I hope, refined."

According to James Lovelock, by 2040, the world population of more than six billion will have been culled by floods, drought and famine. Indeed "[t]he people of Southern Europe, as well as South-East Asia, will be fighting their way into countries such as Canada, Australia and Britain".

"By 2040, parts of the Sahara desert will have moved into middle Europe. We are talking about Paris - as far north as Berlin. In Britain we will escape because of our oceanic position."

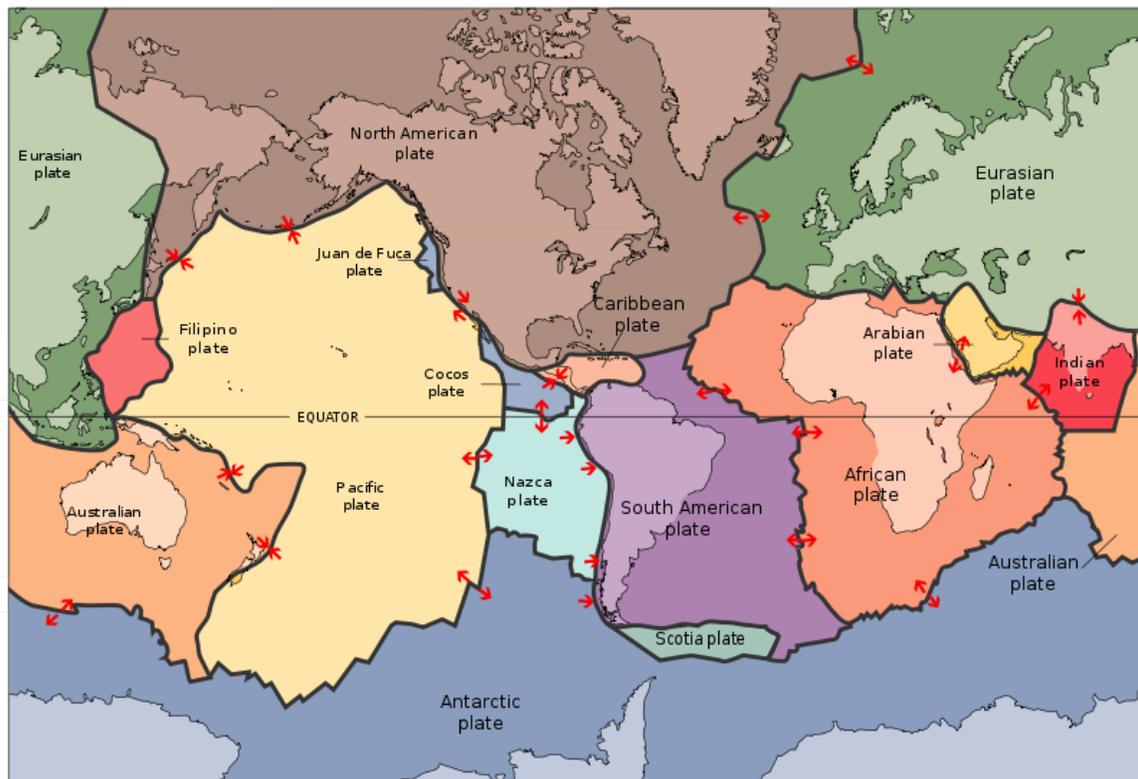
"If you take the Intergovernmental Panel on Climate Change predictions, then by 2040 every summer in Europe will be as hot as it was in 2003 - between 110F and 120F. It is not the death of people that is the main problem, it is the fact that the plants can't grow — there will be almost no food grown in Europe."

"We are about to take an evolutionary step and my hope is that the species will emerge stronger. It would be hubris to think humans as they now are God's chosen race."

## Chapter 6

# Earth's Interior & Related Theories

## 1. Plate tectonics



The tectonic plates of the world were mapped in the second half of the 20th century.

**Plate tectonics** (from the Late Latin *tectonicus*, from the Greek: *τεκτονικός* "pertaining to building") (Little, Fowler & Coulson 1990) is a scientific theory which describes the large scale motions of Earth's lithosphere. The theory builds on the older concepts of continental drift, developed during the first decades of the 20th century (one of the most famous advocates was Alfred Wegener), and was accepted by the majority of the Geoscientific community when the concepts of seafloor spreading were developed in the

late 1950s and early 1960s. The lithosphere is broken up into what are called "tectonic plates". In the case of the Earth, there are currently seven to eight major (depending on how they are defined) and many minor plates. The lithospheric plates ride on the asthenosphere. These plates move in relation to one another at one of three types of plate boundaries: convergent, or collisional boundaries; divergent boundaries, also called spreading centers; and conservative transform boundaries. Earthquakes, volcanic activity, mountain-building, and oceanic trench formation occur along these plate boundaries. The lateral relative movement of the plates varies, though it is typically 0–100 mm annually (Read & Watson 1975).

The tectonic plates are composed of two types of lithosphere: thicker continental and thin oceanic. The upper part is called the crust, again of two types (continental and oceanic). This means that a plate can be of one type, or of both types. One of the main points the theory proposes is that the amount of surface of the (continental and oceanic) plates that disappear in the mantle along the convergent boundaries by subduction is more or less in equilibrium with the new (oceanic) crust that is formed along the divergent margins by seafloor spreading. This is also referred to as the "conveyor belt" principle. In this way, the total surface of the Globe remains the same. This is in contrast with earlier theories advocated before the Plate Tectonics "paradigm", as it is sometimes called, became the main scientific model, theories that proposed gradual shrinking (contraction) or gradual expansion of the Globe, and that still exist in science as alternative models.

Regarding the driving mechanism of the plates various models co-exist: Tectonic plates are able to move because the Earth's lithosphere has a higher strength and lower density than the underlying asthenosphere. Lateral density variations in the mantle result in convection. Their movement is thought to be driven by a combination of the motion of seafloor away from the spreading ridge (due to variations in topography and density of the crust that result in differences in gravitational forces) and drag, downward suction, at the subduction zones. A different explanation lies in different forces generated by the rotation of the Globe and tidal forces of the Sun and the Moon. The relative importance of each of these factors is unclear.

## **Key principles**

The outer layers of the Earth are divided into lithosphere and asthenosphere. This is based on differences in mechanical properties and in the method for the transfer of heat. Mechanically, the lithosphere is cooler and more rigid, while the asthenosphere is hotter and flows more easily. In terms of heat transfer, the lithosphere loses heat by conduction whereas the asthenosphere also transfers heat by convection and has a nearly adiabatic temperature gradient. This division should not be confused with the *chemical* subdivision of these same layers into the mantle (comprising both the asthenosphere and the mantle portion of the lithosphere) and the crust: a given piece of mantle may be part of the lithosphere or the asthenosphere at different times, depending on its temperature and pressure.

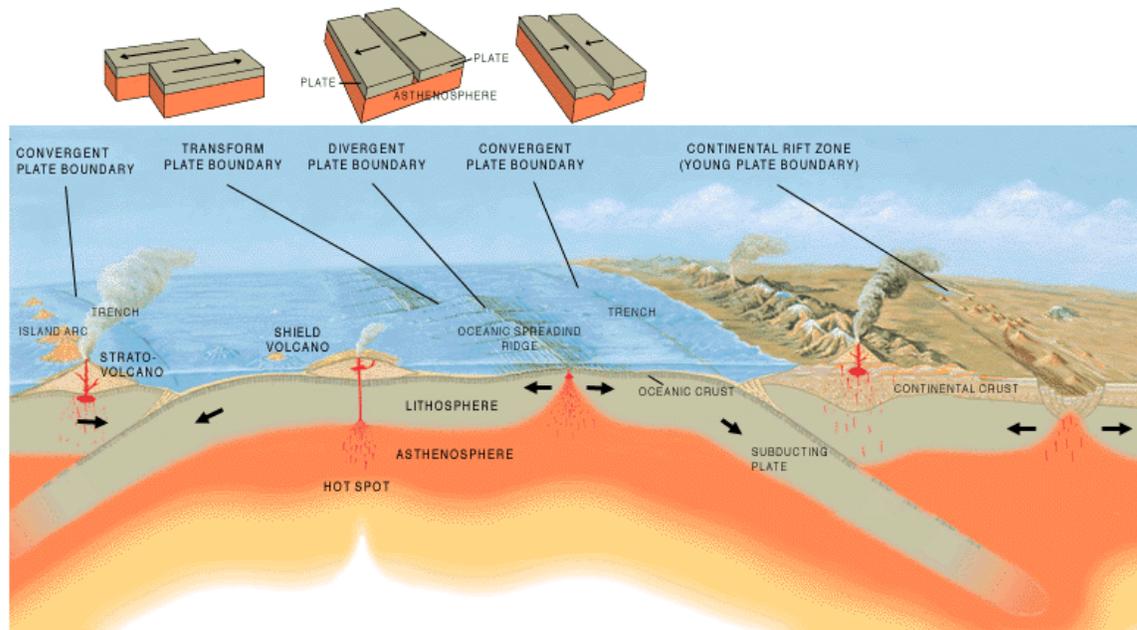
The key principle of plate tectonics is that the lithosphere exists as separate and distinct *tectonic plates*, which ride on the fluid-like (visco-elastic solid) asthenosphere. Plate motions range up to a typical 10–40 mm/a (Mid-Atlantic Ridge; about as fast as fingernails grow), to about 160 mm/a (Nazca Plate; about as fast as hair grows) (Zhen Shao 1997; Hancock, Skinner & Dineley 2000). The driving mechanism behind this movement is described below in a separate section.

Tectonic lithosphere plates consist of lithospheric mantle overlain by either or both of two types of crustal material: oceanic crust (in older texts called *sima* from silicon and magnesium) and continental crust (*sial* from silicon and aluminium). Average oceanic lithosphere is typically 100 km thick (Turcotte & Schubert 2002); its thickness is a function of its age: as time passes, it conductively cools and becomes thicker. Because it is formed at mid-ocean ridges and spreads outwards, its thickness is therefore a function of its distance from the mid-ocean ridge where it was formed. For a typical distance oceanic lithosphere must travel before being subducted, the thickness varies from about 6 km thick at mid-ocean ridges to greater than 100 km at subduction zones; for shorter or longer distances, the subduction zone (and therefore also the mean) thickness becomes smaller or larger, respectively (Turcotte & Schubert 2002). Continental lithosphere is typically ~200 km thick, though this also varies considerably between basins, mountain ranges, and stable cratonic interiors of continents. The two types of crust also differ in thickness, with continental crust being considerably thicker than oceanic (35 km vs. 6 km) (Turcotte & Schubert 2002).

The location where two plates meet is called a *plate boundary*, and plate boundaries are commonly associated with geological events such as earthquakes and the creation of topographic features such as mountains, volcanoes, mid-ocean ridges, and oceanic trenches. The majority of the world's active volcanoes occur along plate boundaries, with the Pacific Plate's Ring of Fire being most active and most widely known. These boundaries are discussed in further detail below.

As explained above, tectonic plates can include continental crust or oceanic crust, and many plates contain both. For example, the African Plate includes the continent and parts of the floor of the Atlantic and Indian Oceans. The distinction between oceanic crust and continental crust is based on their modes of formation. Oceanic crust is formed at sea-floor spreading centers, and continental crust is formed through arc volcanism and accretion of terranes through tectonic processes; though some of these terranes may contain ophiolite sequences, which are pieces of oceanic crust, these are considered part of the continent when they exit the standard cycle of formation and spreading centers and subduction beneath continents. Oceanic crust is also denser than continental crust owing to their different compositions. Oceanic crust is denser because it has less silicon and more heavier elements ("mafic") than continental crust ("felsic") (Schmidt & Harbert 1998). As a result of this density stratification, oceanic crust generally lies below sea level (for example most of the Pacific Plate), while the continental crust buoyantly projects above sea level.

# Types of plate boundaries



Three types of plate boundary.

Basically, three types of plate boundaries exist (Meissner 2002, p. 100), with a fourth, mixed type, characterized by the way the plates move relative to each other. They are associated with different types of surface phenomena. The different types of plate boundaries are:

1. *Transform boundaries (Conservative)* occur where plates slide or, perhaps more accurately, grind past each other along transform faults. The relative motion of the two plates is either sinistral (left side toward the observer) or dextral (right side toward the observer). The San Andreas Fault in California is an example of a transform boundary exhibiting dextral motion.
2. *Divergent boundaries (Constructive)* occur where two plates slide apart from each other. Mid-ocean ridges (e.g., Mid-Atlantic Ridge) and active zones of rifting (such as Africa's Great Rift Valley) are both examples of divergent boundaries.
3. *Convergent boundaries (Destructive)* (or *active margins*) occur where two plates slide towards each other commonly forming either a subduction zone (if one plate moves underneath the other) or a continental collision (if the two plates contain continental crust). Deep marine trenches are typically associated with subduction zones. The subducting slab contains many hydrous minerals, which release their water on heating; this water then causes the mantle to melt, producing volcanism. Examples of this are the Andes mountain range in South America and the Japanese island arc.
4. *Plate boundary zones* occur where the effects of the interactions are unclear and the broad belt boundaries are not well defined.

## **Driving forces of plate motion**

Plate tectonics is basically a kinematic phenomenon: Earth scientists agree upon the observation and deduction that the plates have moved one with respect to the other, and debate and find agreements on how and when. But still a major question remains on what is the motor behind this movement; the geodynamic mechanism, and here science diverges in different theories.

Generally, it is accepted that tectonic plates are able to move because of the relative density of oceanic lithosphere and the relative weakness of the asthenosphere. Dissipation of heat from the mantle is acknowledged to be the original source of energy driving plate tectonics, through convection or large scale upwelling and doming. As a consequence, in the current view, although it is still a matter of some debate, because of the excess density of the oceanic lithosphere sinking in subduction zones a powerful source of plate motion is generated. When the new crust forms at mid-ocean ridges, this oceanic lithosphere is initially less dense than the underlying asthenosphere, but it becomes denser with age, as it conductively cools and thickens. The greater density of old lithosphere relative to the underlying asthenosphere allows it to sink into the deep mantle at subduction zones, providing most of the driving force for plate motions. The weakness of the asthenosphere allows the tectonic plates to move easily towards a subduction zone. Although subduction is believed to be the strongest force driving plate motions, it cannot be the only force since there are plates such as the North American Plate which are moving, yet are nowhere being subducted. The same is true for the enormous Eurasian Plate. The sources of plate motion are a matter of intensive research and discussion among earth scientists. One of the main points is that the kinematic pattern of the movements itself should be separated clearly from the possible geodynamic mechanism that is invoked as the driving force of the observed movements, as some patterns may be explained by more than one mechanism (van Dijk 1992, van Dijk & Okkes 1991). Basically, the driving forces that are advocated at the moment, can be divided in three categories: mantle dynamics related, gravity related (mostly secondary forces), and Earth rotation related.

### **Mantle dynamics related driving forces**

For a considerable period of around 25 years (last quarter of the twentieth century) the leading theory envisaged large scale convection currents in the upper mantle which are transmitted through the asthenosphere as the main driving force of the tectonic plates. This theory was launched by Arthur Holmes and some forerunners in the 1930s and was immediately recognised as the solution for the acceptance of the theory discussed since its occurrence in the papers of Alfred Wegener in the early years of the century. It was, though, long debated because the leading ("fixist") theory was still envisaging a static Earth without moving continents, up until the major break-throughs in the early sixties.

Two- and three-dimensional imaging of the Earth's interior (seismic tomography) shows that there is a laterally varying density distribution throughout the mantle. Such density variations can be material (from rock chemistry), mineral (from variations in mineral structures), or thermal (through thermal expansion and contraction from heat energy).

The manifestation of this varying lateral density is mantle convection from buoyancy forces (Tanimoto & Lay 2000).

How mantle convection relates directly and indirectly to the motion of the plates is a matter of ongoing study and discussion in geodynamics. Somehow, this energy must be transferred to the lithosphere in order for tectonic plates to move. There are essentially two types of forces that are thought to influence plate motion: friction and gravity.

Basal drag (friction): The plate motion is in this way driven by friction between the convection currents in the asthenosphere and the more rigid overlying floating lithosphere.

Slab suction (gravity): Local convection currents exert a downward frictional pull on plates in subduction zones at ocean trenches. Slab suction may occur in a geodynamic setting wherein basal tractions continue to act on the plate as it dives into the mantle (although perhaps to a greater extent acting on both the under and upper side of the slab).

Lately, the convection theory is much debated as modern techniques based on 3D seismic tomography of imaging the internal structure of the Earth's mantle still fail to recognise these predicted large scale convection cells. Therefore, alternative patterns of mantle dynamics have been proposed:

In the theory of plume tectonics developed during the 1990s, a modified concept of mantle convection currents is used, related to super plumes rising from the deeper mantle which would be the drivers or the substitutes of the major convection cells. These ideas, which find their roots in the early 1930s with the so-called "fixistic" ideas of the European and Russian Earth Science Schools, find resonance in the modern theories which envisage hot spots/mantle plumes in the mantle which remain fixed and are overridden by oceanic and continental lithosphere plates during time, and leave their traces in the geological record (though these phenomena are not invoked as real driving mechanisms, but rather as a modulator). The modern theories that continue building on the older mantle doming concepts and see the movements of the plates a secondary phenomena, are beyond the scope of this page and are discussed elsewhere for example on the plume tectonics page.

Another suggestion is that the mantle flows neither in cells nor large plumes, but rather as a series of channels just below the Earth's crust which then provide basal friction to the lithosphere. This theory is called "surge tectonics" and became quite popular in geophysics and geodynamics during the 1980s and 1990s (Smoot et al. 1996).

## **Gravity related driving forces**

Gravity related forces are usually invoked as secondary phenomena within the framework of a more general driving mechanism such as the various forms of mantle dynamics described above.

Gravitational sliding away from a spreading ridge: According to many authors, plate motion is driven by the higher elevation of plates at ocean ridges. As oceanic lithosphere

is formed at spreading ridges from hot mantle material, it gradually cools and thickens with age (and thus distance from the ridge). Cool oceanic lithosphere is significantly denser than the hot mantle material from which it is derived and so with increasing thickness it gradually subsides into the mantle to compensate the greater load. The result is a slight lateral incline with distance from the ridge axis.

This force is regarded as a secondary force often referred to as "ridge-push". This is a misnomer as nothing is "pushing" horizontally and tensional features are dominant along ridges. It is more accurate to refer to this mechanism as gravitational sliding as variable topography across the totality of the plate can vary considerably and the topography of spreading ridges is only the most prominent feature. Other mechanisms generating this gravitational secondary force are for example:

Flexural bulging of the lithosphere before it dives underneath an adjacent plate, for instance, produces a clear topographical feature that can offset or at least affect the influence of topographical ocean ridges.

Mantle plumes and hot spots impinging on the underside of tectonic plates can drastically alter the topography of the ocean floor. Some of these, on a larger scale, are seen as the major driving force of the plates (see below).

Slab-pull: Current scientific opinion is that the asthenosphere is insufficiently competent or rigid to directly cause motion by friction along the base of the lithosphere. Slab pull is therefore most widely thought to be the greatest force acting on the plates. In this current understanding, plate motion is mostly driven by the weight of cold, dense plates sinking into the mantle at trenches (Conrad & Lithgow-Bertelloni 2002). Recent models indicate that trench suction plays an important role as well. However, as the North American Plate is nowhere being subducted, yet it is in motion presents a problem. The same holds for the African, Eurasian, and Antarctic plates. Slab pull is especially invoked in areas where remnants of older lithosphere become trapped along convergence zones e.g. as relicts in collisional belts, which, sinking into the mantle and rolling backwards, exert a pull on the overlying crust.

Gravitational sliding away from mantle doming: According to older theories one of the driving mechanisms of the plates is the existence of large scale asthenosphere/mantle domes, which cause the gravitational sliding of lithosphere plates away from them. This gravitational sliding represents a secondary phenomenon of this, basically vertically oriented mechanism. This can act on various scales, from the small scale of one island arc up to the larger scale of an entire ocean basin.

### **Earth rotation related driving forces**

Alfred Wegener, being a meteorologist, had proposed tidal forces and pole flight Force as main driving mechanisms for continental drift. However, these forces were considered far too small to cause continental motion as the concept then was of continents plowing through oceanic crust. Therefore, also Wegener in his last edition of his book in 1929 converted to convection currents as the main driving force.

In the plate tectonics context (accepted since the seafloor spreading proposals of Heezen, Hess, Dietz, Morley, Vine and Matthews -see below- during the early 1960s), though, oceanic crust in motion *with* the continents which made the proposals related to Earth rotation to be reconsidered, also in more recent literature, these are:

1. Tidal drag due to the gravitational force the Moon (and the Sun) exerts on the crust of the Earth
2. Shear strain of the Earth globe due to N-S compression related to the rotation and modulations of it
3. Pole flight force: equatorial drift due to rotation and centrifugal effects: tendency of the plates to move from the poles to the equator ("*Polflucht*")
4. Coriolis effect the plates suffer when they move around the globe (coriolis effect/law of Buys Ballot)
5. Global deformation of the geoid due to small displacements of rotational pole with respect to the Earth crust
6. Other smaller deformation effects of the crust due to wobbles and spin movements of the Earth rotation on a smaller time scale.

In order for these mechanisms to be overall valid, systematic relationships should exist all over the Globe between the orientation and kinematics of deformation, and the geographical latitudinal and longitudinal grid of the Earth itself. Ironically, these systematic relations studies in the second half of the nineteenth century and the first half of the twentieth century do underline exactly the opposite: that the plates had not moved in time, that the deformation grid was fixed with respect to the Earth equator and axis, and that gravitational driving forces were generally acting vertically and caused only locally horizontal movements (the so-called pre-plate tectonic, "fixist theories"). Later studies (discussed below on this page) therefore invoked many of the relationships recognised during this pre-plate tectonics period, to support their theories.

Of the many forces discussed in this paragraph, tidal force is still highly debated and defended as a possible principle driving force, whereas the other forces are used or in global geodynamic models not using the plate tectonics concepts (therefore beyond the discussions treated in this section), or proposed as minor modulations within the overall plate tectonics model.

In 1973 George W. Moore of the USGS and R. C. Bostrom presented evidence for a general westward drift of the Earth's lithosphere with respect to the mantle, and, therefore, tidal forces or tidal lag or "friction" due to the Earth's rotation and the forces acting upon it by the Moon being a driving force for plate tectonics: as the Earth spins eastward beneath the moon, the moon's gravity ever so slightly pulls the Earth's surface layer back westward, just like proposed by Alfred Wegener (see above). In a more recent 2006 study (Scoppola et al. 2006), scientists rediscussed and advocated these earlier proposed ideas. It has also been suggested recently in Lovett (2006) that this observation may also explain why Venus and Mars have no plate tectonics, since Venus has no moon and Mars' moons are too small to have significant tidal effects on Mars. In a recent paper by Torsvik et al. (2010) it was suggested that, on the other hand, it can easily be observed

that many plates are moving north and eastward, and that the dominantly westward motion of the Pacific ocean basins derives simply from the eastward bias of the Pacific spreading center (which is not a predicted manifestation of such lunar forces). In the same paper the authors admit, however, that relative to the lower mantle, there is a slight westward component in the motions of all the plates. They demonstrated though that the westward drift, seen only for the past 30 Ma, is attributed to the increased dominance of the steadily growing and accelerating Pacific plate. The debate is still open.

### **Relative significance of each driving force mechanism**

The actual vector of a plate's motion must necessarily be a function of all the forces acting upon the plate. However, therein remains the problem regarding what degree each process contributes to the motion of each tectonic plate.

The diversity of geodynamic settings and properties of each plate must clearly result in differences in the degree to which such processes are actively driving the plates. One method of dealing with this problem is to consider the relative rate at which each plate is moving and to consider the available evidence of each driving force upon the plate as far as possible.

One of the most significant correlations found is that lithospheric plates attached to downgoing (subducting) plates move much faster than plates not attached to subducting plates. The Pacific plate, for instance, is essentially surrounded by zones of subduction (the so-called Ring of Fire) and moves much faster than the plates of the Atlantic basin, which are attached (perhaps one could say 'welded') to adjacent continents instead of subducting plates. It is thus thought that forces associated with the downgoing plate (slab pull and slab suction) are the driving forces which determine the motion of plates, except for those plates which are not being subducted (Conrad & Lithgow-Bertelloni 2002). The driving forces of plate motion continue to be active subjects of on-going research within geophysics and tectonophysics.

### **Historical context - development of the theory**

Plate tectonics is the main current theory in Earth Sciences regarding the development of our planet Earth. It is, therefore, appropriate to dedicate some space to explain how the Earth Science community, step by step, has built this theory, from early speculations, through the gathering of proof and severe debates, up to the refinement and quantification, and still ongoing confrontations with alternative ideas.

## Summary



Detailed map showing the tectonic plates with their movement vectors.

In line with other previous and contemporaneous proposals, in 1912 the meteorologist Alfred Wegener amply described what he called continental drift, expanded in his 1915 book *The Origin of Continents and Oceans* and the scientific debate started that would end up fifty years later in the theory of plate tectonics (Hughes 2001a). Starting from the idea (also expressed by his forerunners) that the present continents once formed a single land mass (which was called Pangea later on) that drifted apart, thus releasing the continents from the Earth's mantle and likening them to "icebergs" of low density granite floating on a sea of denser basalt (Wegener 1966; Hughes 2001b).

But without detailed evidence and a force sufficient to drive the movement, the theory was not generally accepted: the Earth might have a solid crust and mantle and a liquid core, but there seemed to be no way that portions of the crust could move around.

Notwithstanding much opposition, the view of continental drift gained support and a lively debate started between "drifters" or "mobilists" (proponents of the theory) and "fixists" (opponents). During the 1920s, 1930s and 1940s, the former reached important milestones proposing that convection currents might have driven the plate movements, and that spreading may have occurred below the sea within the oceanic crust. Concepts close to the elements now incorporated in plate tectonics were proposed by geophysicists and geologists (both fixists and mobilists) like Vening-Meinesz, Holmes, and Umbgrove.

One of the first pieces of geophysical evidence that was used to support the movement of lithospheric plates came from paleomagnetism. This is based on the fact that rocks of different ages show a variable magnetic field direction, evidenced by studies since the mid-nineteenth century. The magnetic north and south reverse through time, and, especially important in paleotectonic studies, the relative position of the magnetic north

varies through time. Initially, during the first half of the twentieth century, the latter phenomena was explained by introducing what was called "polar wander", i.e., it was assumed that the north pole location had been shifting through time. An alternative explanation, though, was that the continents had moved (shifted and rotated) relative to the north pole, and each continent, in fact, shows its own "polar wander path". During the late 1950s it was shown with success that these data could show the validity of continental drift in two occasions: by Keith Runcorn in a paper in 1956, and by Warren Carey in a symposium held in March 1956.

The second piece of evidence in support of continental drift came during the late 1950s and early 60s from data on the bathymetry of the deep ocean floors and the nature of the oceanic crust such as magnetic properties and, more generally, with the development of marine geology which gave evidence for the association of seafloor spreading along the mid-oceanic ridges and magnetic field reversals, published between 1959 and 1963 by Heezen, Dietz, Hess, Mason, Vine & Matthews, and Morley (Korgen 1995; Spiess & Kuperman 2003).

Simultaneous advances in early seismic imaging techniques in and around Wadati-Benioff zones along the trenches bounding many continental margins, together with many other geophysical (e.g. gravimetric) and geological observations, showed how the oceanic crust could disappear into the mantle, providing the mechanism to balance the extension of the ocean basins with shortening along its margins.

All these evidences, both from the ocean floor and from the continental margins made clear around 1965 that continental drift was feasible and the theory of plate tectonics, which was defined in a series of papers between 1965 and 1967, was born, with all its extraordinary explanatory and predictive power. The theory revolutionized the Earth sciences, explaining a diverse range of geological phenomena and their implications in other studies such as paleogeography and paleobiology.

## **Continental drift**

In the late 19th and early 20th centuries, geologists assumed that the Earth's major features were fixed, and that most geologic features such as basin development and mountain ranges could be explained by vertical crustal movement, described in what is called the geosynclinal theory. Generally, this was placed in the context of a contracting planet Earth due to heat loss in the course of a relatively short geological time.

It was observed as early as 1596 that the opposite coasts of the Atlantic Ocean—or, more precisely, the edges of the continental shelves—have similar shapes and seem to have once fitted together (Kious & Tilling 1996).

Since that time many theories were proposed to explain this apparent complementarity, but the assumption of a solid Earth made these various proposals difficult to accept (Frankel 1987).

The discovery of radioactivity and its associated heating properties in 1895 prompted a re-examination of the apparent age of the Earth (Joly 1909) since this had previously been estimated by its cooling rate and assumption the Earth's surface radiated like a black body (Thomson 1863).

Those calculations had implied that, even if it started at red heat, the Earth would have dropped to its present temperature in a few tens of millions of years. Armed with the knowledge of a new heat source, scientists realized that the Earth would be much older, and that its core was still sufficiently hot to be liquid.

By 1915, after having published a first article in 1912 (Wegener 1912). Alfred Wegener was making serious arguments for the idea of continental drift in the first edition of *The Origin of Continents and Oceans*. In that book (re-issued in four successive editions up to the final one in 1936), he noted how the east coast of South America and the west coast of Africa looked as if they were once attached. Wegener wasn't the first to note this (Abraham Ortelius, Snider-Pellegrini, Roberto Mantovani and Frank Bursley Taylor preceded him just to mention a few), but he was the first to marshal significant fossil and paleo-topographical and climatological evidence to support this simple observation (and was supported in this by researchers such as Alex du Toit). Furthermore, when the rock strata of the margins of separate continents are very similar it suggests that these rocks were formed in the same way, implying that they were joined initially. For instance, some parts of Scotland and Ireland contain rocks very similar to those found in Newfoundland and New Brunswick. Furthermore, the Caledonian Mountains of Europe and parts of the Appalachian Mountains of North America are very similar in structure and lithology.

However, his ideas were not taken seriously by many geologists, who pointed out that there was no apparent mechanism for continental drift. Specifically, they did not see how continental rock could plow through the much denser rock that makes up oceanic crust. Wegener could not explain the force that drove continental drift, and his vindication did not come until after his death in 1930.

### **Floating continents - paleomagnetism - seismicity zones**

As it was observed early that although granite existed on continents, seafloor seemed to be composed of denser basalt, the prevailing concept during the first half of the twentieth century was that there were two types of crust, named "sial" (continental type crust), and "sima" (oceanic type crust). Furthermore, it was supposed that a static shells of strata was present under the continents. It therefore looked apparent that a layer of basalt (sial) underlies the continental rocks.

However, based upon abnormalities in plumb line deflection by the Andes in Peru, Pierre Bouguer had deduced that less-dense mountains must have a downward projection into the denser layer underneath. The concept that mountains had "roots" was confirmed by George B. Airy a hundred years later during study of Himalayan gravitation, and seismic studies detected corresponding density variations. Therefore, by the mid-1950s the

question remained unresolved of whether mountain roots were clenched in surrounding basalt or were floating upon it like an iceberg.

During the 20th century, improvements in and greater use of seismic instruments such as seismographs enabled scientists to learn that earthquakes tend to be concentrated in specific areas, most notably along the oceanic trenches and spreading ridges. By the late 1920s, seismologists were beginning to identify several prominent earthquake zones parallel to the trenches that typically were inclined 40–60° from the horizontal and extended several hundred kilometers into the Earth. These zones later became known as Wadati-Benioff zones, or simply Benioff zones, in honor of the seismologists who first recognized them, Kiyoo Wadati of Japan and Hugo Benioff of the United States. The study of global seismicity greatly advanced in the 1960s with the establishment of the Worldwide Standardized Seismograph Network (WWSSN) to monitor the compliance of the 1963 treaty banning above-ground testing of nuclear weapons. The much improved data from the WWSSN instruments allowed seismologists to map precisely the zones of earthquake concentration world wide.

Meanwhile, debates developed around the phenomena of polar Wander. Since the early debates of continental drift, scientists had discussed and used evidence that polar drift had occurred due to the fact that continents seemed to have moved through different climatic zones during the past. Furthermore, paleomagnetic data had shown that the magnetic pole had also shifted during time. Reasoning in an opposite way, the continents might have shifted and rotated, while the pole remained relatively fixed. The first time the evidence of magnetic polar wander was used to support the movements of continents was in a paper by Keith Runcorn in 1956, and successive papers by him and his students Ted Irving (who was actually the first to be convinced of the fact that paleomagnetism supported continental drift) and Ken Creer.

This was immediately followed by a symposium in Tasmania in March 1956. In this symposium, the evidence was used in the theory of an expansion of the global crust. In this hypothesis the shifting of the continents can be simply explained by a large increase in size of the Earth since its formation. However, this was unsatisfactory because its supporters could offer no convincing mechanism to produce a significant expansion of the Earth. Certainly there is no evidence that the moon has expanded in the past 3 billion years; other work would soon show that the evidence was equally in support of continental drift on a globe with a stable radius.

During the thirties up to the late fifties, numerous milestones were reached that would eventually lead to the development of plate tectonics. These are the works of Vening-Meinesz, Holmes, Umbgrove, and numerous others, in which concepts close or near identical to modern plate tectonics theory were defined and outlined. The most important milestone was reached when the English geologist Arthur Holmes proposed in 1920 that plate junctions might lie beneath the sea, and in 1928 that convection currents within the mantle might be the driving force.

Often, all these milestones are forgotten for various reasons:

1. During this timespan, continental drift was not accepted.
2. Some of these ideas were discussed in the context of abandoned fixistic ideas of a deforming globe without continental drift or an expanding Earth.
3. They were published during an episode of extreme political and economic instability and scientific communication was obviously hampered by this.
4. Many of these were published by European scientists and at first not mentioned or given little credit in the papers published by the American researchers which during the 1960s presented evidence for sea floor spreading.

## **Mid oceanic ridge spreading and convection**

In 1947, a team of scientists led by Maurice Ewing utilizing the Woods Hole Oceanographic Institution's research vessel *Atlantis* and an array of instruments, confirmed the existence of a rise in the central Atlantic Ocean, and found that the floor of the seabed beneath the layer of sediments consisted of basalt, not the granite which is the main constituent of continents. They also found that the oceanic crust was much thinner than continental crust. All these new findings raised important and intriguing questions (Lippsett 2001; Lippsett 2006).

The new data that had been collected on the ocean basins also showed particular characteristics regarding the bathymetry. One of the major outcomes of these datasets was that all along the globe, a system of mid-oceanic ridges was detected. An important conclusion was that along this system, new ocean floor was being created, which led to the concept of the "Great Global Rift". This was described in the crucial paper of Bruce Heezen (1960) which would trigger a real revolution in thinking. A profound consequence of seafloor spreading is that new crust was, and is now, being continually created along the oceanic ridges. Therefore, Heezen advocated the so-called "expanding Earth" hypothesis of S. Warren Carey (see above). So, still the question remained: how can new crust be continuously added along the oceanic ridges without increasing the size of the Earth? In reality, this question had been solved already by numerous scientists during the forties and the fifties, like Arthur Holmes, Vening-Meinesz, Coates and many others: The crust in excess disappeared along what were called the oceanic trenches where so-called "subduction" occurred. Therefore, when various scientists during the early sixties started to reason on the data at their disposal regarding ocean floor, the pieces of the theory fell quickly at its place.

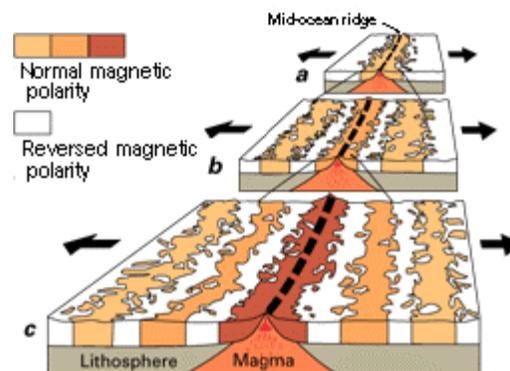
The question particularly intrigued Harry Hammond Hess, a Princeton University geologist and a Naval Reserve Rear Admiral, and Robert S. Dietz, a scientist with the U.S. Coast and Geodetic Survey who first coined the term *seafloor spreading*. Dietz and Hess (the former published the same idea one year earlier in *Nature*, but priority belongs to Hess who had already distributed an unpublished manuscript of his 1962 article by 1960) were among the small handful who really understood the broad implications of sea floor spreading and how it would eventually agree with the, at that time, unconventional

and unaccepted ideas of continental drift and the elegant and mobilistic models proposed by previous workers like Holmes.

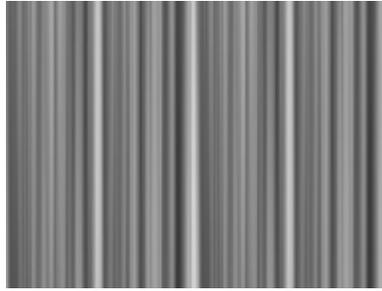
In the same year, Robert R. Coats of the U.S. Geological Survey described the main features of island arc subduction in the Aleutian Islands. His paper, though little-noted (and even ridiculed) at the time, has since been called "seminal" and "prescient". In reality, it actually shows that the work by the European scientists on island arcs and mountain belts performed and published during the 1930s up until the 1950s was applied and appreciated also in the United States.

If the Earth's crust was expanding along the oceanic ridges, Hess and Dietz reasoned like Holmes and others before them, it must be shrinking elsewhere. Hess followed Heezen suggesting that new oceanic crust continuously spreads away from the ridges in a conveyor belt-like motion. And, using the mobilistic concepts developed before, he correctly concluded that many millions of years later, the oceanic crust eventually descends along the continental margins where oceanic trenches – very deep, narrow canyons are present e.g. along the rim of the Pacific Ocean basin – were formed. The important step Hess made was that convection currents would be the driving force in this process, arriving at the same conclusions as Holmes had decades before with the only difference that the thinning of the ocean crust was performed using the mechanism of Heezen of spreading along the ridges. Hess therefore concluded that the Atlantic Ocean was expanding while the Pacific Ocean was shrinking. As old oceanic crust is "consumed" in the trenches, (like Holmes and others, he believed this was done by thickening of the continental lithosphere, not, as nowadays believed, by underthrusting at a larger scale of the oceanic crust itself into the mantle) new magma rises and erupts along the spreading ridges to form new crust. In effect, the ocean basins are perpetually being "recycled," with the creation of new crust and the destruction of old oceanic lithosphere occurring simultaneously, in a way like what later would be called the Wilson cycle (see below). Thus, the new mobilistic concepts neatly explained why the Earth does not get bigger with sea floor spreading, why there is so little sediment accumulation on the ocean floor, and why oceanic rocks are much younger than continental rocks.

### The final proof: magnetic striping



Seafloor magnetic striping



A demonstration of magnetic striping. (The darker the color is the closer it is to normal polarity)

Beginning in the 1950s, scientists like Victor Vacquier, using magnetic instruments (magnetometers) adapted from airborne devices developed during World War II to detect submarines, began recognizing odd magnetic variations across the ocean floor. This finding, though unexpected, was not entirely surprising because it was known that basalt—the iron-rich, volcanic rock making up the ocean floor—contains a strongly magnetic mineral (magnetite) and can locally distort compass readings. This distortion was recognized by Icelandic mariners as early as the late 18th century. More important, because the presence of magnetite gives the basalt measurable magnetic properties, these newly discovered magnetic variations provided another means to study the deep ocean floor. When newly formed rock cools, such magnetic materials recorded the Earth's magnetic field at the time.

As more and more of the seafloor was mapped during the 1950s, the magnetic variations turned out not to be random or isolated occurrences, but instead revealed recognizable patterns. When these magnetic patterns were mapped over a wide region, the ocean floor showed a zebra-like pattern: one stripe with normal polarity and the adjoining stripe with reversed polarity. The overall pattern, defined by these alternating bands of normally and reversely polarized rock, became known as magnetic striping, and was published by Ron G. Mason and co-workers in 1961, who didn't find, though, an explanation for these data in terms of sea floor spreading, like Vine, Matthews and Morley a few years later (Mason & Raff 1961); (Raff & Mason 1961).

The discovery of magnetic striping called for an explanation. In the early 1960s scientists such as Heezen, Hess and Dietz had begun to theorise that mid-ocean ridges mark structurally weak zones where the ocean floor was being ripped in two lengthwise along the ridge crest. New magma from deep within the Earth rises easily through these weak zones and eventually erupts along the crest of the ridges to create new oceanic crust. This process, at first denominated the "conveyer belt hypothesis" and later called seafloor spreading, operating over many millions of years continues to form new ocean floor all across the 50,000 km-long system of mid-ocean ridges.

Only four years after the maps with the "zebra pattern" of magnetic stripes were published, the link between sea floor spreading and these patterns was correctly placed, independently by Lawrence Morley, and by Fred Vine and Drummond Matthews, in 1963 (Vine & Matthews 1963) now called the Vine-Matthews-Morley hypothesis. This

hypothesis linked these patterns to geomagnetic reversals and was supported by several lines of evidence :

1. the stripes are symmetrical around the crests of the mid-ocean ridges; at or near the crest of the ridge, the rocks are very young, and they become progressively older away from the ridge crest;
2. the youngest rocks at the ridge crest always have present-day (normal) polarity;
3. stripes of rock parallel to the ridge crest alternate in magnetic polarity (normal-reversed-normal, etc.), suggesting that they were formed during different epochs documenting the (already known from independent studies) normal and reversal episodes of the Earth's magnetic field.

By explaining both the zebra-like magnetic striping and the construction of the mid-ocean ridge system, the seafloor spreading hypothesis (SFS) quickly gained converts and represented another major advance in the development of the plate-tectonics theory. Furthermore, the oceanic crust now came to be appreciated as a natural "tape recording" of the history of the geomagnetic field reversals (GMFR) of the Earth's magnetic field. Nowadays, extensive studies are dedicated to the calibration of the normal-reversal patterns in the oceanic crust on one hand and known timescales derived from the dating of basalt layers in sedimentary sequences (magnetostratigraphy) on the other, to arrive at estimates of past spreading rates and plate reconstructions.

### **Definition and refining of the theory - from new global tectonics to plate tectonics**

After all these considerations, Plate Tectonics (or, as it was initially called "New Global Tectonics") became quickly accepted in the scientific world, and numerous papers followed that defined the concepts:

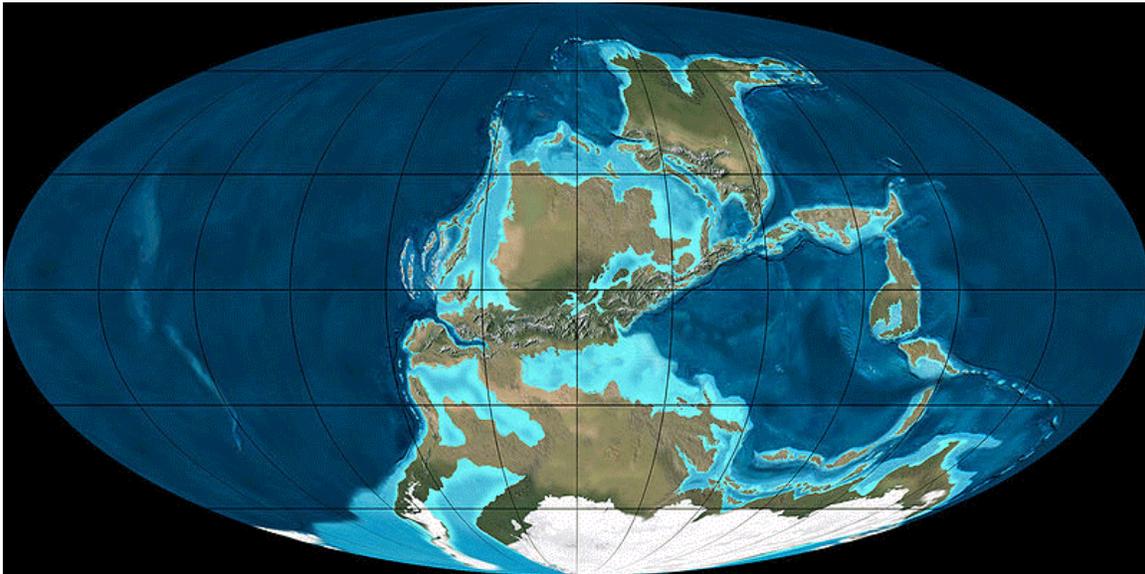
- In 1965, Tuzo Wilson who had been a promotor of the sea floor spreading hypothesis and continental drift from the very beginning (e.g. Wilson 1963) added the concept of transform faults to the model, completing the classes of fault types necessary to make the mobility of the plates on the globe work out (Wilson 1965).
- A symposium on continental drift was held at the Royal Society of London in 1965 which must be regarded as the official start of the acceptance of plate tectonics by the scientific community, and which abstracts are issued as Blacket, Bullard & Runcorn (1965). In this symposium, Edward Bullard and co-workers showed with a computer calculation how the continents along both sides of the Atlantic would best fit to close the ocean, which became known as the famous "Bullard's Fit".
- In 1966 Tuzo Wilson published the paper that referred to previous plate tectonic reconstructions, introducing the concept of what is now known as the "Wilson Cycle" (Wilson 1966).
- In 1967, at the American Geophysical Union's meeting, W. Jason Morgan proposed that the Earth's surface consists of 12 rigid plates that move relative to each other (Morgan 1968).

- Two months later, Xavier Le Pichon published a complete model based on 6 major plates with their relative motions, and we may say that this marks the final acceptance of the scientific community of plate tectonics (Le Pichon 1967).
- In the same year, McKenzie and Parker independently presented a model similar to Morgan's using translations and rotations on a sphere to define the plate motions (McKenzie & Parker 1967).

## Implications for biogeography

Continental drift theory helps biogeographers to explain the disjunct biogeographic distribution of present day life found on different continents but having similar ancestors (Moss & Wilson 1998). In particular, it explains the Gondwanan distribution of ratites and the Antarctic flora.

## Plate reconstruction



Reconstruction of plate configurations for the whole Phanerozoic

Reconstruction is used to establish past (and future) plate configurations, helping determine the shape and make-up of ancient supercontinents and providing a basis for paleogeography.

### Defining plate boundaries

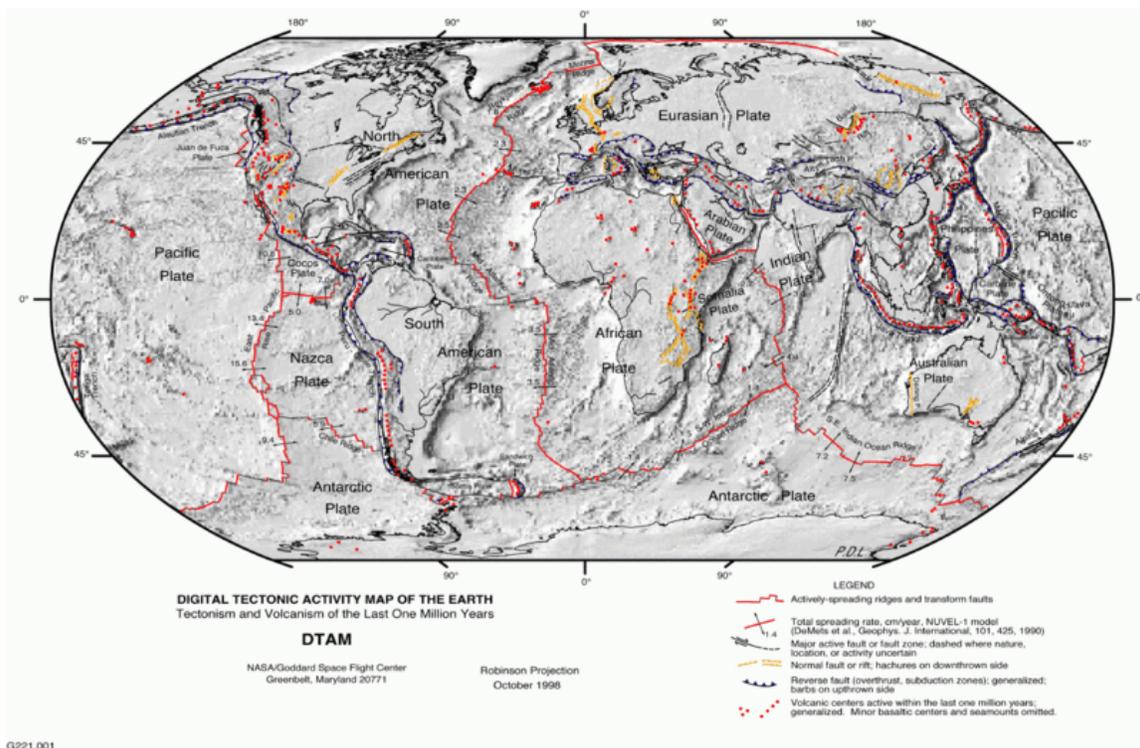
Current plate boundaries are defined by their seismicity (Condie 1997). Past plate boundaries within existing plates are identified from evidence of vanished oceans, such as ophiolites (Lliboutry 2000).

## Past plate motions

The movement of plates has caused the formation and break-up of continents over time, including occasional formation of a supercontinent that contains most or all of the continents. The supercontinent Rodinia is thought to have formed about 1 billion years ago and to have embodied most or all of Earth's continents, and broken up into eight continents around 600 million years ago. The eight continents later re-assembled into another supercontinent called Pangaea; Pangaea broke up into Laurasia (which became North America and Eurasia) and Gondwana (which became the remaining continents).

Various types of quantitative and semi-quantitative information are available to constrain past plate motions. The geometric fit between continents, such as between west Africa and South America is still an important part of plate reconstruction. Magnetic stripe patterns provide a reliable guide to relative plate motions going back into the Jurassic period. The tracks of hotspots give absolute reconstructions but these are only available back to the Cretaceous (Torsvik 2008). Older reconstructions rely mainly on paleomagnetic pole data, although these only constrain the latitude and rotation, but not the longitude. Combining poles of different ages in a particular plate to produce apparent polar wander paths provides a method for comparing the motions of different plates through time (Butler 1992). Additional evidence comes from the distribution of certain sedimentary rock types, faunal provinces shown by particular fossil groups, and the position of orogenic belts (Torsvik 2008).

## Current plates



## Major plates

Depending on how they are defined, there are usually seven or eight "major" plates:

- African Plate
- Antarctic Plate
- Indo-Australian Plate, sometimes subdivided into:
  - Indian Plate
  - Australian Plate
- Eurasian Plate
- North American Plate
- South American Plate
- Pacific Plate

## Minor plates

There are dozens of smaller plates, the seven largest of which are:

- Arabian Plate
- Caribbean Plate
- Juan de Fuca Plate
- Cocos Plate
- Nazca Plate
- Philippine Sea Plate
- Scotia Plate

## Current Motion

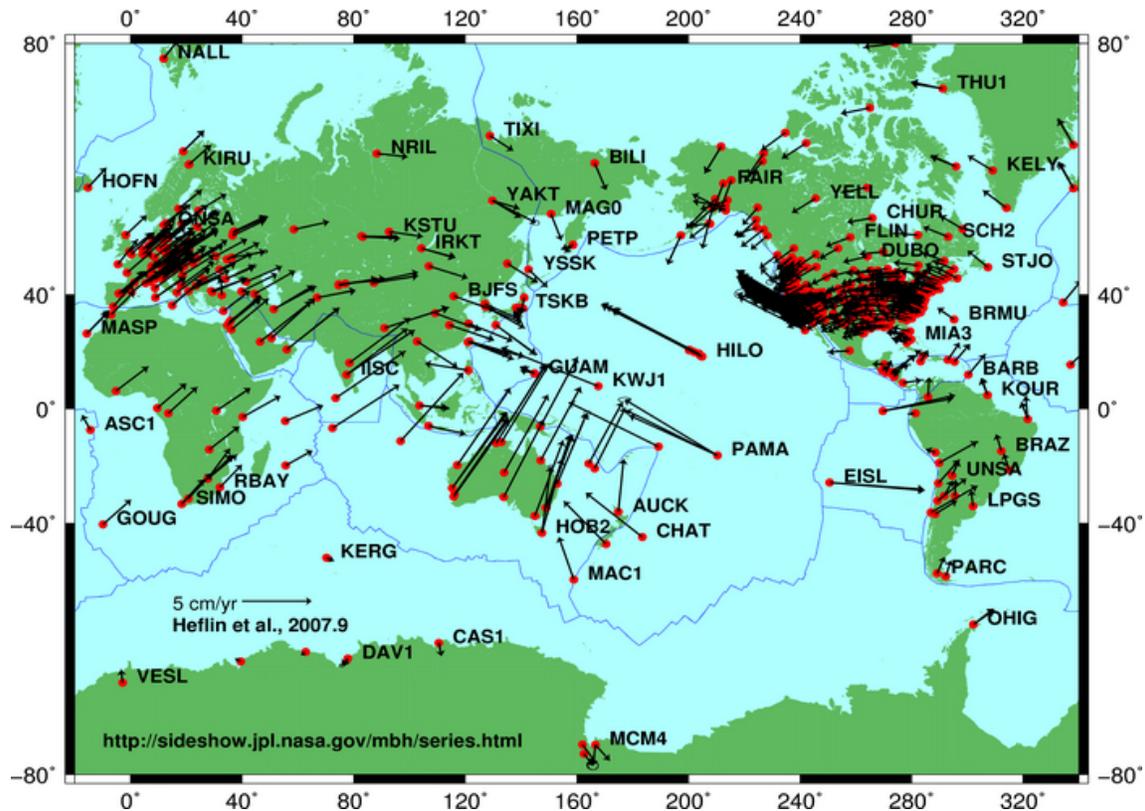


Plate motion based on Global Positioning System (GPS) satellite data from NASA JPL. The vectors show direction and magnitude of motion.

The current motion of the tectonic plates is nowadays revealed from remote sensing satellite data sets, calibrated with ground station measurements.

## Plate tectonics on other celestial bodies (Planets, Moons)

The appearance of plate tectonics on terrestrial planets is related to planetary mass, with more massive planets than Earth expected to exhibit plate tectonics. Earth may be a borderline case, owing its tectonic activity to abundant water (Valencia, O'Connell & Sasselov 2007)

(Silica and water form a deep eutectic.)

### Venus

Venus shows no evidence of active plate tectonics. There is debatable evidence of active tectonics in the planet's distant past; however, events taking place since then (such as the plausible and generally accepted hypothesis that the Venusian lithosphere has thickened greatly over the course of several hundred million years) has made constraining the course of its geologic record difficult. However, the numerous well-preserved impact

craters have been utilized as a dating method to approximately date the Venusian surface (since there are thus far no known samples of Venusian rock to be dated by more reliable methods). Dates derived are dominantly in the range c. 500 to 750 Ma, although ages of up to c. 1.2 Ga have been calculated. This research has led to the fairly well accepted hypothesis that Venus has undergone an essentially complete volcanic resurfacing at least once in its distant past, with the last event taking place approximately within the range of estimated surface ages. While the mechanism of such an impressive thermal event remains a debated issue in Venusian geosciences, some scientists are advocates of processes involving plate motion to some extent.

One explanation for Venus' lack of plate tectonics is that on Venus temperatures are too high for significant water to be present (Kasting 1988). The Earth's crust is soaked with water, and water plays an important role in the development of shear zones. Plate tectonics requires weak surfaces in the crust along which crustal slices can move, and it may well be that such weakening never took place on Venus because of the absence of water. However, some researchers remain convinced that plate tectonics is or was once active on this planet.

## **Mars**

Mars is considerably smaller than Earth and Venus, and there is evidence for ice on its surface and in its crust.

In the 1990s, it was proposed that Martian Crustal Dichotomy was created by plate tectonic processes (Sleep 1994). Scientists today disagree, and believe that it was created either by upwelling within the Martian mantle that thickened the crust of the Southern Highlands and formed Tharsis (Zhong & Zuber 2001) or by a giant impact that excavated the Northern Lowlands (Andrews-Hanna, Zuber & Banerdt 2008).

Observations made of the magnetic field of Mars by the *Mars Global Surveyor* spacecraft in 1999 showed patterns of magnetic striping discovered on this planet. Some scientists interpreted these as requiring plate tectonic processes, such as seafloor spreading (Connerney et al. 1999, Connerney et al. 2005)). However, their data fail a "magnetic reversal test", which is used to see if they were formed by flipping polarities of a global magnetic field (Harrison 2000).

## **Galilean satellites of Jupiter**

Some of the satellites of Jupiter have features that may be related to plate-tectonic style deformation, although the materials and specific mechanisms may be different from plate-tectonic activity on Earth.

## **Titan, moon of Saturn**

Titan, the largest moon of Saturn, was reported to show tectonic activity in images taken by the Huygens Probe, which landed on Titan on January 14, 2005 (Soderblom et al. 2007).

## **Exoplanets**

It is believed that many planets around other stars will have plate tectonics. On Earth-sized planets, plate tectonics is more likely if there are oceans of water, but on larger super-earths plate tectonics is very likely even if the planet is dry (Valencia, O'Connell & Sasselov 2007).

# Mountain



The Matterhorn, Swiss Alps



Five Finger Mountain, Azerbaijan.

A **mountain** is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a **hill**. The adjective **montane** is used to describe mountainous areas and things associated with them. The study of mountains is called Orography.

Exogeology deals with planetary mountains, which in that branch of science are usually called *montes* (singular—*mons*). The highest mountain on earth is the Mount Everest 8,848 m (29,029 ft). The highest known mountain in the Solar System is Olympus Mons on the planet Mars at 21,171 m (69,459 ft).

## Definition



The view of Jeff Davis Peak from the glacier-carved summit of Wheeler Peak, Nevada. Because Boundary Peak, Nevada is partially in California, and is actually a sub-peak of Montgomery Peak, the shorter Wheeler Peak can be considered the tallest mountain in Nevada.

There is no universally accepted definition of a mountain. Elevation, volume, relief, steepness, spacing and continuity has been used as criteria for defining a mountain. In the Oxford English Dictionary a mountain is defined as "a natural elevation of the earth surface rising more or less abruptly from the surrounding level and attaining an altitude which, relatively to the adjacent elevation, is impressive or notable."

Whether a landform is called a mountain may depend on usage among the local people. The highest point in San Francisco, California, is called Mount Davidson, notwithstanding its height of 300 m (980 ft), which makes it ten feet short of the minimum for a mountain in American appellation. Similarly, Mount Scott outside Lawton, Oklahoma is only 251 m (823 ft) from its base to its highest point.

Definitions of "mountain" include:

- Height over base of at least 2,500 m (8,202 ft).
- Height over base of 1,500 m (4,921 ft)–2,500 m (8,202 ft). with a slope greater than 2 degrees
- Height over base of 1,000 m (3,281 ft)–1,500 m (4,921 ft). with a slope greater than 5 degrees
- Local (radius 7,000 m (22,966 ft). elevation greater than 300 m (984 ft)., or 300 m (984 ft)–1,000 m (3,281 ft). if local (radius 7,000 m (22,966 ft). elevation is greater than 300 m (984 ft).

By this definition, mountains cover 64% of Asia, 25% of Europe, 22% of South America, 17% of Australia, and 3% of Africa. As a whole, 24% of the Earth's land mass is mountainous and 10% of people live in mountainous regions. Most of the world's rivers are fed from mountain sources, and more than half of humanity depends on mountains for water.

## Characteristics



Mountain in Carbon County, Utah

High mountains, as well as those located close to the Earth's poles, reach into the colder layers of the atmosphere. They are consequently subject to glaciation, and erosion

through frost action. Such processes produce the peak shape. Some of these mountains have glacial lakes, created by melting glaciers; for example, there are an estimated 3,000 glacial lakes in Bhutan. Mountains can be eroded and weathered, altering their characteristics over time.



Mount Olympus in Greece.

Tall mountains have different climatic conditions at the top than at the base, and will thus have different life zones at different altitudes. At the highest elevations, trees cannot grow, and whatever life may be present will be of the alpine type, resembling tundra. Just below the tree line, one may find subalpine forests of needleleaf trees, which can withstand cold, dry conditions. In regions with dry climates, the tendency of mountains to have higher precipitation as well as lower temperatures also provides for varying conditions, which in turn lead to differing flora and fauna. Some plants and animals found in these zones tend to become isolated since the conditions above and below a particular zone will be inhospitable and thus constrain their movements or dispersal. On the other hand, birds, being capable of flight, may take advantage of montane habitats and migrate into a region that would otherwise not provide appropriate habitat. These isolated ecological systems, or microclimates, are known as sky islands.

The reason mountains are colder than lowlands has to do with how the sun heats the surface of the earth. Practically all the heat at the surface of the Earth comes from the

sun, in the form of solar energy. The sun's radiation is absorbed by land and sea, whence the heat is transferred into the air. Air is an insulator, so conduction of heat from the ground to the atmosphere is negligible. Heat is mainly transferred into the atmosphere through convection and radiation. Warm air rises because of its buoyancy, leading to convective circulation, in the form of thermals, within the lowest layer of the atmosphere, the troposphere. When heat radiates from the surface of the earth, it is released as long-wave radiation, which does not travel through the air efficiently. This radiant heat is absorbed temporarily by gasses in the atmosphere, such as carbon dioxide and water vapor. Thus, the lower portion of the troposphere—more than 50% of all air lies below the altitude of the summit of Mt. Everest—forms a blanket of air keeping the surface warm. This is the Greenhouse Effect. The higher one goes in altitude, the less of this blanket there is to keep in the heat. Thus, higher elevations, such as mountains, are colder than surrounding lowlands. Air temperature in the lowest layer of the atmosphere, the troposphere, decreases with gains in altitude. The rate at which the temperature drops with elevation, called the environmental lapse rate, is not constant (it can fluctuate throughout the day or seasonally and also regionally), but a normal lapse rate is 5.5°C per 1,000 m (3°F per 3,000 ft). The temperature continues to drop up to a height of about 9–16 km, where it does not decrease further. However, this is higher than the highest mountaintop.

Mountains are generally less preferable for human habitation than lowlands; the weather is often harsher, and there is little level ground suitable for agriculture. At very high altitudes, there is less oxygen in the air and less protection against solar radiation (UV). Acute mountain sickness (caused by hypoxia—a lack of oxygen in the blood) affects over half of lowlanders who spend more than a few hours above 3,500 meters (11,483 ft).

Many mountains and mountain ranges throughout the world have been left in their natural state, and are today primarily used for recreation, while others are used for logging, mining, grazing, or see little use. Some mountains offer spectacular views from their summits, while others are densely wooded. Summit accessibility is affected by height, steepness, latitude, terrain, weather. Roads, lifts, or tramways affect accessibility. Hiking, backpacking, mountaineering, rock climbing, ice climbing, downhill skiing, and snowboarding are recreational activities enjoyed on mountains. Mountains that support heavy recreational use (especially downhill skiing) are often the locations of mountain resorts.

Mountains are made up of earth and rock materials. The outermost layer of the Earth or the Earth's crust is composed of six plates. When two plates move or collide each other, vast land areas are uplifted, resulting in the formation of mountains.

## Types

Classified by the geological processes that shape them, there are five major types of mountains:

Fold mountains

Fold mountains are the most common type of mountains. Examples of fold mountains are the Himalayas (Asia), the Alps (Europe). They are formed due to collision of two plates, causing folding of the Earth's crust. The fold that descends on both sides is called anticline; whereas, the fold that ascends from a common low point (on both sides) is called syncline.

#### Fault-Block mountains

As the name suggests, fault-block mountains or fault mountains are formed when blocks of rock materials slide along faults in the Earth's crust. There are two types of block mountains, namely the lifted and tilted. In the former type, the mountain has two steep sides; whereas, the tilted type has one steep side and gentle sloping side. Example of fault-block mountain is the Sierra Nevada mountains (North America).

#### Volcanic mountains

Volcanic mountains are formed due to volcanic eruptions, for e.g. Mount Fuji (Japan). They are formed when volcanic magma erupts and piles up on the surface of the Earth.

#### Dome mountains

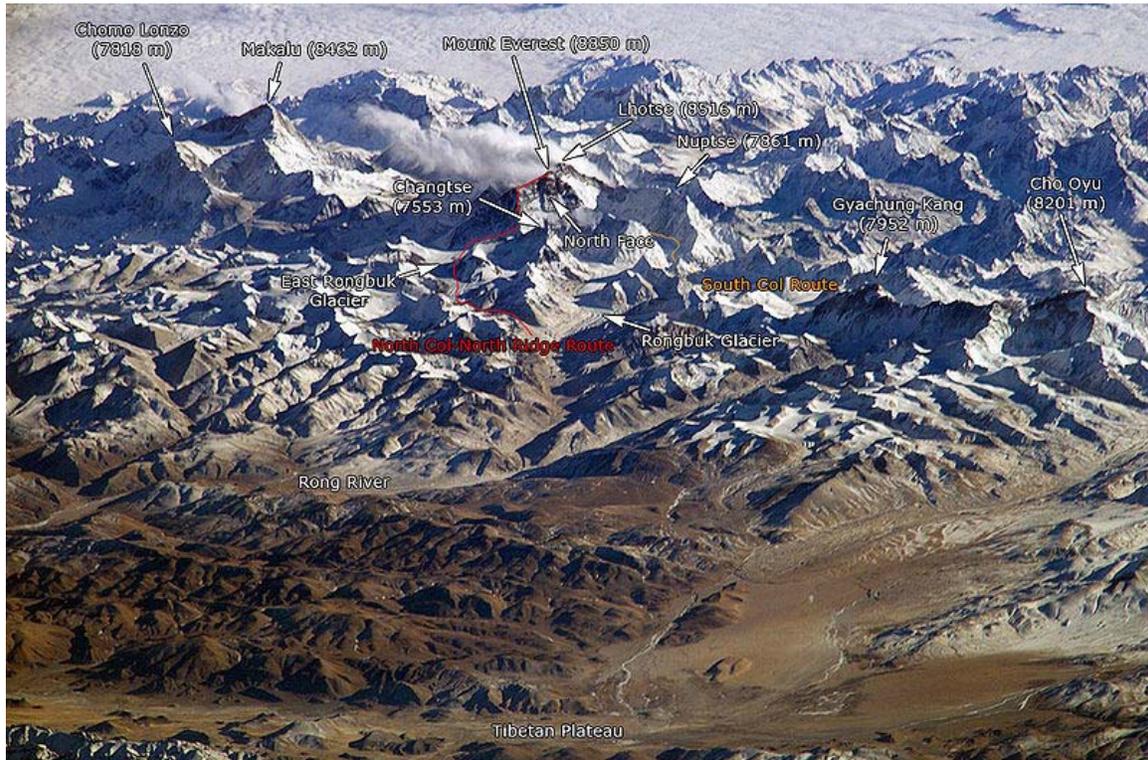
Dome mountains are formed when the hot magma rises from the mantle and uplifts the overlying sedimentary layer of the Earth's crust. In the process, the magma is not erupted, but it cools down and forms the core of the mountain. Example of dome mountain is the Navajo Mountain in Utah. They are called dome mountains due to their appearance that resembles dome shape.

#### Plateau mountains

Plateau mountains are pseudo mountains that are formed because of erosion. An example of plateau mountain is the Catskill Mountains (New York). They usually occur near the fold mountain ranges.

Some mountains are formed as a result of several of the Earth's forces. Though the Rockies in North America is formed due to folding, there are mountains in the same range that are formed by faulting and doming. In nature, there is a continuous process of glaciation, soil erosion, and mechanical and chemical weathering, which altogether play a major role in altering the shape and characteristics of mountains.

# Geology



The Himalayan mountain range with Mount Everest.

A mountain is usually produced by the movement of lithospheric plates, either orogenic movement or epeirogenic movement. The compressional forces, isostatic uplift and intrusion of igneous matter forces surface rock upward, creating a landform higher than the surrounding features. The height of the feature makes it either a hill or, if higher and steeper, a mountain. The absolute heights of features termed mountains and hills vary greatly according to an area's terrain. The major mountains tend to occur in long linear arcs, indicating tectonic plate boundaries and activity. Two types of mountain are formed depending on how the rock reacts to the tectonic forces—block mountains or fold mountains.

Compressional forces in continental collisions may cause the compressed region to thicken, so the upper surface is forced upward. To balance the weight of the earth surface, much of the compressed rock is forced *downward*, producing deep "mountain roots". Mountains therefore form downward as well as upward. However, in some continental collisions part of one continent may simply *override* part of the others, crumpling in the process.

Some isolated mountains were produced by volcanoes, including many apparently small islands that reach a great height above the ocean floor.



Blue Ridge Mountains in Shenandoah National Park, Virginia, USA

Block mountains are created when large areas are widely broken up by faults creating large vertical displacements. This occurrence is fairly common. The uplifted blocks are block mountains or *horsts*. The intervening dropped blocks are termed *graben*: these can be small or form extensive rift valley systems. This form of landscape can be seen in East Africa, the Vosges, the Basin and Range province of Western North America and the Rhine valley. These areas often occur when the regional stress is extensional and the crust is thinned.

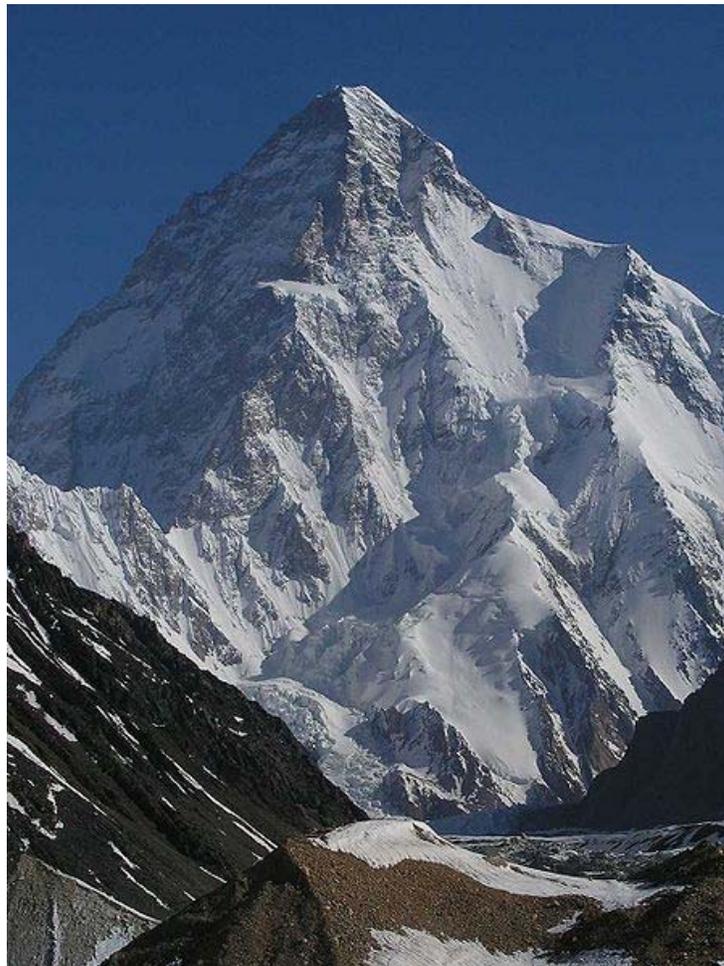
The mid-ocean ridges are often referred to as undersea mountain ranges due to their bathymetric prominence.

Rock that does not fault may fold, either symmetrically or asymmetrically. The upfolds are *anticlines* and the downfolds are *synclines*: in asymmetric folding there may also be recumbent and overturned folds. The Jura Mountains are an example of folding. Over time, erosion can bring about an inversion of relief: the soft upthrust rock is worn away so the anticlines are actually lower than the tougher, more compressed rock of the synclines.

## Gallery



Mount Everest, 8,848 metres (29,029 ft), Himalayas, Nepal, China.



K2, 8,611 metres (28,251 ft), Gilgit-Baltistan, Pakistan.



Nilkantha (mountain), 6,597 metres (21,644 ft), Himalayas, India.

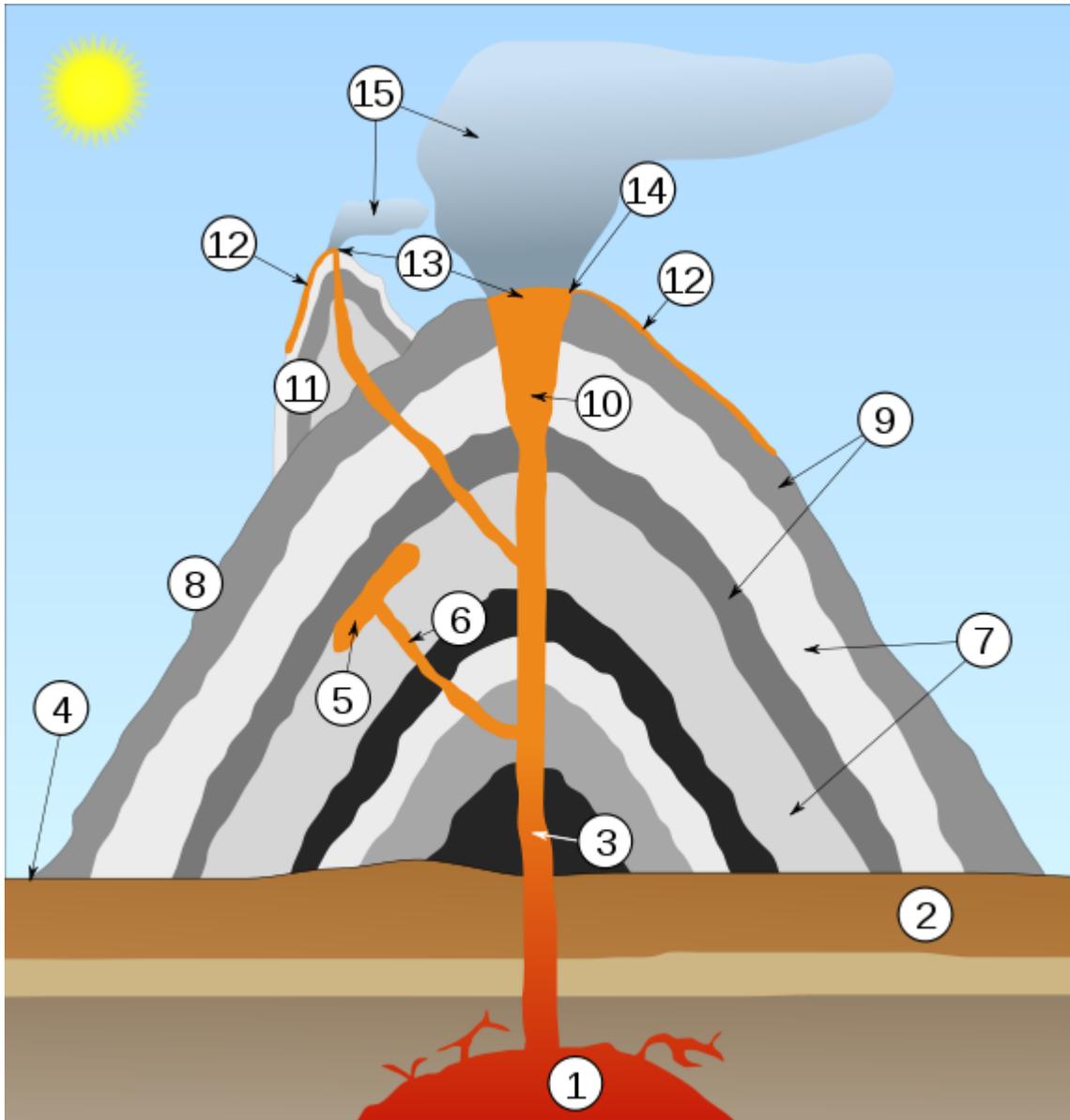


Mount Kilimanjaro, 5,895 metres (19,341 ft), Tanzania.

# Volcano



Cleveland Volcano in the Aleutian Islands of Alaska photographed from the International Space Station, May 2006



**Cross-section through a stratovolcano (vertical scale is exaggerated):**

- |   |  |
|---|--|
| 1. Large magma chamber                  | 9. Layers of lava emitted by the volcano |
| 2. Bedrock                              | 10. Throat                               |
| 3. Conduit (pipe)                       | 11. Parasitic cone                       |
| 4. Base                                 | 12. Lava flow                            |
| 5. Sill                                 | 13. Vent                                 |
| 6. Dike                                 | 14. Crater                               |
| 7. Layers of ash emitted by the volcano | 15. Ash cloud                            |
| 8. Flank                                |  |



Pinatubo ash plume reaching a height of 19 km, 3 days before the climactic eruption of 15 June 1991

A **volcano** is an opening, or rupture, in a planet's surface or crust, which allows hot magma, volcanic ash and gases to escape from below the surface.

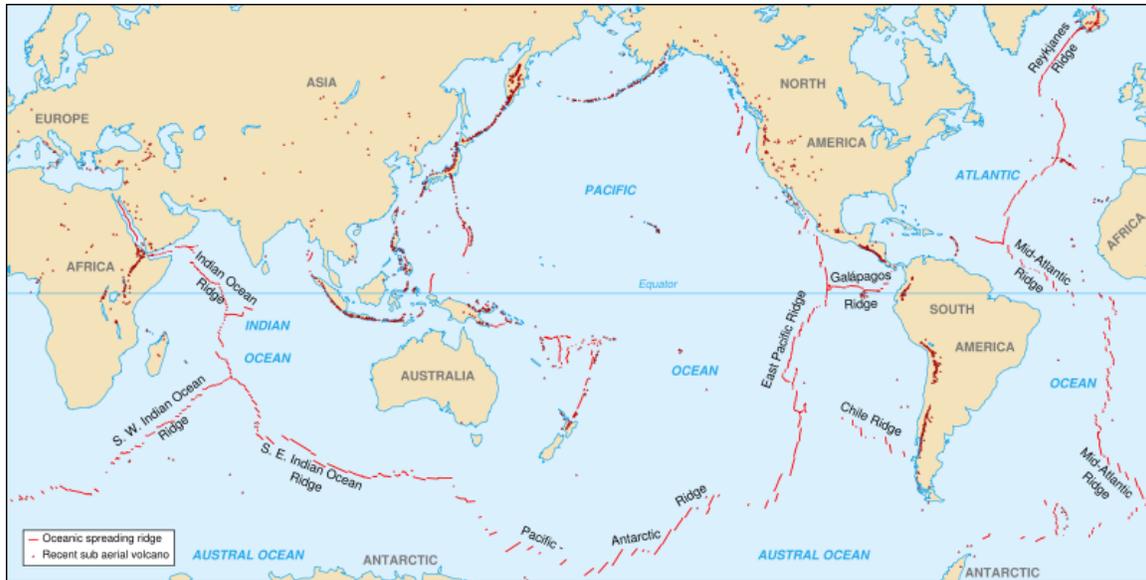
Volcanoes are generally found where tectonic plates are diverging or converging. A mid-oceanic ridge, for example the Mid-Atlantic Ridge, has examples of volcanoes caused by divergent tectonic plates pulling apart; the Pacific Ring of Fire has examples of volcanoes caused by convergent tectonic plates coming together. By contrast, volcanoes are usually not created where two tectonic plates slide past one another. Volcanoes can also form where there is stretching and thinning of the Earth's crust (called "non-hotspot intraplate volcanism"), such as in the East African Rift, the Wells Gray-Clearwater volcanic field and the Rio Grande Rift in North America.

Volcanoes can be caused by mantle plumes. These so-called hotspots, for example at Hawaii, can occur far from plate boundaries. Hotspot volcanoes are also found elsewhere in the solar system, especially on rocky planets and moons.

## Etymology

The word *volcano* is derived from the name of Vulcano, a volcanic island in the Aeolian Islands of Italy whose name in turn originates from Vulcan, the name of a god of fire in Roman mythology. The study of volcanoes is called volcanology, sometimes spelled *vulcanology*.

## Plate tectonics and hotspots



Map showing the divergent plate boundaries (OSR – Oceanic Spreading Ridges) and recent sub aerial volcanoes.

### Divergent plate boundaries

At the mid-oceanic ridges, two tectonic plates diverge from one another. New oceanic crust is being formed by hot molten rock slowly cooling and solidifying. The crust is very thin at mid-oceanic ridges due to the pull of the tectonic plates. The release of pressure due to the thinning of the crust leads to adiabatic expansion, and the partial melting of the mantle causing volcanism and creating new oceanic crust. Most divergent plate boundaries are at the bottom of the oceans, therefore most volcanic activity is submarine, forming new seafloor. Black smokers or deep sea vents are an example of this kind of volcanic activity. Where the mid-oceanic ridge is above sea-level, volcanic islands are formed, for example, Iceland.



Mount Rinjani eruption in 1994, in Lombok, Indonesia

### **Convergent plate boundaries**

Subduction zones are places where two plates, usually an oceanic plate and a continental plate, collide. In this case, the oceanic plate subducts, or submerges under the continental plate forming a deep ocean trench just offshore. Water released from the subducting plate lowers the melting temperature of the overlying mantle wedge, creating magma. This magma tends to be very viscous due to its high silica content, so often does not reach the surface and cools at depth. When it does reach the surface, a volcano is formed. Typical examples for this kind of volcano are Mount Etna and the volcanoes in the Pacific Ring of Fire.



Lava enters the Pacific at the Big Island of Hawaii

## Hotspots

Hotspots are not usually located on the ridges of tectonic plates, but above mantle plumes, where the convection of the Earth's mantle creates a column of hot material that rises until it reaches the crust, which tends to be thinner than in other areas of the Earth. The temperature of the plume causes the crust to melt and form pipes, which can vent magma. Because the tectonic plates move whereas the mantle plume remains in the same place, each volcano becomes dormant after a while and a new volcano is then formed as the plate shifts over the hotspot. The Hawaiian Islands are thought to be formed in such a manner, as well as the Snake River Plain, with the Yellowstone Caldera being the part of the North American plate currently above the hot spot.

## Volcanic features



Conical Mount Fuji in Japan, at sunrise from Lake Kawaguchi (2005)

The most common perception of a volcano is of a conical mountain, spewing lava and poisonous gases from a crater at its summit. This describes just one of many types of volcano, and the features of volcanoes are much more complicated. The structure and behavior of volcanoes depends on a number of factors. Some volcanoes have rugged peaks formed by lava domes rather than a summit crater, whereas others present landscape features such as massive plateaus. Vents that issue volcanic material (lava, which is what magma is called once it has escaped to the surface, and ash) and gases (mainly steam and magmatic gases) can be located anywhere on the landform. Many of these vents give rise to smaller cones such as Pu'u 'Ō'ō on a flank of Hawaii's Kīlauea.



Lakagigar fissure vent in Iceland, source of the major world climate alteration of 1783-84. Volcanic eruptions are experienced somewhere in Iceland on an average of once every five years.



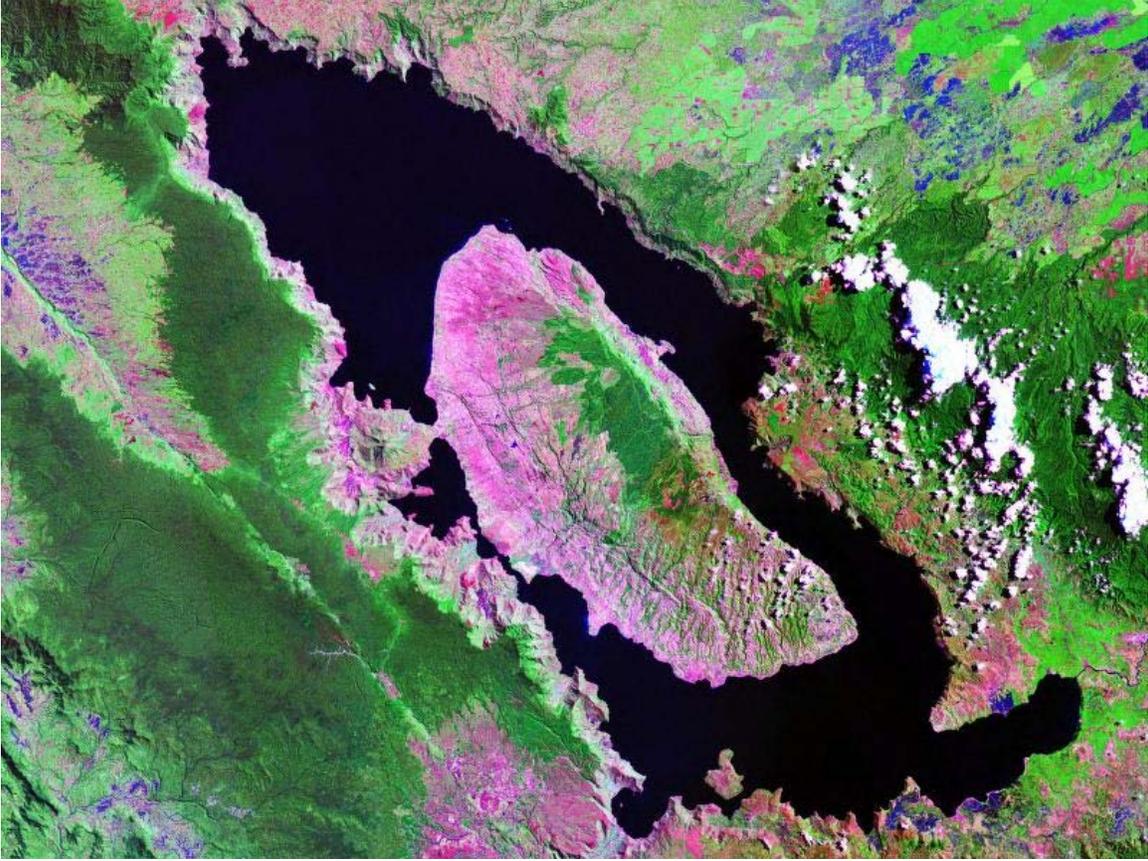
January 2009 image of the rhyolitic lava dome of Chaitén Volcano, southern Chile during its 2008-2009 eruption



Holocene cinder cone volcano on State Highway 18 near Veyo, Utah



Mayon, near-perfect stratovolcano in the Philippines



The Lake Toba volcano created a caldera 100 km long.



Pillow lava (NOAA)



Herðubreið, one of the tuyas in Iceland



Mud volcano on Taman Peninsula, Russia

Other types of volcano include cryovolcanoes (or ice volcanoes), particularly on some moons of Jupiter, Saturn and Neptune; and mud volcanoes, which are formations often not associated with known magmatic activity. Active mud volcanoes tend to involve temperatures much lower than those of igneous volcanoes, except when a mud volcano is actually a vent of an igneous volcano.

### **Fissure vents**

Volcanic **fissure vents** are flat, linear cracks through which lava emerges.

### **Shield volcanoes**

**Shield volcanoes**, so named for their broad, shield-like profiles, are formed by the eruption of low-viscosity lava that can flow a great distance from a vent, but not generally explode catastrophically. Since low-viscosity magma is typically low in silica, shield volcanoes are more common in oceanic than continental settings. The Hawaiian volcanic chain is a series of shield cones, and they are common in Iceland, as well.

## **Lava domes**

**Lava domes** are built by slow eruptions of highly viscous lavas. They are sometimes formed within the crater of a previous volcanic eruption (as in Mount Saint Helens), but can also form independently, as in the case of Lassen Peak. Like stratovolcanoes, they can produce violent, explosive eruptions, but their lavas generally do not flow far from the originating vent.

## **Cryptodomes**

**Cryptodomes** are formed when viscous lava forces its way up and causes a bulge. The 1980 eruption of Mount St. Helens was an example. Lava was under great pressure and forced a bulge in the mountain, which was unstable and slid down the north side.

## **Volcanic cones (cinder cones)**

**Volcanic cones** or **cinder cones** are the result from eruptions that erupt mostly small pieces of scoria and pyroclastics (both resemble cinders, hence the name of this volcano type) that build up around the vent. These can be relatively short-lived eruptions that produce a cone-shaped hill perhaps 30 to 400 meters high. Most cinder cones erupt only once. Cinder cones may form as flank vents on larger volcanoes, or occur on their own. Parícutin in Mexico and Sunset Crater in Arizona are examples of cinder cones. In New Mexico, Caja del Rio is a volcanic field of over 60 cinder cones.

## **Stratovolcanoes (composite volcanoes)**

**Stratovolcanoes** or **composite volcanoes** are tall conical mountains composed of lava flows and other ejecta in alternate layers, the strata that give rise to the name. Stratovolcanoes are also known as composite volcanoes, created from several structures during different kinds of eruptions. Strato/composite volcanoes are made of cinders, ash and lava. Cinders and ash pile on top of each other, lava flows on top of the ash, where it cools and hardens, and then the process begins again. Classic examples include Mt. Fuji in Japan, Mayon Volcano in the Philippines, and Mount Vesuvius and Stromboli in Italy.

In recorded history, explosive eruptions by stratovolcanoes have posed the greatest hazard to civilizations, as ash is produced by an explosive eruption. No supervolcano erupted in recorded history. Shield volcanoes have not an enormous pressure build up from the lava flow. Fissure vents and monogenetic volcanic fields (volcanic cones) have not powerful explosive eruptions, as they are many times under extension. Stratovolcanoes (30-35°) are steeper than shield volcanoes (generally 5-10°), their loose tephra are material for dangerous lahars.

## **Supervolcanoes**

A **supervolcano** is a large volcano that usually has a large caldera and can potentially produce devastation on an enormous, sometimes continental, scale. Such eruptions would

be able to cause severe cooling of global temperatures for many years afterwards because of the huge volumes of sulfur and ash erupted. They are the most dangerous type of volcano. Examples include Yellowstone Caldera in Yellowstone National Park and Valles Caldera in New Mexico (both western United States), Lake Taupo in New Zealand, Lake Toba in Sumatra, Indonesia and Ngorogoro Crater in Tanzania. Supervolcanoes are hard to identify centuries later, given the enormous areas they cover. Large igneous provinces are also considered supervolcanoes because of the vast amount of basalt lava erupted, but are non-explosive.

## **Submarine volcanoes**

**Submarine volcanoes** are common features on the ocean floor. Some are active and, in shallow water, disclose their presence by blasting steam and rocky debris high above the surface of the sea. Many others lie at such great depths that the tremendous weight of the water above them prevents the explosive release of steam and gases, although they can be detected by hydrophones and discoloration of water because of volcanic gases. Pumice rafts may also appear. Even large submarine eruptions may not disturb the ocean surface. Because of the rapid cooling effect of water as compared to air, and increased buoyancy, submarine volcanoes often form rather steep pillars over their volcanic vents as compared to above-surface volcanoes. They may become so large that they break the ocean surface as new islands. Pillow lava is a common eruptive product of submarine volcanoes. Hydrothermal vents are common near these volcanoes, and some support peculiar ecosystems based on dissolved minerals.

## **Subglacial volcanoes**

**Subglacial volcanoes** develop underneath icecaps. They are made up of flat lava which flows at the top of extensive pillow lavas and palagonite. When the icecap melts, the lavas on the top collapse, leaving a flat-topped mountain. These volcanoes are also called table mountains, tuyas or (uncommonly) mobergs. Very good examples of this type of volcano can be seen in Iceland, however, there are also tuyas in British Columbia. The origin of the term comes from Tuya Butte, which is one of the several tuyas in the area of the Tuya River and Tuya Range in northern British Columbia. Tuya Butte was the first such landform analyzed and so its name has entered the geological literature for this kind of volcanic formation. The Tuya Mountains Provincial Park was recently established to protect this unusual landscape, which lies north of Tuya Lake and south of the Jennings River near the boundary with the Yukon Territory.

## **Mud volcanoes**

**Mud volcanoes** or **mud domes** are formations created by geo-excreted liquids and gases, although there are several different processes which may cause such activity. The largest structures are 10 kilometers in diameter and reach 700 meters high.

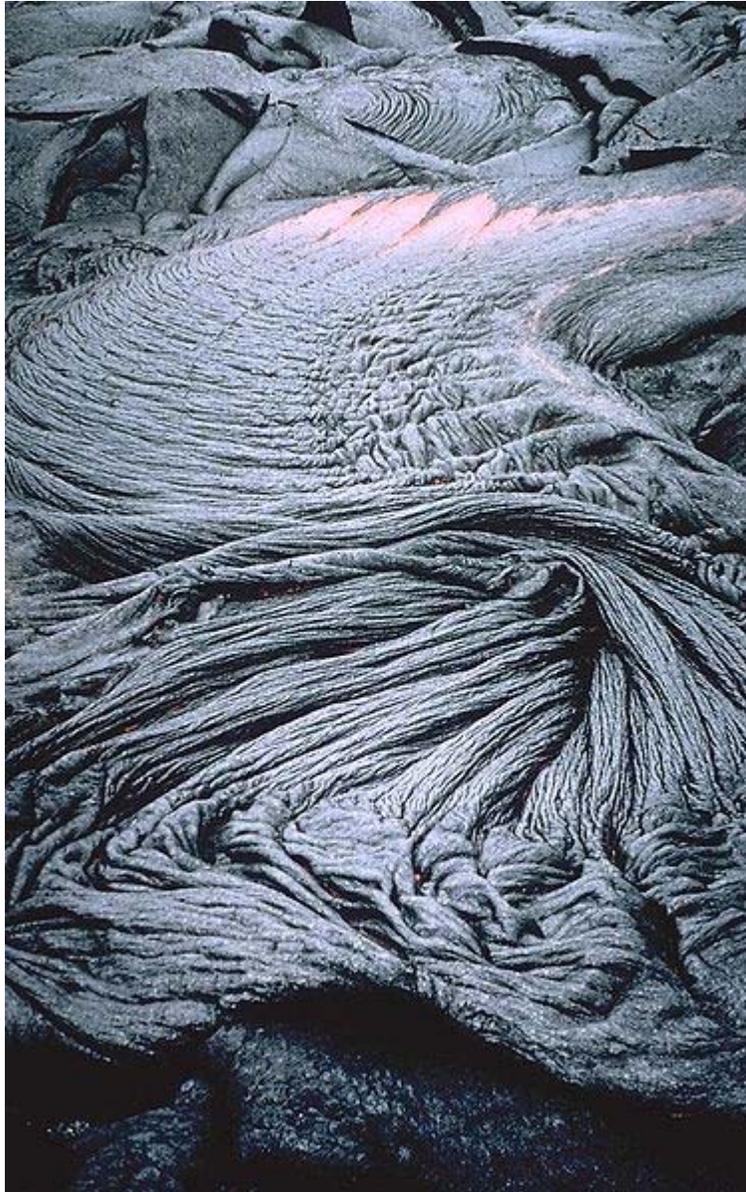
## Erupted material



Pāhoehoe Lava flow on Hawaii. The picture shows overflows of a main lava channel



The Stromboli volcano off the coast of Sicily has erupted continuously for thousands of years, giving rise to the term strombolian eruption.



Pāhoehoe lava from Kīlauea, Hawaii

## Lava composition

Another way of classifying volcanoes is by the *composition of material erupted* (lava), since this affects the shape of the volcano. Lava can be broadly classified into 4 different compositions (Cas & Wright, 1987):

- If the erupted magma contains a high percentage (>63%) of silica, the lava is called felsic.
  - Felsic lavas (dacites or rhyolites) tend to be highly viscous (not very fluid) and are erupted as domes or short, stubby flows. Viscous lavas tend to form stratovolcanoes or lava domes. Lassen Peak in California is an

example of a volcano formed from felsic lava and is actually a large lava dome.

- Because siliceous magmas are so viscous, they tend to trap volatiles (gases) that are present, which cause the magma to erupt catastrophically, eventually forming stratovolcanoes. Pyroclastic flows (ignimbrites) are highly hazardous products of such volcanoes, since they are composed of molten volcanic ash too heavy to go up into the atmosphere, so they hug the volcano's slopes and travel far from their vents during large eruptions. Temperatures as high as 1,200 °C are known to occur in pyroclastic flows, which will incinerate everything flammable in their path and thick layers of hot pyroclastic flow deposits can be laid down, often up to many meters thick. Alaska's Valley of Ten Thousand Smokes, formed by the eruption of Novarupta near Katmai in 1912, is an example of a thick pyroclastic flow or ignimbrite deposit. Volcanic ash that is light enough to be erupted high into the Earth's atmosphere may travel many kilometres before it falls back to ground as a tuff.
- If the erupted magma contains 52–63% silica, the lava is of *intermediate* composition.
  - These "andesitic" volcanoes generally only occur above subduction zones (e.g. Mount Merapi in Indonesia).
  - Andesitic lava is typically formed at convergent boundary margins of tectonic plates, by several processes:
    - Hydration melting of peridotite and fractional crystallization
    - Melting of subducted slab containing sediments
    - Magma mixing between felsic rhyolitic and mafic basaltic magmas in an intermediate reservoir prior to emplacement or lava flow.
- If the erupted magma contains <52% and >45% silica, the lava is called mafic (because it contains higher percentages of magnesium (Mg) and iron (Fe)) or basaltic. These lavas are usually much less viscous than rhyolitic lavas, depending on their eruption temperature; they also tend to be hotter than felsic lavas. Mafic lavas occur in a wide range of settings:
  - At mid-ocean ridges, where two oceanic plates are pulling apart, basaltic lava erupts as pillows to fill the gap;
  - Shield volcanoes (e.g. the Hawaiian Islands, including Mauna Loa and Kilauea), on both oceanic and continental crust;
  - As continental flood basalts.
- Some erupted magmas contain ≤45% silica and produce ultramafic lava. Ultramafic flows, also known as komatiites, are very rare; indeed, very few have been erupted at the Earth's surface since the Proterozoic, when the planet's heat flow was higher. They are (or were) the hottest lavas, and probably more fluid than common mafic lavas.

## Lava texture

Two types of lava are named according to the surface texture: 'A'a and pāhoehoe, both Hawaiian words. 'A'a is characterized by a rough, clinkery surface and is the typical

texture of viscous lava flows. However, even basaltic or mafic flows can be erupted as 'a'a flows, particularly if the eruption rate is high and the slope is steep.

Pāhoehoe is characterized by its smooth and often ropey or wrinkly surface and is generally formed from more fluid lava flows. Usually, only mafic flows will erupt as pāhoehoe, since they often erupt at higher temperatures or have the proper chemical make-up to allow them to flow with greater fluidity.

## **Volcanic activity**



Active volcano Mount St. Helens shortly after the eruption of 18 May 1980



Damavand, the highest volcano in Asia, is a potentially active volcano with fumaroles and solfatara near its summit.



Fresco of Bacchus and Agathodaemon with Mount Vesuvius, as seen in Pompeii's House of the Centenary.



Fourpeaked volcano, Alaska, in September 2007, after being thought extinct for over 10,000 years.

## Popular classification of volcanoes

### Active

A popular way of classifying magmatic volcanoes is by their frequency of eruption, with those that erupt regularly called **active**, those that have erupted in historical *times* but are now quiet called **dormant**, and those that have not erupted in historical times called **extinct**. However, these popular classifications—extinct in particular—are practically meaningless to scientists. They use classifications which refer to a particular volcano's formative and eruptive processes and resulting shapes, which was explained above.

There is no real consensus among volcanologists on how to define an "active" volcano. The lifespan of a volcano can vary from months to several million years, making such a distinction sometimes meaningless when compared to the lifespans of humans or even civilizations. For example, many of Earth's volcanoes have erupted dozens of times in the past few thousand years but are not currently showing signs of eruption. Given the long lifespan of such volcanoes, they are very active. By human lifespans, however, they are not.

Scientists usually consider a volcano to be **erupting or likely to erupt** if it is currently erupting, or showing signs of unrest such as unusual earthquake activity or significant new gas emissions. Most scientists consider a volcano *active* if it has erupted in holocene times. **Historic times** is another timeframe for *active*. But it is important to note that the span of recorded history differs from region to region. In China and the Mediterranean, recorded history reaches back more than 3,000 years but in the Pacific Northwest of the United States and Canada, it reaches back less than 300 years, and in Hawaii and New Zealand, only around 200 years. The Smithsonian Global Volcanism Program's definition of *active* is having erupted within the last 10,000 years (the 'holocene' period).

Presently there are about 500 active volcanoes in the world – the majority following along the Pacific 'Ring of Fire' – and around 50 of these erupt each year. The United States is home to 50 active volcanoes. There are more than 1,500 potentially active volcanoes. An estimated 500 million people live near active volcanoes.

### Extinct

**Extinct** volcanoes are those that scientists consider unlikely to erupt again, because the volcano no longer has a lava supply. Examples of extinct volcanoes are many volcanoes on the Hawaiian – Emperor seamount chain in the Pacific Ocean (extinct because the Hawaii hotspot is centered near the Big Island), Hohentwiel, Shiprock, and Paricutin (which is monogenetic). Otherwise, whether a volcano is truly extinct is often difficult to determine. Since "supervolcano" calderas can have eruptive lifespans sometimes measured in millions of years, a caldera that has not produced an eruption in tens of thousands of years is likely to be considered dormant instead of extinct.

## Dormant

It is difficult to distinguish an extinct volcano from a **dormant** one. Volcanoes are often considered to be extinct if there are no written records of its activity. Nevertheless volcanoes may remain dormant for a long period of time, Yellowstone has a repose/recharge period of around 700 ka and Toba of around 380 ka. Vesuvius was described by Roman writers as having been covered with gardens and vineyards before its famous eruption of AD 79, which destroyed the towns of Herculaneum and Pompeii. Before the catastrophic eruption of 1991, Pinatubo was an inconspicuous volcano, unknown to most people in the surrounding areas. More recently, the long-dormant Soufrière Hills volcano on the island of Montserrat was thought to be extinct before activity resumed in 1995. Another recent example is Fourpeaked Mountain in Alaska, which, prior to its eruption in September 2006, had not erupted since before 8000 BC and was long thought to be extinct.

## Notable volcanoes



Koryaksky volcano towering over Petropavlovsk-Kamchatsky on Kamchatka Peninsula, Far Eastern Russia.

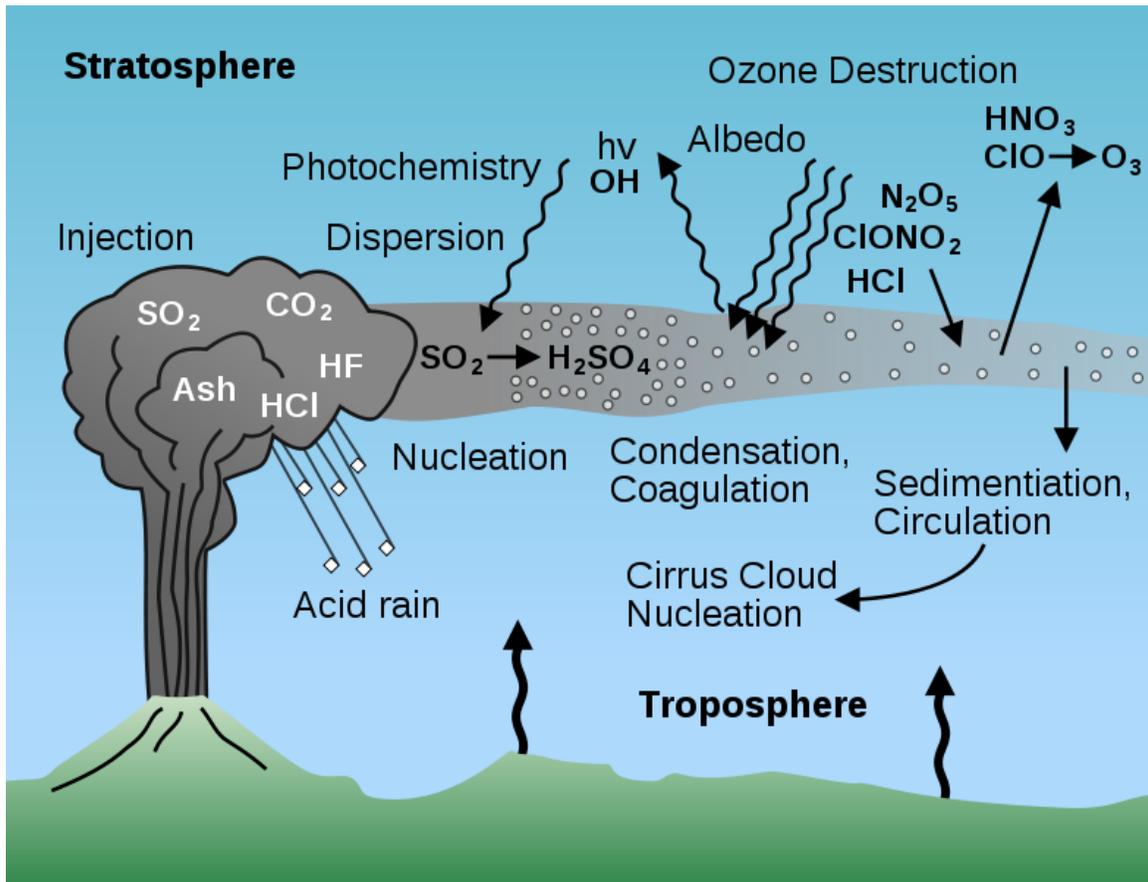


Mount Teide on the island of Tenerife (Spain).

The 16 current Decade Volcanoes are:

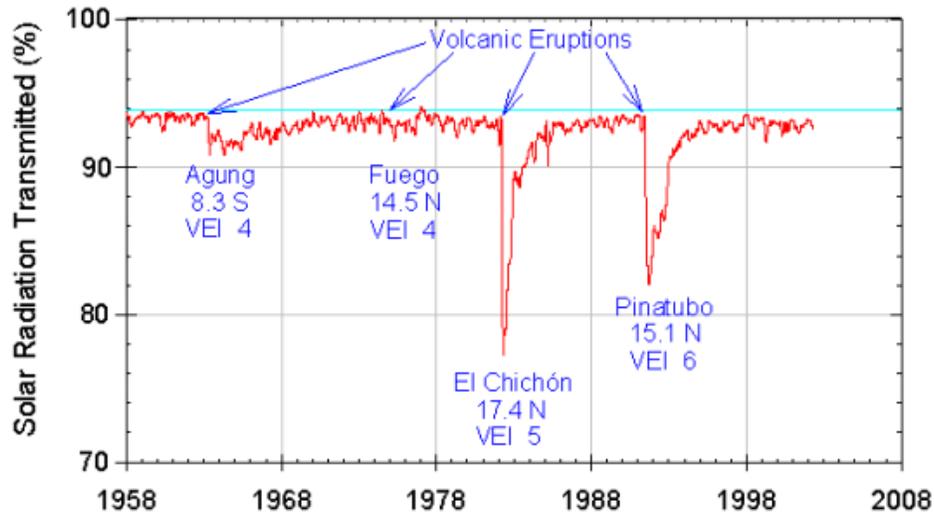
- Avachinsky-Koryaksky, Kamchatka, Russia
- Nevado de Colima, Jalisco and Colima, Mexico
- Mount Etna, Sicily, Italy
- Galeras, Nariño, Colombia
- Mauna Loa, Hawaii, USA
- Mount Merapi, Central Java, Indonesia
- Mount Nyiragongo, Democratic Republic of the Congo
- Mount Rainier, Washington, USA
- Sakurajima, Kagoshima Prefecture, Japan
- Santa Maria/Santiaguito, Guatemala
- Santorini, Cyclades, Greece
- Taal Volcano, Luzon, Philippines
- Teide, Canary Islands, Spain
- Ulawun, New Britain, Papua New Guinea
- Mount Unzen, Nagasaki Prefecture, Japan
- Vesuvius, Naples, Italy

# Effects of volcanoes

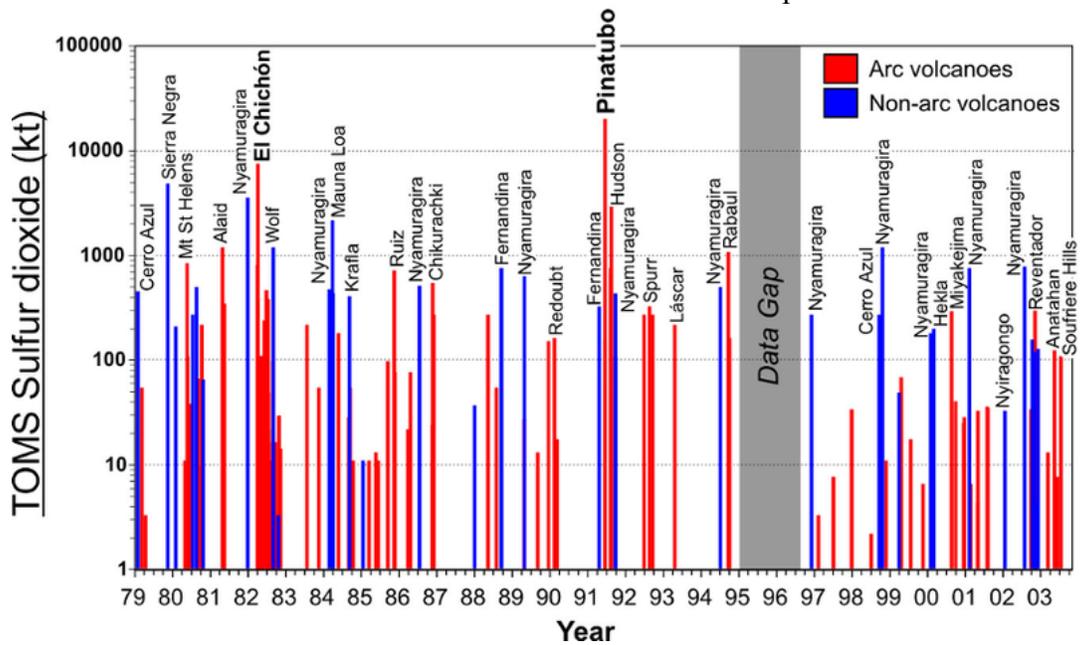


Volcanic "injection"

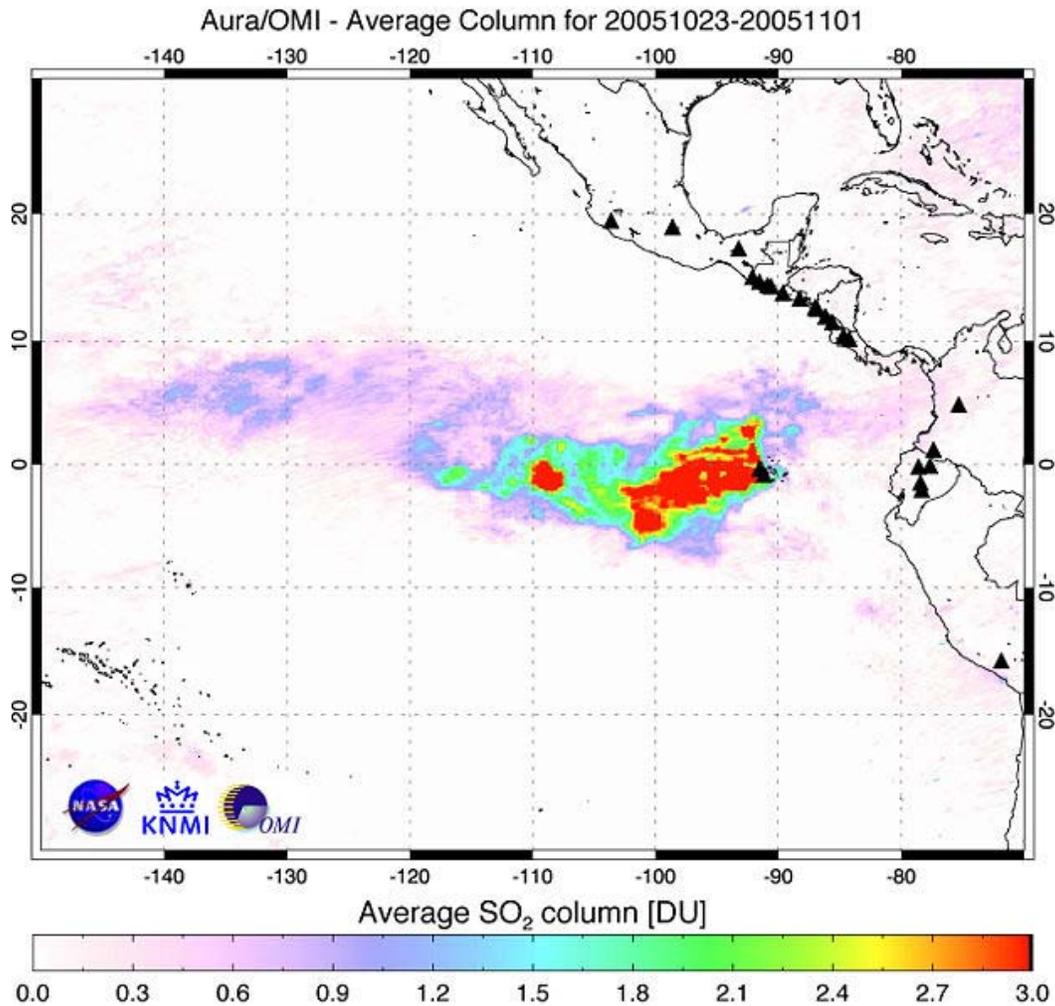
## Mauna Loa Observatory Atmospheric Transmission



Solar radiation reduction from volcanic eruptions



Sulfur dioxide emissions by volcanoes.



Average concentration of sulfur dioxide over the Sierra Negra Volcano (Galapagos Islands) from October 23–November 1, 2005

There are many different types of volcanic eruptions and associated activity: phreatic eruptions (steam-generated eruptions), explosive eruption of high-silica lava (e.g., rhyolite), effusive eruption of low-silica lava (e.g., basalt), pyroclastic flows, lahars (debris flow) and carbon dioxide emission. All of these activities can pose a hazard to humans. Earthquakes, hot springs, fumaroles, mud pots and geysers often accompany volcanic activity.

The concentrations of different volcanic gases can vary considerably from one volcano to the next. Water vapor is typically the most abundant volcanic gas, followed by carbon dioxide and sulfur dioxide. Other principal volcanic gases include hydrogen sulfide, hydrogen chloride, and hydrogen fluoride. A large number of minor and trace gases are also found in volcanic emissions, for example hydrogen, carbon monoxide, halocarbons, organic compounds, and volatile metal chlorides.

Large, explosive volcanic eruptions inject water vapor ( $\text{H}_2\text{O}$ ), carbon dioxide ( $\text{CO}_2$ ), sulfur dioxide ( $\text{SO}_2$ ), hydrogen chloride ( $\text{HCl}$ ), hydrogen fluoride ( $\text{HF}$ ) and ash (pulverized rock and pumice) into the stratosphere to heights of 16–32 kilometres (10–20 mi) above the Earth's surface. The most significant impacts from these injections come from the conversion of sulfur dioxide to sulfuric acid ( $\text{H}_2\text{SO}_4$ ), which condenses rapidly in the stratosphere to form fine sulfate aerosols. The aerosols increase the Earth's albedo—its reflection of radiation from the Sun back into space - and thus cool the Earth's lower atmosphere or troposphere; however, they also absorb heat radiated up from the Earth, thereby warming the stratosphere. Several eruptions during the past century have caused a decline in the average temperature at the Earth's surface of up to half a degree (Fahrenheit scale) for periods of one to three years — sulfur dioxide from the eruption of Huaynaputina probably caused the Russian famine of 1601 - 1603.

One proposed volcanic winter happened c. 70,000 years ago following the supereruption of Lake Toba on Sumatra island in Indonesia. According to the Toba catastrophe theory to which some anthropologists and archeologists subscribe, it had global consequences, killing most humans then alive and creating a population bottleneck that affected the genetic inheritance of all humans today. The 1815 eruption of Mount Tambora created global climate anomalies that became known as the "Year Without a Summer" because of the effect on North American and European weather. Agricultural crops failed and livestock died in much of the Northern Hemisphere, resulting in one of the worst famines of the 19th century. The freezing winter of 1740-41, which led to widespread famine in northern Europe, may also owe its origins to a volcanic eruption.

It has been suggested that volcanic activity caused or contributed to the End-Ordovician, Permian-Triassic, Late Devonian mass extinctions, and possibly others. The massive eruptive event which formed the Siberian Traps, one of the largest known volcanic events of the last 500 million years of Earth's geological history, continued for a million years and is considered to be the likely cause of the "Great Dying" about 250 million years ago, which is estimated to have killed 90% of species existing at the time.

The sulfate aerosols also promote complex chemical reactions on their surfaces that alter chlorine and nitrogen chemical species in the stratosphere. This effect, together with increased stratospheric chlorine levels from chlorofluorocarbon pollution, generates chlorine monoxide ( $\text{ClO}$ ), which destroys ozone ( $\text{O}_3$ ). As the aerosols grow and coagulate, they settle down into the upper troposphere where they serve as nuclei for cirrus clouds and further modify the Earth's radiation balance. Most of the hydrogen chloride ( $\text{HCl}$ ) and hydrogen fluoride ( $\text{HF}$ ) are dissolved in water droplets in the eruption cloud and quickly fall to the ground as acid rain. The injected ash also falls rapidly from the stratosphere; most of it is removed within several days to a few weeks. Finally, explosive volcanic eruptions release the greenhouse gas carbon dioxide and thus provide a deep source of carbon for biogeochemical cycles.



Rainbow and volcanic ash with sulfur dioxide emissions from Halema`uma`u vent.

Gas emissions from volcanoes are a natural contributor to acid rain. Volcanic activity releases about 130 to 230 teragrams (145 million to 255 million short tons) of carbon dioxide each year. Volcanic eruptions may inject aerosols into the Earth's atmosphere. Large injections may cause visual effects such as unusually colorful sunsets and affect global climate mainly by cooling it. Volcanic eruptions also provide the benefit of adding nutrients to soil through the weathering process of volcanic rocks. These fertile soils assist the growth of plants and various crops. Volcanic eruptions can also create new islands, as the magma cools and solidifies upon contact with the water.

Ash thrown into the air by eruptions can present a hazard to aircraft, especially jet aircraft where the particles can be melted by the high operating temperature. Dangerous encounters in 1982 after the eruption of Galunggung in Indonesia, and 1989 after the eruption of Mount Redoubt in Alaska raised awareness of this phenomenon. Nine Volcanic Ash Advisory Centers were established by the International Civil Aviation Organization to monitor ash clouds and advise pilots accordingly. The 2010 eruptions of Eyjafjallajökull caused major disruptions to air travel in Europe.

## Volcanoes on other planetary bodies



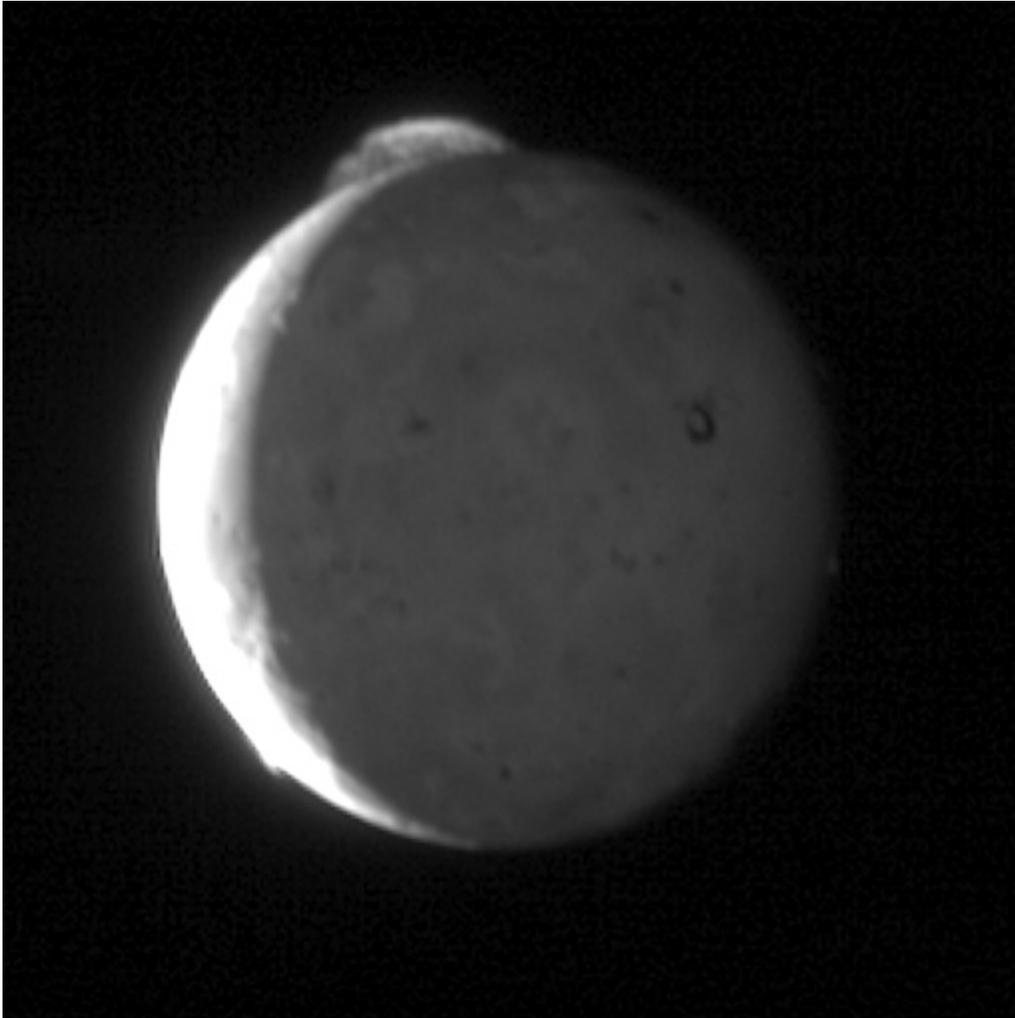
Olympus Mons (Latin, "Mount Olympus") is the tallest known mountain in our solar system, located on the planet Mars.

The Earth's Moon has no large volcanoes and no current volcanic activity, although recent evidence suggests it may still possess a partially molten core. However, the Moon does have many volcanic features such as maria (the darker patches seen on the moon), rilles and domes.

The planet Venus has a surface that is 90% basalt, indicating that volcanism played a major role in shaping its surface. The planet may have had a major global resurfacing

event about 500 million years ago, from what scientists can tell from the density of impact craters on the surface. Lava flows are widespread and forms of volcanism not present on Earth occur as well. Changes in the planet's atmosphere and observations of lightning have been attributed to ongoing volcanic eruptions, although there is no confirmation of whether or not Venus is still volcanically active. However, radar sounding by the Magellan probe revealed evidence for comparatively recent volcanic activity at Venus's highest volcano Maat Mons, in the form of ash flows near the summit and on the northern flank.

There are several extinct volcanoes on Mars, four of which are vast shield volcanoes far bigger than any on Earth. They include Arsia Mons, Ascraeus Mons, Hecates Tholus, Olympus Mons, and Pavonis Mons. These volcanoes have been extinct for many millions of years, but the European *Mars Express* spacecraft has found evidence that volcanic activity may have occurred on Mars in the recent past as well.



The Tvashhtar volcano erupts a plume 330 km (205 mi) above the surface of Jupiter's moon Io.

Jupiter's moon Io is the most volcanically active object in the solar system because of tidal interaction with Jupiter. It is covered with volcanoes that erupt sulfur, sulfur dioxide and silicate rock, and as a result, Io is constantly being resurfaced. Its lavas are the hottest known anywhere in the solar system, with temperatures exceeding 1,800 K (1,500 °C). In February 2001, the largest recorded volcanic eruptions in the solar system occurred on Io. Europa, the smallest of Jupiter's Galilean moons, also appears to have an active volcanic system, except that its volcanic activity is entirely in the form of water, which freezes into ice on the frigid surface. This process is known as cryovolcanism, and is apparently most common on the moons of the outer planets of the solar system.

In 1989 the Voyager 2 spacecraft observed cryovolcanoes (ice volcanoes) on Triton, a moon of Neptune, and in 2005 the Cassini-Huygens probe photographed fountains of frozen particles erupting from Enceladus, a moon of Saturn. The ejecta may be composed of water, liquid nitrogen, dust, or methane compounds. Cassini-Huygens also found evidence of a methane-spewing cryovolcano on the Saturnian moon Titan, which is believed to be a significant source of the methane found in its atmosphere. It is theorized that cryovolcanism may also be present on the Kuiper Belt Object Quaoar.

A 2010 study of the exoplanet COROT-7b, which was detected by transit in 2009, studied that tidal heating from the host star very close to the planet and neighboring planets could generate intense volcanic activity similar to Io.

## **Traditional beliefs about volcanoes**

Many ancient accounts ascribe volcanic eruptions to supernatural causes, such as the actions of gods or demigods. To the ancient Greeks, volcanoes' capricious power could only be explained as acts of the gods, while 16th/17th-century German astronomer Johannes Kepler believed they were ducts for the Earth's tears. One early idea counter to this was proposed by Jesuit Athanasius Kircher (1602–1680), who witnessed eruptions of Mount Etna and Stromboli, then visited the crater of Vesuvius and published his view of an Earth with a central fire connected to numerous others caused by the burning of sulfur, bitumen and coal.

Various explanations were proposed for volcano behavior before the modern understanding of the Earth's mantle structure as a semisolid material was developed. For decades after awareness that compression and radioactive materials may be heat sources, their contributions were specifically discounted. Volcanic action was often attributed to chemical reactions and a thin layer of molten rock near the surface.

## Panoramas



Mount Bromo, East Java, Indonesia.



Crater of Mount Tangkuban Perahu, West Java, Indonesia.



Irazú Volcano, Costa Rica.



Black Rock Volcano an extinct cinder cone near Fillmore, Utah.



Taal Volcano, Philippines.



Crater of Sierra Negra volcano, Isabela island, Galapagos, Ecuador.



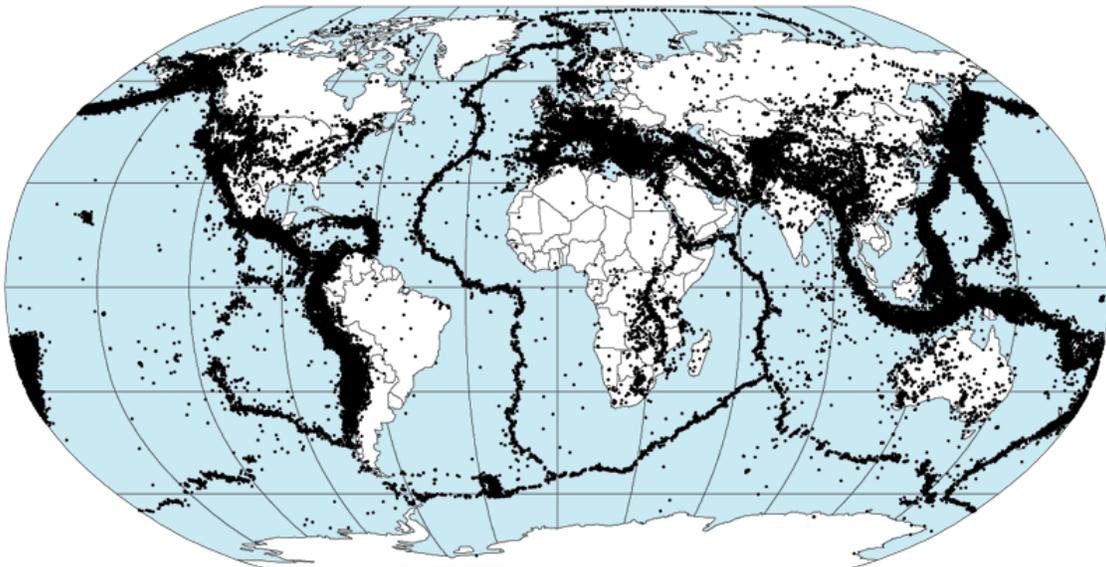
Vulcano island with the north coast of Sicily in the background.



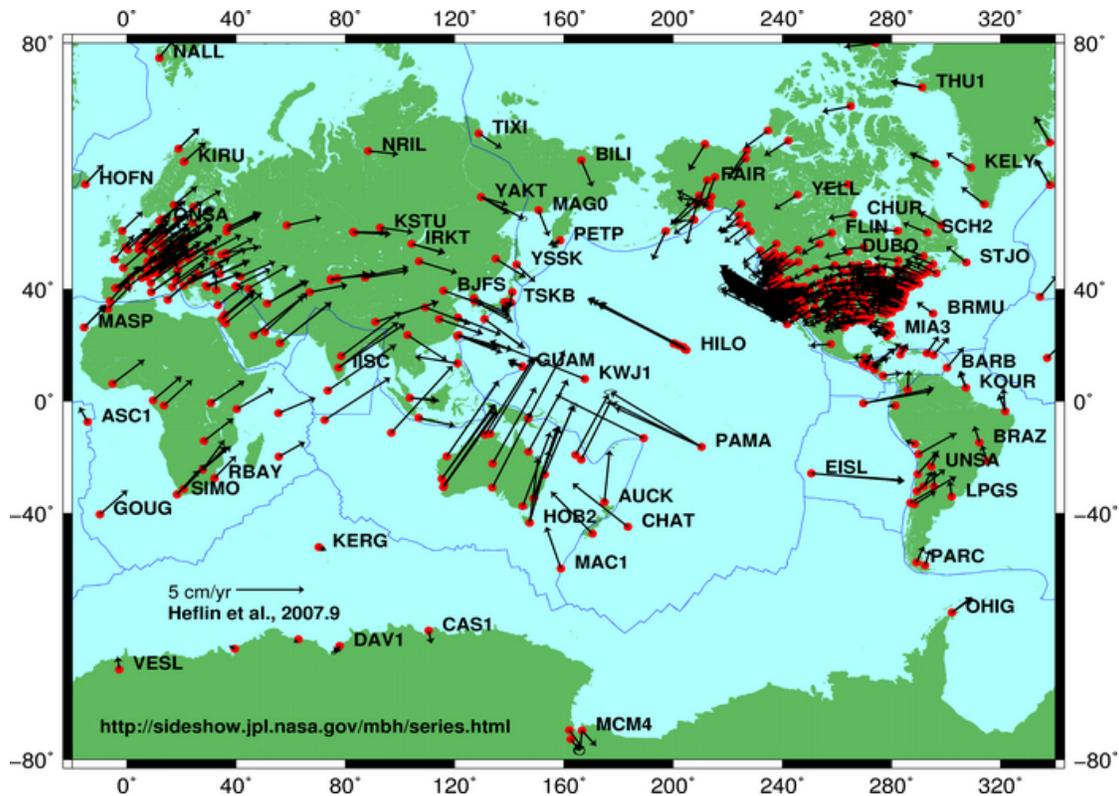
Remote Binubulauan in Kalinga province, central northern Luzon, Philippines, April 2009

## Earthquake

Preliminary Determination of Epicenters  
358,214 Events, 1963 - 1998



Global earthquake epicenters, 1963–1998



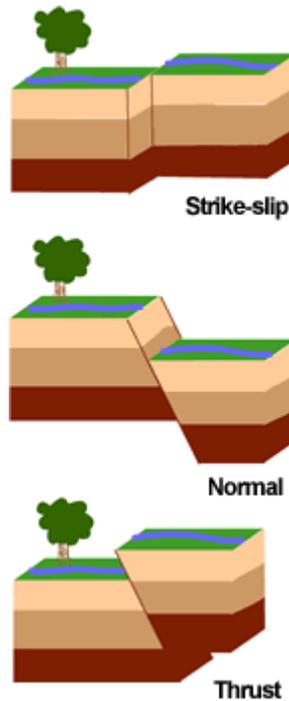
Global plate tectonic movement

An **earthquake** (also known as a **quake**, **tremor** or **temblor**) is the result of a sudden release of energy in the Earth's crust that creates seismic waves. The **seismicity** or **seismic activity** of an area refers to the frequency, type and size of earthquakes experienced over a period of time. Earthquakes are measured with a seismometer; a device which also records is known as a *seismograph*. The moment magnitude (or the related and mostly obsolete Richter magnitude) of an earthquake is conventionally reported, with magnitude 3 or lower earthquakes being mostly imperceptible and magnitude 7 causing serious damage over large areas. Intensity of shaking is measured on the modified Mercalli scale.

At the Earth's surface, earthquakes manifest themselves by shaking and sometimes displacing the ground. When a large earthquake epicenter is located offshore, the seabed sometimes suffers sufficient displacement to cause a tsunami. The shaking in earthquakes can also trigger landslides and occasionally volcanic activity.

In its most generic sense, the word *earthquake* is used to describe any seismic event—whether a natural phenomenon or an event caused by humans—that generates seismic waves. Earthquakes are caused mostly by rupture of geological faults, but also by volcanic activity, landslides, mine blasts, and nuclear tests. An earthquake's point of initial rupture is called its focus or hypocenter. The term epicenter refers to the point at ground level directly above the hypocenter.

## Naturally occurring earthquakes



Fault types

Tectonic earthquakes will occur anywhere within the earth where there is sufficient stored elastic strain energy to drive fracture propagation along a fault plane. In the case of transform or convergent type plate boundaries, which form the largest fault surfaces on earth, they will move past each other smoothly and aseismically only if there are no irregularities or asperities along the boundary that increase the frictional resistance. Most boundaries do have such asperities and this leads to a form of stick-slip behaviour. Once the boundary has locked, continued relative motion between the plates leads to increasing stress and therefore, stored strain energy in the volume around the fault surface. This continues until the stress has risen sufficiently to break through the asperity, suddenly allowing sliding over the locked portion of the fault, releasing the stored energy. This energy is released as a combination of radiated elastic strain seismic waves, frictional heating of the fault surface, and cracking of the rock, thus causing an earthquake. This process of gradual build-up of strain and stress punctuated by occasional sudden earthquake failure is referred to as the Elastic-rebound theory. It is estimated that only 10 percent or less of an earthquake's total energy is radiated as seismic energy. Most of the earthquake's energy is used to power the earthquake fracture growth or is converted into heat generated by friction. Therefore, earthquakes lower the Earth's available elastic potential energy and raise its temperature, though these changes are negligible compared to the conductive and convective flow of heat out from the Earth's deep interior.

## **Earthquake fault types**

There are three main types of fault that may cause an earthquake: normal, reverse (thrust) and strike-slip. Normal and reverse faulting are examples of dip-slip, where the displacement along the fault is in the direction of dip and movement on them involves a vertical component. Normal faults occur mainly in areas where the crust is being extended such as a divergent boundary. Reverse faults occur in areas where the crust is being shortened such as at a convergent boundary. Strike-slip faults are steep structures where the two sides of the fault slip horizontally past each other; transform boundaries are a particular type of strike-slip fault. Many earthquakes are caused by movement on faults that have components of both dip-slip and strike-slip; this is known as oblique slip.

## **Earthquakes away from plate boundaries**

Where plate boundaries occur within continental lithosphere, deformation is spread out over a much larger area than the plate boundary itself. In the case of the San Andreas fault continental transform, many earthquakes occur away from the plate boundary and are related to strains developed within the broader zone of deformation caused by major irregularities in the fault trace (e.g. the “Big bend” region). The Northridge earthquake was associated with movement on a blind thrust within such a zone. Another example is the strongly oblique convergent plate boundary between the Arabian and Eurasian plates where it runs through the northwestern part of the Zagros mountains. The deformation associated with this plate boundary is partitioned into nearly pure thrust sense movements perpendicular to the boundary over a wide zone to the southwest and nearly pure strike-slip motion along the Main Recent Fault close to the actual plate boundary itself. This is demonstrated by earthquake focal mechanisms.

All tectonic plates have internal stress fields caused by their interactions with neighbouring plates and sedimentary loading or unloading (e.g. deglaciation). These stresses may be sufficient to cause failure along existing fault planes, giving rise to intraplate earthquakes.

## **Shallow-focus and deep-focus earthquakes**

The majority of tectonic earthquakes originate at the ring of fire in depths not exceeding tens of kilometers. Earthquakes occurring at a depth of less than 70 km are classified as 'shallow-focus' earthquakes, while those with a focal-depth between 70 and 300 km are commonly termed 'mid-focus' or 'intermediate-depth' earthquakes. In subduction zones, where older and colder oceanic crust descends beneath another tectonic plate, deep-focus earthquakes may occur at much greater depths (ranging from 300 up to 700 kilometers). These seismically active areas of subduction are known as Wadati-Benioff zones. Deep-focus earthquakes occur at a depth at which the subducted lithosphere should no longer be brittle, due to the high temperature and pressure. A possible mechanism for the generation of deep-focus earthquakes is faulting caused by olivine undergoing a phase transition into a spinel structure.

## Earthquakes and volcanic activity

Earthquakes often occur in volcanic regions and are caused there, both by tectonic faults and the movement of magma in volcanoes. Such earthquakes can serve as an early warning of volcanic eruptions, as during the Mount St. Helens eruption of 1980.

Earthquake swarms can serve as markers for the location of the flowing magma throughout the volcanoes. These swarms can be recorded by seismometers and tiltmeters (a device which measures the ground slope) and used as sensors to predict imminent or upcoming eruptions.

## Rupture dynamics

A tectonic earthquake begins by an initial rupture at a point on the fault surface, a process known as nucleation. The scale of the nucleation zone is uncertain, with some evidence, such as the rupture dimensions of the smallest earthquakes, suggesting that it is smaller than 100 m while other evidence, such as a slow component revealed by low-frequency spectra of some earthquakes, suggest that it is larger. The possibility that the nucleation involves some sort of preparation process is supported by the observation that about 40% of earthquakes are preceded by foreshocks. Once the rupture has initiated it begins to propagate along the fault surface. The mechanics of this process are poorly understood, partly because it is difficult to recreate the high sliding velocities in a laboratory. Also the effects of strong ground motion make it very difficult to record information close to a nucleation zone.

Rupture propagation is generally modelled using a fracture mechanics approach, likening the rupture to a propagating mixed mode shear crack. The rupture velocity is a function of the fracture energy in the volume around the crack tip, increasing with decreasing fracture energy. The velocity of rupture propagation is orders of magnitude faster than the displacement velocity across the fault. Earthquake ruptures typically propagate at velocities that are in the range 70–90 % of the S-wave velocity and this is independent of earthquake size. A small subset of earthquake ruptures appear to have propagated at speeds greater than the S-wave velocity. These supershear earthquakes have all been observed during large strike-slip events. The unusually wide zone of coseismic damage caused by the 2001 Kunlun earthquake has been attributed to the effects of the sonic boom developed in such earthquakes. Some earthquake ruptures travel at unusually low velocities and are referred to as slow earthquakes. A particularly dangerous form of slow earthquake is the *tsunami earthquake*, observed where the relatively low felt intensities, caused by the slow propagation speed of some great earthquakes, fail to alert the population of the neighbouring coast, as in the 1896 Meiji-Sanriku earthquake.

## Earthquake clusters

Most earthquakes form part of a sequence, related to each other in terms of location and time. Most earthquake clusters consist of small tremors which cause little to no damage, but there is a theory that earthquakes can recur in a regular pattern.

## **Aftershocks**

An aftershock is an earthquake that occurs after a previous earthquake, the mainshock. An aftershock is in the same region of the main shock but always of a smaller magnitude. If an aftershock is larger than the main shock, the aftershock is redesignated as the main shock and the original main shock is redesignated as a foreshock. Aftershocks are formed as the crust around the displaced fault plane adjusts to the effects of the main shock.

## **Earthquake swarms**

Earthquake swarms are sequences of earthquakes striking in a specific area within a short period of time. They are different from earthquakes followed by a series of aftershocks by the fact that no single earthquake in the sequence is obviously the main shock, therefore none have notable higher magnitudes than the other. An example of an earthquake swarm is the 2004 activity at Yellowstone National Park.

## **Earthquake storms**

Sometimes a series of earthquakes occur in a sort of earthquake storm, where the earthquakes strike a fault in clusters, each triggered by the shaking or stress redistribution of the previous earthquakes. Similar to aftershocks but on adjacent segments of fault, these storms occur over the course of years, and with some of the later earthquakes as damaging as the early ones. Such a pattern was observed in the sequence of about a dozen earthquakes that struck the North Anatolian Fault in Turkey in the 20th century and has been inferred for older anomalous clusters of large earthquakes in the Middle East.

## **Size and frequency of occurrence**

There are around 500,000 earthquakes each year. About 100,000 of these can actually be felt. Minor earthquakes occur nearly constantly around the world in places like California and Alaska in the U.S., as well as in Guatemala, Chile, Peru, Indonesia, Iran, Pakistan, the Azores in Portugal, Turkey, New Zealand, Greece, Italy, and Japan, but earthquakes can occur almost anywhere, including New York City, London, and Australia. Larger earthquakes occur less frequently, the relationship being exponential; for example, roughly ten times as many earthquakes larger than magnitude 4 occur in a particular time period than earthquakes larger than magnitude 5. In the (low seismicity) United Kingdom, for example, it has been calculated that the average recurrences are: an earthquake of 3.7 - 4.6 every year, an earthquake of 4.7 - 5.5 every 10 years, and an earthquake of 5.6 or larger every 100 years. This is an example of the Gutenberg-Richter law.



The Messina earthquake and tsunami took as many as 200,000 lives on December 28, 1908 in Sicily and Calabria.

The number of seismic stations has increased from about 350 in 1931 to many thousands today. As a result, many more earthquakes are reported than in the past, but this is because of the vast improvement in instrumentation, rather than an increase in the number of earthquakes. The USGS estimates that, since 1900, there have been an average of 18 major earthquakes (magnitude 7.0-7.9) and one great earthquake (magnitude 8.0 or greater) per year, and that this average has been relatively stable. In recent years, the number of major earthquakes per year has decreased, although this is thought likely to be a statistical fluctuation rather than a systematic trend. More detailed statistics on the size and frequency of earthquakes is available from the USGS.

Most of the world's earthquakes (90%, and 81% of the largest) take place in the 40,000-km-long, horseshoe-shaped zone called the circum-Pacific seismic belt, known as the Pacific Ring of Fire, which for the most part bounds the Pacific Plate. Massive earthquakes tend to occur along other plate boundaries, too, such as along the Himalayan Mountains.

With the rapid growth of mega-cities such as Mexico City, Tokyo and Tehran, in areas of high seismic risk, some seismologists are warning that a single quake may claim the lives of up to 3 million people.

## **Induced seismicity**

While most earthquakes are caused by movement of the Earth's tectonic plates, human activity can also produce earthquakes. Four main activities contribute to this phenomenon: constructing large dams and buildings, drilling and injecting liquid into wells, and by coal mining and oil drilling. Perhaps the best known example is the 2008 Sichuan earthquake in China's Sichuan Province in May; this tremor resulted in 69,227 fatalities and is the 19th deadliest earthquake of all time. The Zipingpu Dam is believed to have fluctuated the pressure of the fault 1,650 feet (503 m) away; this pressure probably increased the power of the earthquake and accelerated the rate of movement for the fault. The greatest earthquake in Australia's history was also induced by humanity, through coal mining. The city of Newcastle was built over a large sector of coal mining

areas. The earthquake was spawned from a fault which reactivated due to the millions of tonnes of rock removed in the mining process.

## Measuring and locating earthquakes

Earthquakes can be recorded by seismometers up to great distances, because seismic waves travel through the whole Earth's interior. The absolute magnitude of a quake is conventionally reported by numbers on the Moment magnitude scale (formerly Richter scale, magnitude 7 causing serious damage over large areas), whereas the felt magnitude is reported using the modified Mercalli intensity scale (intensity II-XII).

Every tremor produces different types of seismic waves which travel through rock with different velocities: the longitudinal P-waves (shock- or pressure waves), the transverse S-waves (both body waves) and several surface waves (Rayleigh and Love waves). The propagation velocity of the seismic waves ranges from approx. 3 km/s up to 13 km/s, depending on the density and elasticity of the medium. In the Earth's interior the shock- or P waves travel much faster than the S waves (approx. relation 1.7 : 1). The differences in travel time from the epicentre to the observatory are a measure of the distance and can be used to image both sources of quakes and structures within the Earth. Also the depth of the hypocenter can be computed roughly.

In solid rock P-waves travel at about 6 to 7 km per second; the velocity increases within the deep mantle to ~13 km/s. The velocity of S-waves ranges from 2–3 km/s in light sediments and 4–5 km/s in the Earth's crust up to 7 km/s in the deep mantle. As a consequence, the first waves of a distant earth quake arrive at an observatory via the Earth's mantle.

**Rule of thumb:** On the average, the kilometer distance to the earthquake is the number of seconds between the P and S wave **times 8**. Slight deviations are caused by inhomogeneities of subsurface structure. By such analyses of seismograms the Earth's core was located in 1913 by Beno Gutenberg.

Earthquakes are not only categorized by their magnitude but also by the place where they occur. The world is divided into 754 Flinn-Engdahl regions (F-E regions), which are based on political and geographical boundaries as well as seismic activity. More active zones are divided into smaller F-E regions whereas less active zones belong to larger F-E regions.

## Effects/impacts of earthquakes



1755 copper engraving depicting Lisbon in ruins and in flames after the 1755 Lisbon earthquake, which killed an estimated 60,000 people. A tsunami overwhelms the ships in the harbor.

The effects of earthquakes include, but are not limited to, the following:

### **Shaking and ground rupture**

Shaking and ground rupture are the main effects created by earthquakes, principally resulting in more or less severe damage to buildings and other rigid structures. The severity of the local effects depends on the complex combination of the earthquake magnitude, the distance from the epicenter, and the local geological and geomorphological conditions, which may amplify or reduce wave propagation. The ground-shaking is measured by ground acceleration.

Specific local geological, geomorphological, and geostructural features can induce high levels of shaking on the ground surface even from low-intensity earthquakes. This effect is called site or local amplification. It is principally due to the transfer of the seismic motion from hard deep soils to soft superficial soils and to effects of seismic energy focalization owing to typical geometrical setting of the deposits.

Ground rupture is a visible breaking and displacement of the Earth's surface along the trace of the fault, which may be of the order of several metres in the case of major earthquakes. Ground rupture is a major risk for large engineering structures such as dams,

bridges and nuclear power stations and requires careful mapping of existing faults to identify any likely to break the ground surface within the life of the structure.

## **Landslides and avalanches**

Earthquakes, along with severe storms, volcanic activity, coastal wave attack, and wildfires, can produce slope instability leading to landslides, a major geological hazard. Landslide danger may persist while emergency personnel are attempting rescue.

## **Fires**



Fires of the 1906 San Francisco earthquake

Earthquakes can cause fires by damaging electrical power or gas lines. In the event of water mains rupturing and a loss of pressure, it may also become difficult to stop the spread of a fire once it has started. For example, more deaths in the 1906 San Francisco earthquake were caused by fire than by the earthquake itself.

## **Soil liquefaction**

Soil liquefaction occurs when, because of the shaking, water-saturated granular material (such as sand) temporarily loses its strength and transforms from a solid to a liquid. Soil liquefaction may cause rigid structures, like buildings and bridges, to tilt or sink into the

liquefied deposits. This can be a devastating effect of earthquakes. For example, in the 1964 Alaska earthquake, soil liquefaction caused many buildings to sink into the ground, eventually collapsing upon themselves.

## Tsunami



The tsunami of the 2004 Indian Ocean earthquake

Tsunamis are long-wavelength, long-period sea waves produced by the sudden or abrupt movement of large volumes of water. In the open ocean the distance between wave crests can surpass 100 kilometers (62 miles), and the wave periods can vary from five minutes to one hour. Such tsunamis travel 600-800 kilometers per hour (373–497 miles per hour), depending on water depth. Large waves produced by an earthquake or a submarine landslide can overrun nearby coastal areas in a matter of minutes. Tsunamis can also travel thousands of kilometers across open ocean and wreak destruction on far shores hours after the earthquake that generated them.

Ordinarily, subduction earthquakes under magnitude 7.5 on the Richter scale do not cause tsunamis, although some instances of this have been recorded. Most destructive tsunamis are caused by earthquakes of magnitude 7.5 or more.

## **Floods**

A flood is an overflow of any amount of water that reaches land. Floods occur usually when the volume of water within a body of water, such as a river or lake, exceeds the total capacity of the formation, and as a result some of the water flows or sits outside of the normal perimeter of the body. However, floods may be secondary effects of earthquakes, if dams are damaged. Earthquakes may cause landslips to dam rivers, which then collapse and cause floods.

The terrain below the Sarez Lake in Tajikistan is in danger of catastrophic flood if the landslide dam formed by the earthquake, known as the Usoi Dam, were to fail during a future earthquake. Impact projections suggest the flood could affect roughly 5 million people.

## **Tidal forces**

Research work has shown a robust correlation between small tidally induced forces and non-volcanic tremor activity.

## Human impacts



Damaged infrastructure, one week after the 2007 Peru earthquake

Earthquakes may lead to disease, lack of basic necessities, loss of life, higher insurance premiums, general property damage, road and bridge damage, and collapse or destabilization (potentially leading to future collapse) of buildings. Earthquakes can also precede volcanic eruptions, which cause further problems; for example, substantial crop damage, as in the "Year Without a Summer" (1816).

## Major earthquakes

The largest earthquake that has been measured on a seismograph reached 9.5 magnitude, occurring on 22 May 1960. Its epicenter was near Cañete, Chile. The energy released was approximately twice that of the next most powerful earthquake, the Good Friday Earthquake, which was centered in Prince William Sound, Alaska. The ten largest recorded earthquakes have all been megathrust earthquakes; however, of these ten, only the 2004 Indian Ocean earthquake is simultaneously one of the deadliest earthquakes in history.

The earthquakes with the greatest amount of loss of life, while powerful, were deadly because of their proximity to either heavily populated areas or the ocean, where earthquakes can potentially create tsunamis which can devastate communities thousands of miles away. Regions that are most at risk for great loss of life include those where earthquakes are relatively rare but powerful, and poor regions with lax, unenforced, or nonexistent seismic building codes.

## Preparation

In order to determine the likelihood of future seismic activity, geologists and other scientists examine the rock of an area to determine if the rock appears to be "strained". Studying the faults of an area to study the buildup time it takes for the fault to build up stress sufficient for an earthquake also serves as an effective prediction technique. Measurements of the amount of accumulated strain energy on the fault each year, time passed since the last major temblor, and the energy and power of the last earthquake are made. Together the facts allow scientists to determine how much pressure it takes for the fault to generate an earthquake. Though this method is useful, it has only been implemented on California's San Andreas Fault.

Today, there are ways to protect and prepare possible sites of earthquakes from severe damage, through the following processes: earthquake engineering, earthquake preparedness, household seismic safety, seismic retrofit (including special fasteners, materials, and techniques), seismic hazard, mitigation of seismic motion, and earthquake prediction. Seismic retrofitting is the modification of existing structures to make them more resistant to seismic activity, ground motion, or soil failure due to earthquakes. With better understanding of seismic demand on structures and with our recent experiences with large earthquakes near urban centers, the need of seismic retrofitting is well acknowledged. Prior to the introduction of modern seismic codes in the late 1960s for developed countries (US, Japan etc.) and late 1970s for many other parts of the world (Turkey, China etc.), many structures were designed without adequate detailing and reinforcement for seismic protection. In view of the imminent problem, various research work has been carried out. Furthermore, state-of-the-art technical guidelines for seismic assessment, retrofit and rehabilitation have been published around the world - such as the ASCE-SEI 41 and the New Zealand Society for Earthquake Engineering (NZSEE)'s guidelines.