

Introduction to Climate and Environmental Physics



Proff. Thomas Stocker, Hubertus Fischer,
Fortunat Joos, Markus Leuenberger,
Thomas Frölicher and Christoph Raible

Lecture Notes, Autumn Semester 2022, University of Bern

Front page: The Blue Marble. Photograph taken by the crew of Apollo 17 on December 7, 1972 on their way to the moon (NASA AS17-148-22727).

"I believe it is in our nature to explore, to reach out into the unknown. The only true failure would be not to explore at all."

(Sir Ernest Shackleton)

Contents

1	The atmosphere	5
1.1	Global radiation balance	5
1.2	Global warming	12
1.3	Aerosols in the atmosphere	15
1.4	Atmospheric structure and circulation	25
1.5	Statistical description – modes of variability	34
1.6	A selected mechanism: Delayed Oscillator	40
2	Stable isotopes in climate and environmental physics	45
2.1	Introduction	45
2.2	Fractionation of isotopes	45
2.2.1	Non-equilibrium fractionation	46
2.2.2	Equilibrium fractionation	47
2.2.3	Definition of the fractionation factor	48
2.2.4	Delta-notation	49
2.3	Stable isotope in the water cycle	50
2.3.1	Comparison with observations	53
2.3.2	Relationship between $\delta^{18}\text{O}$ and $\delta^2\text{H}$	54
3	Radio nuclides as tracers of environmental processes	57
3.1	Introduction and overview	57
3.1.1	Overview of reservoirs and processes	58
3.2	Production of radio nuclides	59
3.2.1	Production in the atmosphere	59
3.3	Subsurface production	64
3.3.1	Secondary particles in the soil	64
3.3.2	Radon and its daughter products	64
3.3.3	Uranium and thorium	64
3.3.4	Bomb induced isotopes	65
3.4	Concentration of radioactive nuclides in the atmosphere	66
3.4.1	Range of activity of radionuclides	66
3.4.2	Temporal variation of ^{90}Sr due to bomb tests	66
3.4.3	^{85}Kr emitted from nuclear reprocessing systems	68
3.4.4	Temporal variation of isotope abundances	68

3.5	Transfer functions of radio nuclides	71
3.5.1	Dry and wet deposition	71
3.5.2	Penetration of nuclides into the soil, or sedimentation	73
3.5.3	Transfer functions	73
3.6	Radio nuclides for dating and process studies	74
3.6.1	Radioactive decay	76
3.6.2	¹⁴ C method	77
4	Biogeochemical Cycles in the Climate System	83
4.1	Introduction and Motivation	83
4.1.1	A brief historical perspective	83
4.1.2	The key role of biogeochemical cycles	86
4.2	Greenhouse gas increase and climate change	86
4.3	The main carbon reservoirs	87
5	Glacial-interglacial atmospheric CO₂ Variations	91
5.1	Milankovitch cycles of climate over the past 800,000 years	91
5.2	Mechanisms of natural CO ₂ variations	92
5.2.1	Temperature	94
5.2.2	Export of organic material	94
5.2.3	Export of CaCO ₃	97
6	The uptake of anthropogenic carbon by the ocean	99
6.1	Dissolved gases in the ocean	99
6.1.1	Solubility of gases	99
6.1.2	Bunsen's solubility coefficient α and Henry's law	99
6.1.3	Relationship between α and s	100
6.2	Uptake of anthropogenic CO ₂ by the Ocean	101
6.2.1	CO ₂ exchange: Atmosphere – ocean	102
6.2.2	Carbonate chemistry	102
6.2.3	CO ₂ uptake by the ocean on long timescales	104
6.2.4	Tracers reveal how fast the ocean is ventilated	105
6.2.5	Oceanic CO ₂ uptake: A simple model	106
7	Oxygen in the climate system and links between O₂ and CO₂	109
7.1	Observed links between atmospheric O ₂ and CO ₂	109
7.2	Overturning time of atmospheric O ₂ in the biosphere	110
7.3	How much O ₂ can be consumed by fossil fuel burning?	110
7.4	How fast can the atmospheric O ₂ content vary naturally?	111
7.5	Origin of free oxygen	111
7.5.1	Development of atmospheric oxygen	111
7.6	The anthropogenic CO ₂ perturbation and O ₂ changes	112

8 Ocean: Vertical structure, water masses and sea level rise	119
8.1 Inventories and fluxes	119
8.2 Density of ocean water	120
8.3 Vertical structure and water masses	122
8.4 Sea level rise	127
9 Large-scale ocean circulation	131
9.1 Simplified coordinate system on the Earth and Coriolis acceleration	131
9.2 Hydrostatic balance	132
9.3 Pressure gradient geostrophic balance	133
9.4 Ekman circulation at the surface	134
9.5 Sverdrup balance	136
9.6 Western boundary currents	139
9.7 Deep circulation in the ocean	140
9.8 Synthesis view of the basin-scale circulation	142
10 Multiple equilibria in the climate system	145
10.1 Abrupt climate transitions recorded by polar ice cores	145
10.2 The Bipolar Climate Seesaw	145
10.3 Multiple equilibria in a simple ocean model	147
10.4 Multiple equilibria in coupled climate models	150
Table of constants	155
References	157

Chapter 1

The atmosphere

1.1 Global radiation balance

The energy driving any climatic processes on the Earth is derived from the Sun (apart of the very small geothermal heat flux that can be neglected here). As all bodies at a certain temperature emit radiation in a certain frequency range, so does the Sun. This process is called black body radiation and the spectral energy density of the emitted frequency spectrum is given according to Planck's law and illustrated in Fig. 1.1.

$$u(\lambda, T) = \frac{8\pi hc}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{k\lambda T}} - 1}, \quad (1.1)$$

where λ is the wavelength, T the temperature, $h = 6.626 \cdot 10^{-34}$ Js is the Planck constant, $c = 2.998 \cdot 10^8$ m/s is the speed of light and $k = R/N_A = 1.38 \cdot 10^{-23}$ J/K the Boltzmann constant, which is the gas constant $R = 8.314$ J/molK divided by the number of molecules in one mol of particles $N_A = 6.022 \cdot 10^{23}$ mol⁻¹ (Avogadro number). As illustrated in Fig. 1.1 this energy spectrum has a clear peak, which is dependent on the temperature of the black body, and has a strong tail at longer wavelengths.

The total energy flux H emitted by such a body (integrated over the entire spectrum) is given by the Stefan-Boltzmann law

$$H = \sigma T^4, \quad (1.2)$$

where $\sigma = 5.67 \cdot 10^{-8}$ W/(m²K⁴) is the Stefan Boltzmann number. The emission temperature of the Sun is about 5780 K, accordingly the Sun emits about $63.3 \cdot 10^6$ W/m² into space equally in all directions. According to Planck's law, this temperature leads to a frequency spectrum in the visible electromagnetic range with a maximum at a wavelength of about 500 nm. At the distance of the Earth to the Sun (about $150 \cdot 10^9$ m) the radiation density is "diluted" to only $S_0 = 1361$ W/m². The value S_0 is called the solar constant, although it is not absolutely constant and changes by about 1–3 W/m² on seasonal, decadal to centennial time scales due to changes in the Sun's activity or due to the eccentricity of the Earth's orbit around the Sun. The Earth absorbs $S_0(1 - \alpha)\pi R_E^2$ of this radiation, where $\alpha \approx 0.3$ is the planetary albedo and R_E is the Earth radius of about $6.4 \cdot 10^6$ m.

The Earth absorbs this energy and heats up until its own black body radiation emission equals the absorbed solar energy. In radiative equilibrium this implies that

$$E_{\text{in}} = S_0(1 - \alpha)\pi R_E^2 = 4\pi R_E^2\sigma T_E^4 = E_{\text{out}}. \quad (1.3)$$

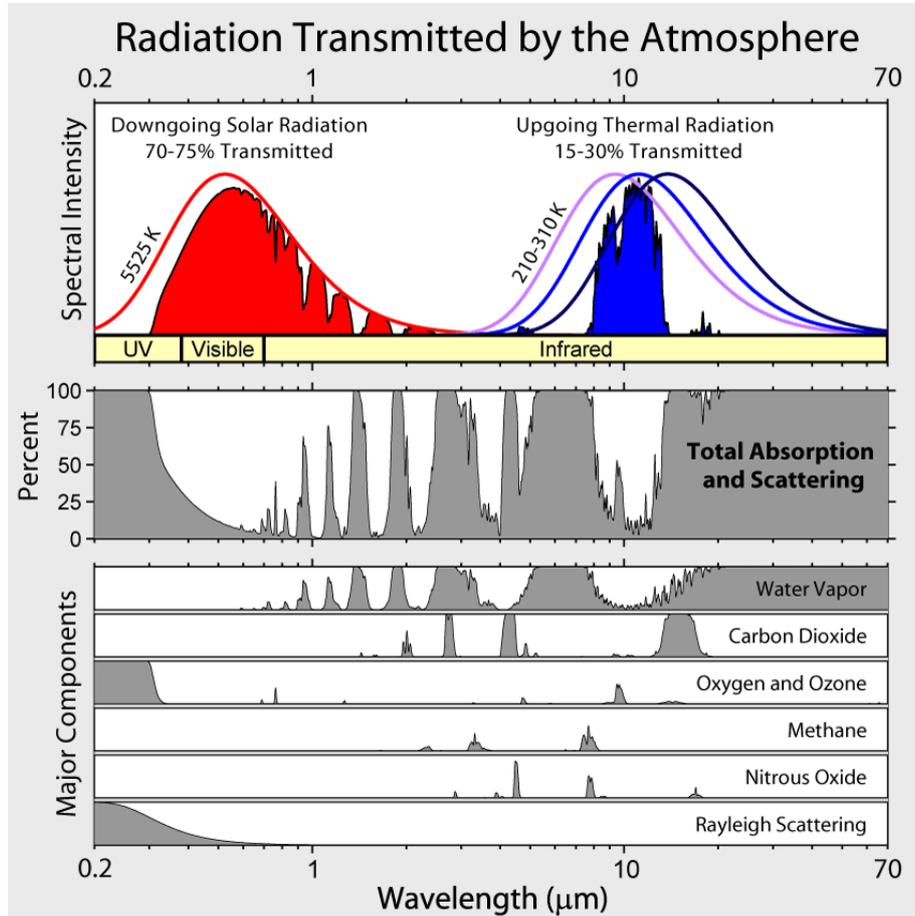


Figure 1.1: Theoretical black body spectra of incoming and outgoing radiation on the Earth together with the real transmitted spectra due to scattering and absorption.

The respective black body temperature of the Earth (also called emission temperature at the top of the atmosphere T_E) is about $255 \text{ K} = -18 \text{ }^\circ\text{C}$ and leads to a spectrum in the infrared with a maximum at a wavelength of about 11360 nm . This emission temperature is significantly colder than the mean global surface temperature. The reason for this is that the atmosphere absorbs part of the emitted terrestrial black body radiation and re-emits it back to the Earth. This is the natural greenhouse effect, which is dependent on the concentration of greenhouse active gases in the atmosphere. Thus, part of the thermal energy is trapped in the atmosphere and only the energy defined by the emission temperature is lost to space at the top of the atmosphere by longwave radiation to achieve a radiative equilibrium.

On a planet with an atmosphere such as the Earth we obtain not only the remaining transmitted shortwave radiation SW_{surf} at the surface but also the downward flux of longwave radiation emitted by the atmosphere. Then (a simplified) radiative equilibrium at the surface requires

$$4\pi R_E^2 \sigma T_S^4 = 4\pi R_E^2 SW_{\text{surf}} + 4\pi R_E^2 \sigma T_a^4, \quad (1.4)$$

with T_a the temperature of the atmosphere (assumed to be the same everywhere). If we assume, that the atmosphere has the same temperature as the emission temperature of the Earth at the top of the atmosphere, this equation leads to

$$T_S = \sqrt[4]{\frac{SW_{\text{surf}}}{\sigma} + T_E^4}. \quad (1.5)$$

With the emission temperature being always 255 K and taking a value of $SW_{\text{surf}} = 161 \text{ W/m}^2$ (Fig.1.2) this leads to a surface temperature of 290 K, much closer to the observed global mean annual temperature of the Earth surface of about 288 K. Clearly our model is still an oversimplification, since the atmosphere has not a uniform temperature distribution.

In Fig. 1.1 the black body spectra of the Sun and the Earth are plotted together with the real transmitted radiation. As explained before, the incoming solar radiation covers mainly the visible and UV spectrum, while the terrestrial radiation emits in the infrared. Luckily for radiation balance studies both spectra are nearly disjunct, i.e., there is little overlap of the incoming short wave and the outgoing long wave radiation. Note also in Fig. 1.1 that trace gases in the atmosphere filter out parts of both the incoming solar as well as the outgoing terrestrial radiation by absorbing specific wave lengths. For example oxygen and ozone prevent harmful, incoming UV radiation from reaching the surface, protecting life at the surface. Vice versa greenhouse gases (and here especially water vapor) absorb outgoing radiation in the IR, heating the surface to temperatures above freezing point, making the presence of liquid water and, thus, also life possible.

Fig. 1.2 illustrates the globally averaged radiation balance of the Earth. In total, absorbed incoming and net outgoing radiation at the top of the atmosphere have to be equal to retain a constant temperature of our planet. Increasing concentrations of atmospheric greenhouse gases disturb this equilibrium and temporarily lead to a global warming based on physical laws until a new equilibrium is achieved. Apart from the incoming shortwave and the outgoing long wave radiation there also exist sensible and latent heat fluxes (Fig. 1.2). The sensible heat flux derives from heat exchange of air at the surface with convective transport of this warm air. The latent heat flux comprises the evaporation and condensation of water vapour at the surface and in the atmosphere. Note that the evaporation / condensation of 1 g water vapour at 25 °C requires/releases 2444 J. Globally, latent and sensible heat fluxes are relatively small compared to the radiative fluxes but they are quite substantial on a local scale.

As discussed above, the Earth receives $S_0/4 = 340 \text{ W/m}^2$ at the top of the atmosphere on a global average and, taking the planetary albedo into account, absorbs about 240 W/m^2 . The rest is reflected directly back to space.

Although this global mean already gives us a first order estimate of the mean global temperature, it cannot explain the diurnal, seasonal and geographical variations in temperature. Obviously, regions close to the equator receive more short wave radiation compared to higher latitudes when averaged over one year than the poles, leading to a geographical gradient in the radiation budget. Due to the inclination of the Earth axis relative to the planetary plane, however, the polar regions receive more incoming radiation in the respective summer season.

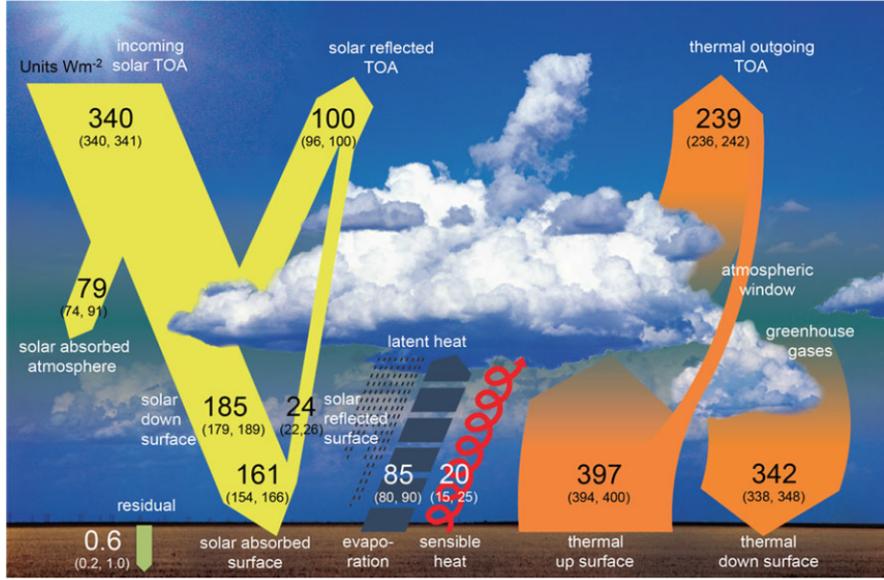


Figure 1.2: Radiative balance of the globe. All energy fluxes are given in W/m^2 [Wild *et al.*, 2013].

This diminishes the geographic gradient. In fact due to the long daytime hours in polar regions the maximum in daily mean energy flux of more than $500 \text{ W}/\text{m}^2$ is encountered at the poles in the respective summer (Fig. 1.3). Vice versa, the highest incoming radiative flux encountered in the course of one day are found in the tropics when the Sun stands vertically above the observer (Fig. 1.3).

Thus, the energy received from the Sun changes with latitude ϕ , the season as well as time of the day. The daily cycle of incoming shortwave radiation is

$$F(\phi, d, t) = S_0 \left(\frac{r_0}{r} \right)^2 \cos \gamma, \quad (1.6)$$

with r being the distance of the Sun during a specific day d of the year, which changes according to Kepler's law according to

$$r = \frac{a(1 - \epsilon^2)}{1 + \epsilon \cos \nu}. \quad (1.7)$$

$\epsilon = 0.0167$ is the current eccentricity of the Earth orbit, $a = 149.6 \cdot 10^9 \text{ m}$ is the semimajor axis of the orbit and $\nu = 360^\circ/365 \cdot d$ the angle passed since the perihelion of the orbit. Note that we assume that the Earth passes the same angle per day ignoring the effect of Kepler's third law (which is a reasonable simplification for the nearly circular orbit of the Earth) and of the obliquity, i.e. ignoring the effect known as "equation of time" in astronomy.

r_0 is the mean distance to the Sun ($149.6 \cdot 10^9 \text{ m}$, nearly identical to the semimajor axis for recent eccentricity conditions) and γ the zenith angle of the incoming radiation, which depends on the time t of the day, the day d in the year and the latitude ϕ of the point on the Earth according to

$$\cos \gamma = \sin \phi \sin \delta + \cos \phi \cos \delta \cos h, \quad (1.8)$$

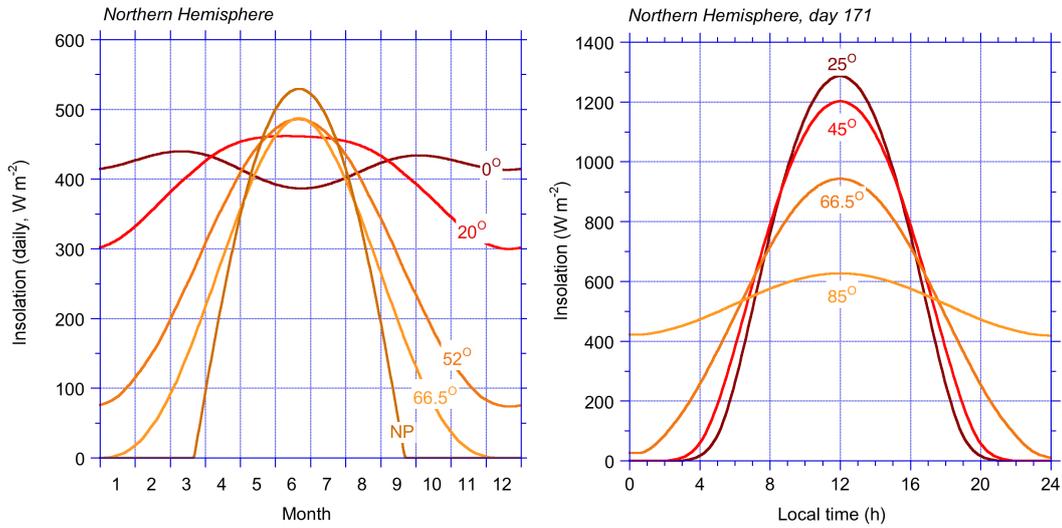


Figure 1.3: Daily cycle of the incoming shortwave radiation at the top of the atmosphere as a function of latitude for day 171 (right) and the seasonal cycle of the incoming shortwave radiation (left).

where $h(t)$ is the solar hour angle $h = 15^\circ(t - 12)$, with t the local time of the day in hours since noon. Obviously, each point on the globe receives radiation only between sunrise and sunset. We can determine the respective hour angle of the sunrise or sunset by setting $\cos \gamma$ in equation 1.8 to zero. This leads to a sunrise/sunset hour angle of

$$H = \pm \arccos(-\tan \phi \tan \delta) \quad (1.9)$$

Note, however, that this equation has only a solution for those latitudes and declinations, where a sunrise or a sunset occurs within a day. In the regions where a polar day occurs (sunlight for all 24 hours of the day), the sunrise angle would be (in radians) $-\pi$ (-180°) and the sunset angle $+\pi$ ($+180^\circ$).

The last component needed for the calculation of Eq. 1.6 is the solar declination δ , i.e., the angle between the equatorial plane of the Earth and the straight line connecting the Earth and the Sun for each day of the year. Because of the tilt of the Earth axis and its rotation around the Sun this parameter changes between -23.45° and $+23.45^\circ$ during the course of the year according to

$$\delta \approx 23.45^\circ \cdot \sin \beta, \quad (1.10)$$

with $\beta = 360^\circ/365 \cdot (d' - 81)$ and d' the day since the vernal equinox. The respective daily cycle for day of the year $d = 171$ is displayed in Fig. 1.3 on the right side for a few latitudes.

To derive the daily mean (Fig. 1.3 on the left side) of the incoming solar radiation equation Eq. 1.6 has to be averaged over $h(t)$, where we assume that the change in r and δ is small during one day (note that H is in radians here)

$$\bar{F} = S_0 \left(\frac{r_0}{r}\right)^2 \int_{-H}^H \frac{1}{2\pi} \cos \gamma dh \quad (1.11)$$

$$\bar{F} = \frac{S_0}{\pi} \left(\frac{r_0}{r}\right)^2 (H \sin \phi \sin \delta + \sin H \cos \phi \cos \delta). \quad (1.12)$$

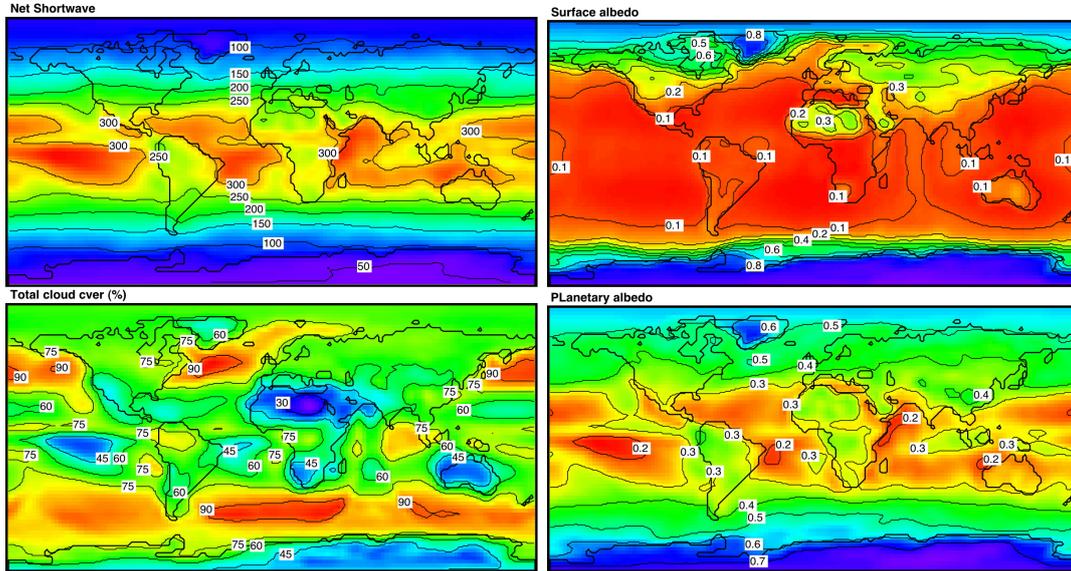


Figure 1.4: Geographical distribution of the net shortwave radiation at the top of the atmosphere (upper left, in W/m^2 with positive values referring to energy absorbed by the Earth/atmosphere system), the cloud cover (lower left), the surface albedo (upper right) and the planetary albedo as observed from space (lower right) (provided by van den Broeke, pers.comm).

In the following we will first discuss in detail the geographical variation of the annual mean radiative budget at the top of the atmosphere (TOA). When looking at Fig. 1.4, regional variations in the net TOA shortwave radiation (Fig. 1.3) $SW_{\text{net}} = SW_{\text{in}} - SW_{\text{out}} = SW_{\text{in}}(1 - \alpha_p)$ can be observed, which deviate significantly from the pure geometrically increase of the incoming radiation from the pole to the equator. The reason for this geographical variation is the fact that clouds reflect part of the incoming shortwave radiation and that different surfaces reflect the incoming shortwave radiation with different efficiency. E.g. the albedo of the ocean is about 0.1 (i.e. about 90% of the radiation are absorbed, while dry snow zones on the polar ice sheets have an albedo > 0.8 (i.e. less than 20% of the incoming shortwave radiation are absorbed).

The second term of the radiation balance (observed from out of space at the top of the atmosphere) is the annual mean net outgoing longwave (LW_{out}) radiation, i.e. the longwave radiation transmitted through or emitted from the atmosphere as displayed in Fig. 1.5. Averaged over the entire globe the net annual mean outgoing longwave radiation must equal the net annual mean incoming shortwave radiation, if the globe is in radiative equilibrium. Looking at Fig. 1.5, this is globally the case but the geographical distributions of incoming and outgoing radiation are different. Accordingly, the total net radiation $R_{\text{net}} = SW_{\text{net}} + LW_{\text{out}}$ at the top of the atmosphere is positive in the tropics (energy is absorbed) while it is negative over the poles (energy is lost to space). In essence, there is a surplus of incoming shortwave radiation at the equator and an excess of outgoing longwave radiation at the poles. This geographic imbalance is the reason for our latitudinal temperature gradients and, thus, our

climate zones. Due to this imbalance, the equator to pole energy/temperature gradient drives also the global atmosphere and ocean circulation, which try to equilibrate this gradient. If those energy fluxes by atmosphere/ocean circulation were not in place, a local energy balance between incoming and outgoing radiation would establish and the mean annual temperature gradient between pole and equator would be much more pronounced (think about planets without an atmosphere/ocean).

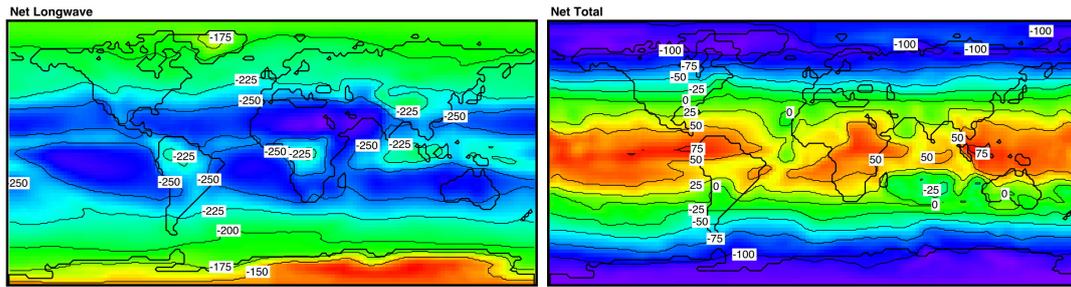


Figure 1.5: Geographical distribution of the net longwave radiation lost to space (left) and the total radiative balance at the top of the atmosphere (right). All values are given in W/m^2 with downward net radiation being positive (provided by van den Broeke, pers.comm).

In the following, we will have a look at the energy balance at the surface of our planet. Overall, less energy reaches the surface due to scattering of incoming shortwave radiation in the atmosphere. However, the geographical distribution of the incoming net shortwave radiation at the surface should be similar to the one at the top of the atmosphere. This is not the case for the net longwave radiation budget at the surface. Areas with little cloud cover emit longwave radiation more efficiently out to space, while clouds block outgoing radiation (Fig. 1.6).

Despite this regional difference in the longwave budget at the surface, the net radiation balance $R_{\text{net}} = SW_{\text{net}} + LW_{\text{net}}$ at the surface looks similar to the one at the top of the atmosphere. Note however, that the global average of the net radiation balance at the surface does not equal zero. To close the regional budget in the atmosphere sensible and latent heat fluxes are required (Fig. 1.7). Although their global average is relatively small compared to the incoming shortwave radiation, they can reach fluxes of 100–200 W/m^2 in certain areas and seasons.

As displayed in Fig. 1.7, the sensible heat flux is generally directed upward in low latitudes while it is directed downward in high latitudes during the winter season. The latter reflects the pronounced temperature inversion over snow covered regions that leads to a downwelling of heat from the atmosphere to the surface. In contrast in low latitudes, the surface is heated up by incoming radiation and transfers this heat to the atmosphere by conduction and convection.

Clearly, the latent heat flux is directed upward at the surface and is highest where incoming radiation is abundant to allow for evaporation of water (a downward latent heat flux is only possible in the case of surface hoar, where water is condensing at the surface). Accordingly, latent heat fluxes are the highest over the subtropical oceans and lowest over cold continental

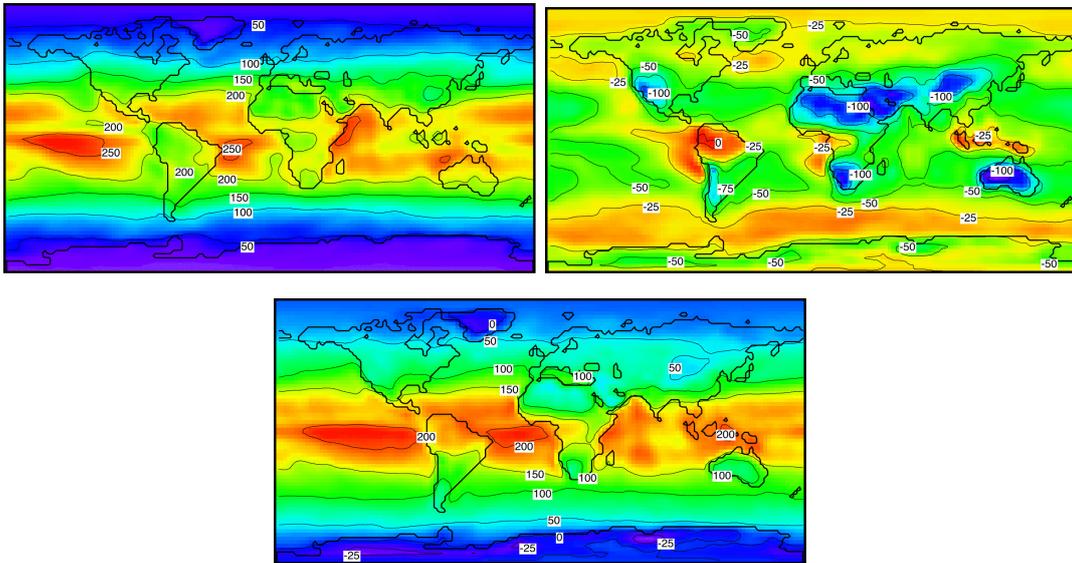


Figure 1.6: Geographical distribution of the net shortwave radiation (upper left) and the net longwave radiation (upper right) at the surface together with the net radiation balance at the surface (lower middle). All values are given in W/m^2 with downward net radiation being positive (provided by van den Broeke, pers.comm).

and polar regions.

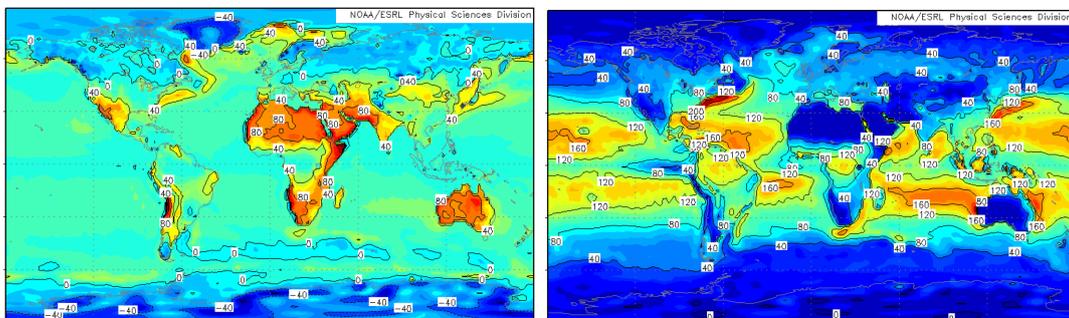


Figure 1.7: Geographical variation of the mean annual sensible heat flux at the surface (left) and the mean annual surface latent heat flux (right). All values given in W/m^2 where downward energy fluxes are negative in contrast to the figures above!

1.2 Global warming

Based on the different components of the global radiation balance discussed above, temporal and spatial changes in the global radiation budget can be caused by

- greenhouse gas concentrations in the atmosphere (trapping terrestrial outgoing longwave radiation) including the fast feedback of a warming atmosphere on its water vapour content,

- changes in the surface albedo (land and sea ice cover, vegetation) and planetary albedo (surface albedo and clouds),
- aerosol concentrations in the atmosphere (scattering and absorption of radiation as well as changes in cloud coverage),
- changes in solar insolation either by changes in the Sun's activity (changing the emission of energy at the Sun's surface on decadal and centennial time scales) or through changes in the orbital parameters of the Earth (changing mainly the seasonal and geographical distribution of the energy reaching the Earth on time scales of tenths of thousands of years).

The most important global warming factor today is undoubtedly the increase in emissions of greenhouse gases due to anthropogenic activity (see also the chapter on biogeochemical cycles). This increase has been documented by atmospheric monitoring over the last 2-5 decades but has been extended back in time by the use of polar ice cores representing an archive of greenhouse gas concentrations before the industrialization set in (Fig. 1.8). This shows that the current mean annual CO₂ (> 410 ppm), CH₄ (> 1870 ppb) and N₂O (> 330 ppb) levels were never reached during the last 800,000 years. In fact the atmospheric CH₄ level has more than doubled while CO₂ increased by more than 40% since the start of the industrialization. The influence of increasing greenhouse gas concentrations on the radiative balance of the Earth can be quantified in W/m² according to

$$\begin{aligned}
 \Delta Q_{\text{CO}_2} &= 5.35 \frac{\text{W}}{\text{m}^2} \ln \left(\frac{\text{CO}_2}{280} \right) \\
 \Delta Q_{\text{CH}_4} &= 0.036 \frac{\text{W}}{\text{m}^2} \left(\sqrt{\text{CH}_4} - \sqrt{700} \right) \\
 \Delta Q_{\text{N}_2\text{O}} &= 0.12 \frac{\text{W}}{\text{m}^2} \left(\sqrt{\text{N}_2\text{O}} - \sqrt{270} \right),
 \end{aligned} \tag{1.13}$$

where the CO₂ concentrations are given in ppm and the CH₄ and N₂O concentrations in ppb [IPCC, 2013]. The specified concentrations in the equations above refer to the preindustrial atmospheric concentrations of these three greenhouse gases as measured in polar ice cores. For the recent atmospheric CO₂ concentration this lead to a radiative forcing by 2011 of about +1.7 W/m² by CO₂ alone as also shown in Fig. 1.10, which continues to increase due to the slow implementation of CO₂ reduction measures worldwide

Apart from the increase in greenhouse gas concentration, anthropogenic activities lead to the emission of atmospheric aerosol species, which either absorb radiation (soot) or lead to an increased backscatter of radiation (e.g. sulphate aerosol). In addition, aerosols act as efficient cloud condensation nuclei with increased cloud coverage changing the planetary albedo. In contrast to greenhouse gases (with atmospheric lifetimes of many years), the atmospheric residence time of aerosol particles is on the order of days to weeks. Accordingly, their influence on the radiative balance is highly localized and most pronounced close to their sources. Whether these effects lead to a local cooling or warming is highly dependent on the aerosol species but on a global average anthropogenic aerosol particles are believed to partly counteract the warming by greenhouse gases.

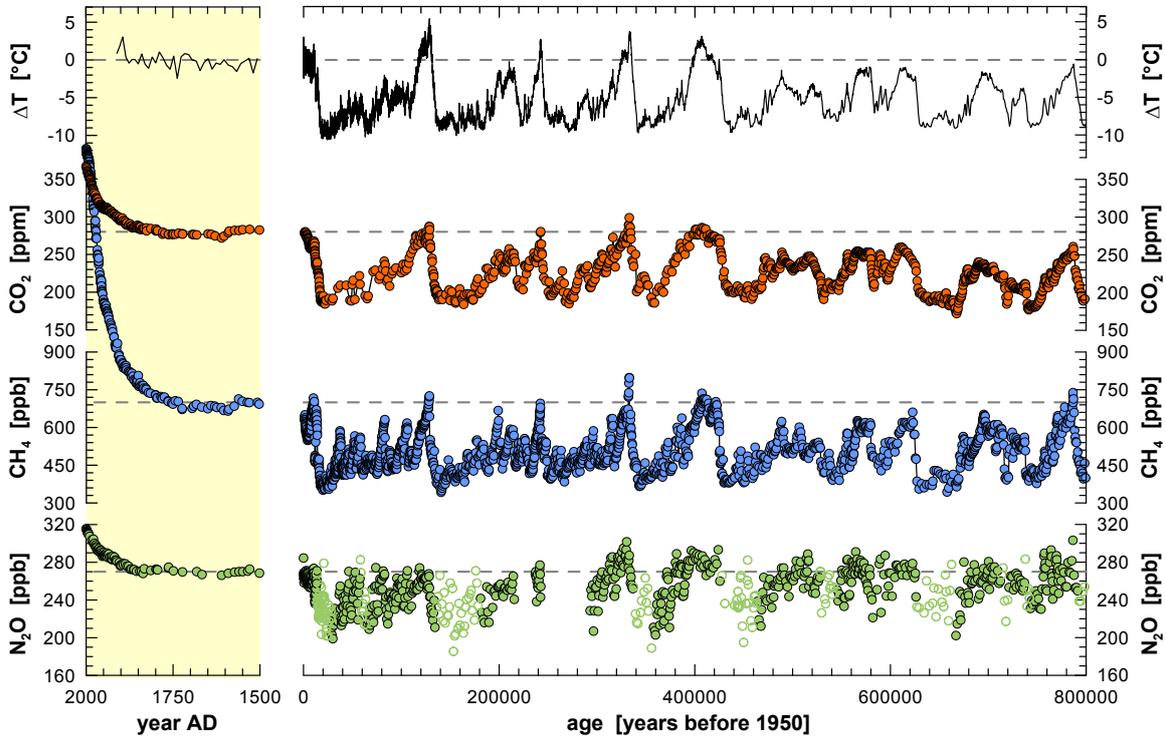


Figure 1.8: Isotope temperature (δD) proxy reconstructions and greenhouse gas concentrations over the last 800,000 years as measured on the EPICA (European Project for Ice Coring in Antarctica) ice core from Dome C [EPICA community members, 2004; Jouzel et al., 2007; Louergue et al., 2008; Lüthi et al., 2008; Petit et al., 1999; Siegenthaler et al., 2005; Spahni et al., 2005].

Changes in the solar irradiance related to variations in the Sun's activity exist but are rather small (Fig. 1.9). Those changes comprise the well known 11 year sunspot cycle as well as longer centennial cycles such as the 87 years Gleissberg cycle and are related to changes in the solar constant S_0 by 1-3 W/m^2 . On longer time scales (100,000 years) changes in the eccentricity of the Earth's orbit lead to changes in S_0 of similar magnitude. In contrast changes in the obliquity (tilt of the Earth's axis varies between 21.5° and 24.5° with a 41,000 years cycle) and the precession (gyroscopic motion of the Earth's axis relative to a fixed star and precession of the major axis of the Earth orbit around the Sun with an efficient cycle of approximately 21,000 years) lead only to redistribution of the incoming energy within different seasons or within different latitudinal bands.

Figure 1.10 summarizes the knowledge of the radiative forcing changes until the first decade after the new millennium. As can be clearly seen, the greenhouse gases contribute most to the current warming with their level of scientific understanding (LOSU) being high allowing for an unambiguous attribution. In contrast, solar irradiance changes contribute little to the warming, while the net aerosol effect leads to a cooling on a global scale. Note, however, that the LOSU of the aerosol effects is still rather low. Overall the radiative balance has changed by about $2.3 W/m^2$ since 1750 AD.

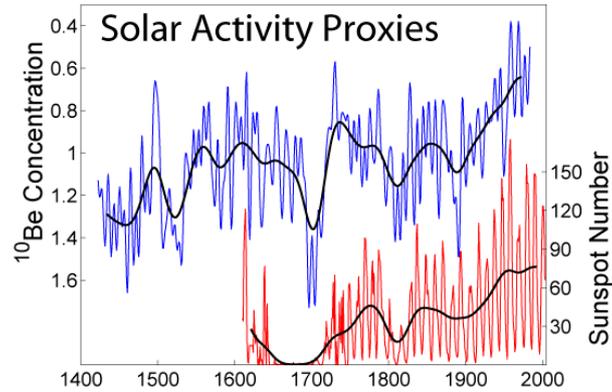


Figure 1.9: Sunspot number (red) and solar activity proxy ^{10}Be produced in the stratosphere as measured on a Greenland ice core (blue) [Beer *et al.*, 1994]. A change in the sunspot number on the order of 100 equals approximately a change in S_0 by 1 W/m^2 .

In the last decade this forcing has increased further, leading to a $1.1 \text{ }^\circ\text{C}$ average global warming observed for the decade 2010-2019 relative to the preindustrial reference interval of 1850-1900. Note that this is a global average warming, while continental areas have experienced typically a larger surface warming. For example the average temperature increase in Switzerland is about twice the global value!

1.3 Aerosols in the atmosphere

The spatial and temporal changes in the radiative balance together with the rotation of the Earth control wind speeds and steer the atmospheric circulation. Aerosols (i.e. the suspension of fine particles in the μm range in air) provide tracers that are passively transported together with the circulation and, thus, allow reconstructing transport patterns. In addition, the production of many aerosol species is directly dependent on wind speed as described below. Moreover, as already described above, the aerosols have a direct and an indirect effect on the radiative balance of the Earth.

Let us start the discussion on aerosols with a short introduction of the role of particles in the atmosphere on the radiative balance. Overall the energy transfer in the atmosphere (and thus also the remaining shortwave radiation at the surface) is dependent on the wavelength-dependent extinction of radiation, where extinction is the sum of absorption and scattering. Absorption implies energy uptake by molecules or particles in the atmosphere, while scattering on gas molecules, particles and cloud droplets comprises various processes that redistribute/redirect the incoming radiation. For example the planetary albedo is the consequence of backscattering by the atmosphere and reflection of the incoming shortwave radiation at the surface (see Fig. 1.11 for a schematic overview of the related processes). The contribution of scattering to the total extinction of light is also commonly called the single scattering albedo.

Let us assume a downward shortwave radiation intensity I , which is penetrating the atmosphere from above. The intensity of this downward radiation is declining due to absorption and scattering of radiation into other directions by molecules and particles. Due to the scat-

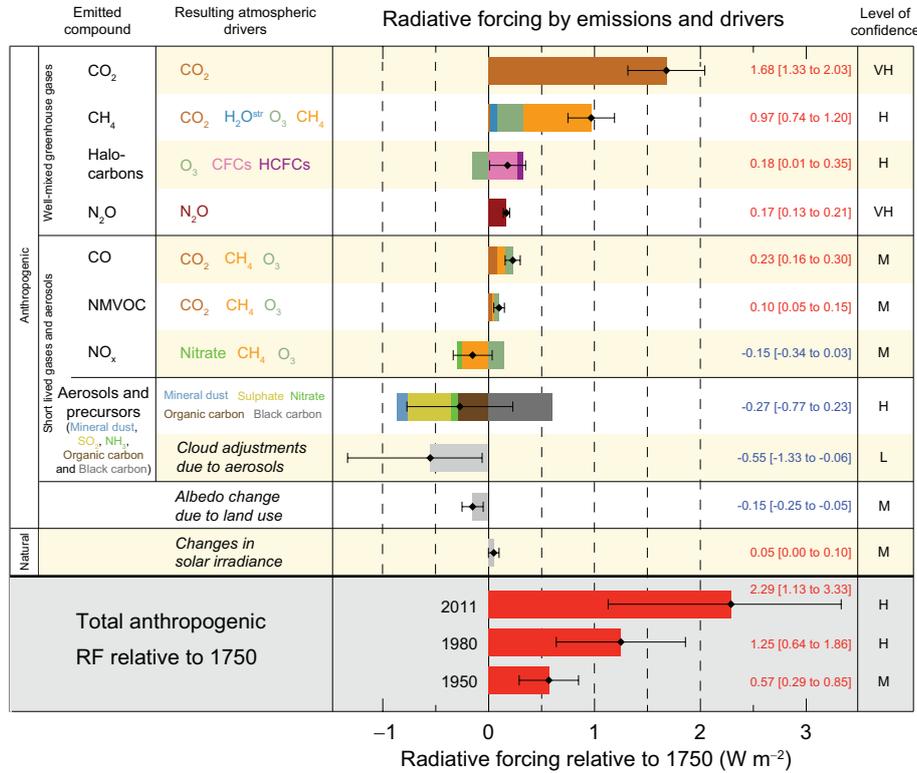


Figure 1.10: Radiative forcing changes since 1750 by natural and anthropogenic changes [IPCC, 2013].

tering processes, a thin horizontal layer of the atmosphere becomes also a source of radiation into other directions. Vice versa the layer is not only irradiated from the initial radiation from above but also by scattered light from other atmospheric layers. Moreover, molecules and particles also emit black body radiation.

Light that reaches the surface without scattering is called direct radiation, while scattered light is called diffuse radiation. Obviously the more scattering occurs, the stronger becomes the diffuse light and the weaker the direct radiation. An observation we make, when scattered light from thin clouds coming from all directions blinds our eyes much more than the direct light from a blue sky.

The change in the (frequency dependent) intensity is given by the so-called radiation transfer equation

$$\frac{dI}{dz} = -\beta_{ext}I + J, \quad (1.14)$$

where β_{ext} is the frequency dependent extinction coefficient (that is the downward intensity reduction by absorption and scattering) and J is the emission term, i.e. black body radiation emitted by the molecules or particles themselves. Note that the extinction coefficient is the sum of both absorption as well as scattering, which are both wavelength dependent, i.e. $\beta_{ext}(\lambda) = \beta_{abs}(\lambda) + \beta_{scat}(\lambda)$.

In the following we will neglect the emission term, as we are only interested in the short wave part of the spectrum. Based on this simplified differential equation the change in intensity

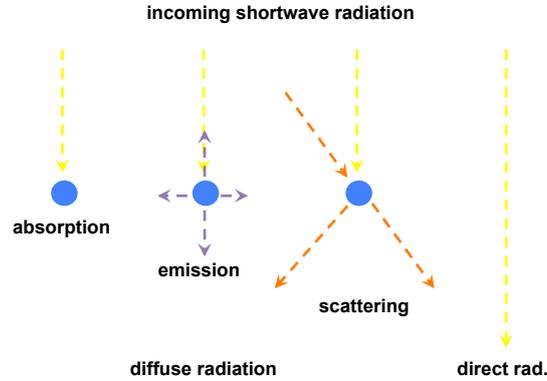


Figure 1.11: Schematic of various processes leading to extinction and scattering of incoming shortwave radiation.

along its pathway through the entire atmosphere from the top of the atmosphere (TOA) to the ground can be described according to

$$I(z = 0, \lambda) = I(z = TOA, \lambda) e^{-\int_0^{TOA} \beta_{ext}(z, \lambda) dz}. \quad (1.15)$$

where $I(z = 0, \lambda)$ is the wavelength dependent intensity at the ground and $I(z = TOA, \lambda)$ is the wavelength dependent intensity at the top of the atmosphere. The integral in the exponent is called the column integrated optical depth of the atmosphere and takes into account that the extinction coefficient is a function of altitude z due to changing density and aerosol concentration.

Clearly, the full solution of this equation, which is dependent on the angular distribution of scattered light, on wavelengths and the vertical change in β_{ext} is beyond the scope of this introduction. However, if we neglect the emission term and assume that the extinction coefficient is constant throughout the atmosphere, the solution of the equation above is the well-known Lambert-Beer law

$$I(z, \lambda) = I_0 e^{-K(\lambda)z}, \quad (1.16)$$

which shows that the intensity of light at a certain wavelength declines exponentially with the thickness of the irradiated medium, where in this case $I_0 = I(z = TOA, \lambda)$.

The scattering of light is largely dependent on the size of the scattering particles. For example if the scattering particles are gas atoms or molecules in the atmosphere, their size is much smaller than the wavelength of the incident radiation. In this case, molecules are polarized by the incoming radiation (i.e. the electrons of the atoms are interacting with the radiation) making the molecule itself an emitter of electromagnetic radiation. This so-called Rayleigh scattering is strongly dependent on wavelength (proportional to $1/\lambda^4$), thus radiation with shorter wavelengths is scattered much more strongly than for longer wavelengths. This process is the reason that the sky appears to be blue, when you look away from the Sun, because the light with short wavelengths is scattered more efficiently towards your eyes. Vice versa, the sky during a sunset becomes red in direction of the Sun, because the shorter wavelength

light is more efficiently scattered away from the incident light direction. For particle sizes much larger than the wavelength of the incoming light, the geometric scattering on such (approximately spherical) particles becomes dominant. In contrast to Rayleigh scattering, this process is not dependent on wavelength. E.g. the light scattered in a cloud (where the cloud droplets are much larger than the wavelength of light) is subject to this process and therefore clouds appear to be white. In between these two extremes, i.e., when the particle size is similar to the wavelength of light, we speak of Mie scattering which is weakly dependent on wavelength. The overall occurrence of various scattering processes dependent on wavelength of the radiation and the particle radius of the scattering particle is depicted in Fig. 1.12.

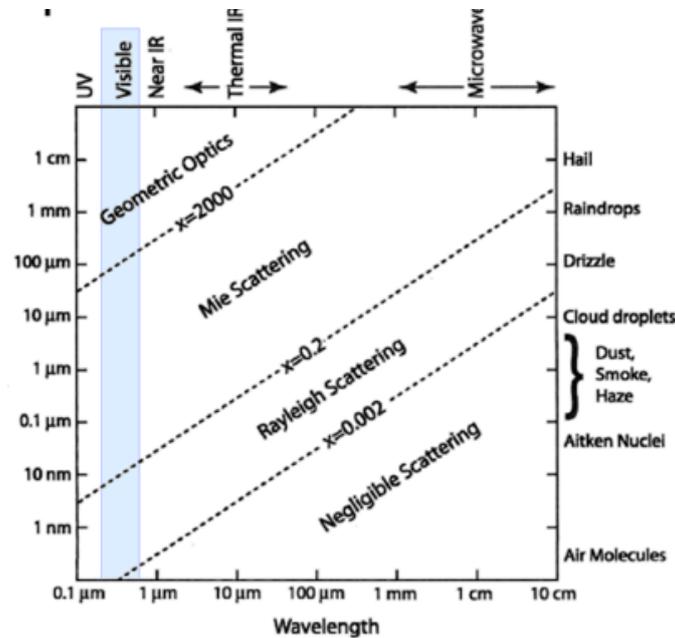


Figure 1.12: Dependence of various scattering processes on the diameter of the particle that leads to scattering and the wavelength of the incident light.

In the following, we will discuss the general physical and chemical properties of aerosol particles. In principle, there are two ways to characterize different aerosol species. Either by their chemical composition, which is reflecting the source they are derived from, or by their physical properties such as total mass and size. The most important aerosol components are summarized in Table 1.1. The most important contributors to the total global aerosol load are sea salt aerosol and mineral dust aerosol [Fischer *et al.*, 2007]. Sea salt aerosol is mainly produced by dispersion of sea water and the chemical composition of sea salt aerosol, thus, reflects the composition of sea water (Table 1.2). The process that leads to the entrainment of sea salt particles is the bursting of small air bubbles in the surface water that are formed by wave breaking. These bubbles ascend to the sea surface where they burst and produce small film droplets, that can be easily picked up by the wind and enter the troposphere by turbulent mixing. Sea salt aerosol formation is strongly dependent on wind speed with the sea salt aerosol flux out of the ocean being an exponential function of the surface wind speed.

In addition to this open ocean production of sea salt aerosol, sea ice has been recently identified as an alternative source of sea salt aerosol. The formation of so called frost flowers at the sea surface during the sea ice formation process [Wagenbach *et al.*, 1998] as well as easily wind mobilized snow lying on the sea ice [Yang *et al.*, 2008], which is enriched in sea salt by sea spray and brine soaking, regionally represent also very efficient sea salt aerosol sources. This sea salt aerosol from sea ice, however, shows a chemical signature that is depleted in sulfate and slightly depleted in sodium compared to the sea water composition due to the precipitation of mirabilite ($\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$) during the sea ice freezing process. This chemical depletion has been clearly observed in aerosol samples from coastal Antarctic sites during winter months, when the sea ice aerosol source becomes dominant [Wagenbach *et al.*, 1998].

The second most important aerosol is mineral dust, which is produced in arid and semiarid areas in both hemispheres [Prospero *et al.*, 2002] by chemical and physical weathering of crustal material. The process that leads to entrainment of small dust particles into the boundary layer is called saltation [Marticorena and Bergametti, 1995]: Larger particles are horizontally accelerated by surface wind and collide with other particles. The impact of that collision is able to provide vertical momentum to the smaller particles that leads to entrainment into the boundary layer and subsequently efficient uplift. The strength of the dust aerosol source is highly dependent on climate conditions in the source regions [Genthon, 1992], i.e.

- the lack of vegetation cover,
- the intermittent availability of water that accelerates weathering followed by drying of the soil,
- the soil moisture which determines the threshold wind speed u_t that is required to mobilize any dust particles,
- the wind speed u (above the threshold velocity u_t) that determines the dust flux into the atmosphere. In general this dust source flux is proportional to $u^2(u - u_t)$.

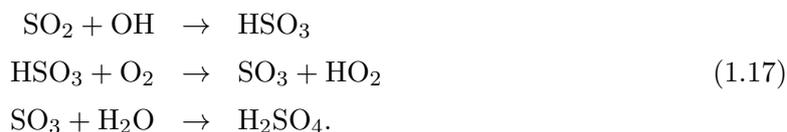
Similar to sea salt, the chemical composition of mineral dust aerosol reflects the composition of its source, i.e. it reflects crustal material (Table 1.3). Thus Si, Al, Fe, Ca, and K are the most important contributions. However, the spatial variability of the chemical composition of rocks is large and fractionation processes may occur during weathering and uplift processes.

Both sea salt and mineral dust aerosol are formed by mechanical dispersion of its source material. The second possibility to produce aerosol is condensation of supersaturated gaseous precursors. The most important example for this pathway is the production of sulfate aerosol. The latter is produced by the oxidation of SO_2 (e.g. derived from industrial emissions, volcano eruptions or biological activity) in the atmosphere by the reaction with the OH radical to

Source	Estimated flux in Tg yr ⁻¹
Natural	
Primary	
Mineral dust	
0.1 – 1.0 μm	48
1.0 – 2.5 μm	260
2.5 – 5.0 μm	609
5.0 – 10.0 μm	573
0.1 – 10.0 μm	1490
Seasalt	10,100
Volcanic dust	30
Biological debris	50
Secondary	
Sulfates from DMS	12.4
Sulfates from volcanic SO ₂	20
Organic aerosol from biogenic VOC	11.2
Anthropogenic	
Primary	
Industrial dust (except black carbon)	100
Black carbon	12 ^a
organic aerosols	81 ^a
Secondary	
Sulfates from SO ₂	48.6 ^b
Sulfates from SO ₂	21.3 ^c

Table 1.1: Annual fluxes of the most important aerosol sources. From *Seinfeld and Pandis* [2006]. ^a Tg C; ^b Tg S; ^c Tg NO₃⁻.

form sulfuric acid according to



In the presence of water vapour, sulfuric acid (H₂SO₄) has a very low vapour pressure and immediately forms very small aerosol particles. This process of new aerosol formation is called homogeneous nucleation. Alternatively, the SO₂ can be dissolved in already existing aerosol particles or cloud droplets and can react in the water phase with oxidants such as H₂O₂ to

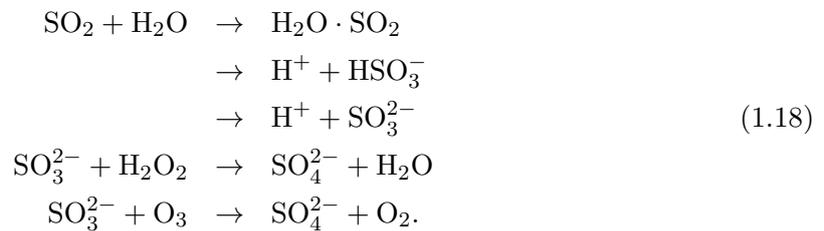
Species	Percent by weight
Cl	55.04
Na	30.61
Mg	3.69
Ca	1.16
K	1.1
Br	0.19
NO_3^-	$3 \cdot 10^{-6} - 2 \cdot 10^{-3}$
NH_4^+	$1.4 \cdot 10^{-6} - 1.4 \cdot 10^{-5}$

Table 1.2: Chemical composition of sea water. From *Seinfeld and Pandis* [2006].

Elemental abundances (ppmw)	
Element	Crustal rock
Si	277,200
Al	81,300
Fe	50,000
Ca	36,300
Mg	20,900
Na	28,300
K	25,900
Ti	4,400
Mn	950

Table 1.3: Elemental composition of different rocks, soil, shale and continental aerosol [*Warneck*, 1987]. Accordingly, the chemical composition of mineral dust is slightly different from the mean crustal composition and may vary from one dust source region to the other.

form sulfate in the aerosol phase:



The two different pathways (dispersion vs. homogeneous nucleation) for sea salt and mineral dust aerosol on the one side and sulfate aerosol on the other side differ significantly in the size of the aerosol produced. Dispersion particles are typically a few μm large while nucleation particles are smaller than $0.1 \mu\text{m}$.

This brings us to the physical characterization of aerosol particles. The most convenient way to characterize aerosol particles physically is by their size distribution (Fig. 1.13). We distinguish

Aerosol Size Distribution

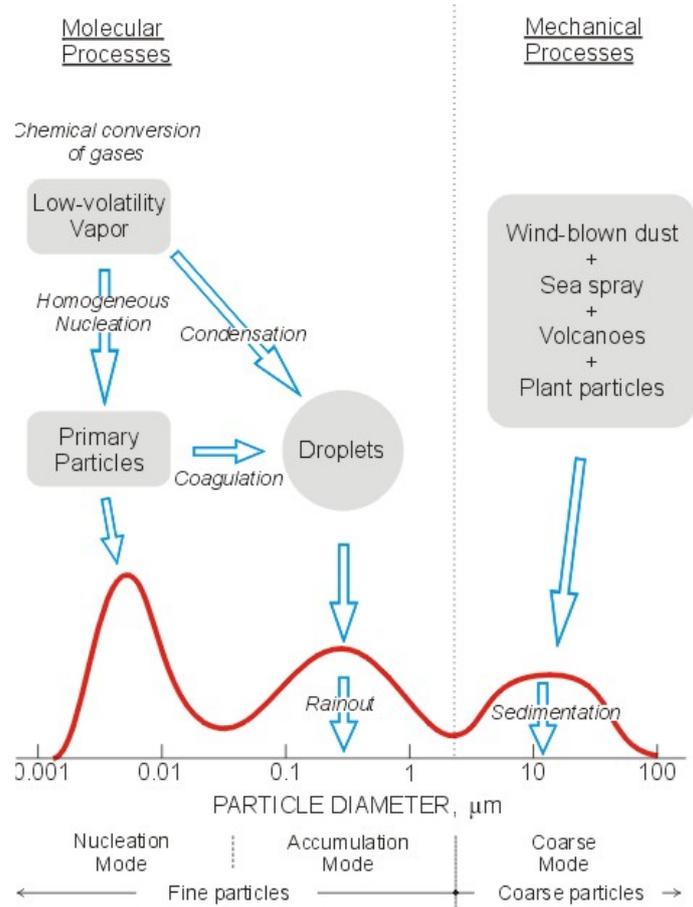


Figure 1.13: Idealized volume size distribution of an aged particle.

- the number distribution $dN/d(\log r)$, i.e. the number of aerosol particles in a specific (logarithmic) size bin,
- the area distribution $dA/d(\log r) = 4\pi r^2 dN/d(\log r)$, i.e. the total aerosol surface in a specific size bin,
- the volume distribution $dV/d(\log r) = 4/3\pi r^3 dN/d(\log r)$, i.e. the total volume of particles in each size bin,
- the mass distribution $dM/d(\log r) = \rho dV/d(\log r)$, i.e. the aerosol mass accumulated in each size bin. If the density of the particles is the same in all size classes, then the mass distribution and the volume distribution are directly proportional.

In the formulas above we assumed that all particles are spheres. In this case the number, area, volume, and mass distribution can be easily transformed into each other. From the formulas above it becomes clear that the mass of the aerosol is dominated by larger particles (proportional to r^3) while the majority (in number) of particles is small.

Close to the source areas the size distribution of the aerosol body is dominated by the production process. E.g. over the ocean, the dispersion mode is most pronounced due to the predominance of sea salt aerosol formation. In contrast, for large atmospheric SO_2 concentrations the nucleation mode becomes more important. When looking at Fig. 1.13, there exists a third size mode with particles between 0.1 and 1 μm . This mode is called accumulation mode and is derived from smaller particles that gain in size by coagulation. Accordingly, this process is the sink that "removes" small nucleation particles from the atmosphere. Remote from any aerosol sources, where large particles have been largely lost from the aerosol body by sedimentation (see below) and where small particles had the chance to coagulate, the accumulation mode is generally dominant in the volume size spectrum. The rate of coagulation of particles of two different size bins (the chance of a particle to hit another one) is dependent on the mobility of the particles (controlled by its diffusion constant, which is the larger the smaller the particles are) and the cross section of the particle (which obviously is the larger the greater the radius of the particle is). This coagulation rate J_{coaag} can be written as

$$J_{coaag} = 4\pi(r_1 + r_2)(D_1 + D_2)n_1n_2, \quad (1.19)$$

where $r_{1,2}$ are the radii of the two particles with $n_{1,2}$ the number of particles in a certain size bin and $D_{1,2}$ their diffusion constants. From this equation it becomes clear that the coagulation rate is highest for the collision of one small and one large particle.

We have now seen how aerosols are produced, so the question remains how are accumulation and dispersion particles removed from the atmosphere, i.e. how are they deposited back onto the surface. In general we can distinguish dry and wet deposition. The dry deposition is controlled by the effective dry deposition velocity v_{dry} of the particles in the air column (which for large particles is controlled by gravitational sedimentation but for small particles by other processes such turbulent aerodynamic resistance, impaction of wind transported particles at the surface or adhesion forces close to the surface) and the dry deposition flux J_{dry} can be written as

$$J_{dry} = v_{dry}C_{air}, \quad (1.20)$$

where C_{air} is the atmospheric concentration of the particles. The sedimentation of large particles (tens of μm) can be approximated by Stokes law for the friction of spherical particles in a medium of dynamic viscosity η ($\eta = 18.27 \cdot 10^{-6}$ Pa s for air). The sedimentation velocity v_{sed} is determined by the balance between gravitation (corrected for buoyancy of the particle of density ρ in air of density ρ_{air}) and the friction force (Stoke's force, Fig. 1.14)

$$\frac{4}{3}\pi r^3(\rho - \rho_{air})g = 6\pi\eta r v_{sed}, \quad (1.21)$$

and, therefore

$$v_{sed} = \frac{2}{9}(\rho - \rho_{air})g \frac{r^2}{\eta}. \quad (1.22)$$

As can be seen from this equation, the sedimentation velocity and, thus, the sedimentational deposition flux is proportional to r^2 and, thus, is larger the larger the particles are.

Accordingly, the larger dispersion particles are mainly removed from the atmosphere by sedimentation. Note however that for small particles sedimentation is not the controlling dry deposition process and this equation does not apply.

The second process removing aerosols from the atmosphere is wet deposition by liquid scavenging processes. These can either occur in the cloud (nucleation scavenging or rain out) or below the cloud, where falling droplets collect aerosol particles when they fall to the surface (wash out). These processes are rather independent of the particle radius and the wet deposition flux J_{wet} can be estimated using an effective scavenging ratio ϵ (the ratio of the aerosol concentration in the precipitation and the air at the same time). Thus

$$J_{wet} = \epsilon P C_{air}, \quad (1.23)$$

where P is the precipitation rate.

The total aerosol deposition is the sum of the wet and dry deposition

$$J_{tot} = J_{wet} + J_{dry} = \epsilon P C_{air} + v_{dry} C_{air} = (\epsilon P + v_{dry}) C_{air} = v_{tot} C_{air}. \quad (1.24)$$

If we assume an air column of height H with a (well mixed) aerosol concentration C_{air} (Fig. 1.15), then the change in the air concentration with time in this column is dependent on the aerosol deposition to the surface

$$\frac{dC_{air}}{dt} = -\frac{v_{tot}}{H} C_{air}. \quad (1.25)$$

The solution to this equation is

$$C_{air} = C_{air,0} e^{-\frac{v_{tot}}{H} t} = C_{air,0} e^{-\frac{t}{\tau}}, \quad (1.26)$$

with $\tau = H/v_{tot}$ being the atmospheric residence time of the aerosol particles.

Accordingly, in absence of any aerosol production (i.e. away from its source) the aerosol concentration declines exponentially with time dependent on the wet and dry deposition.

As we have seen above, for large particles the dry deposition velocity is a function of the radius squared of the aerosol particle (when sedimentation dominates) with larger particles

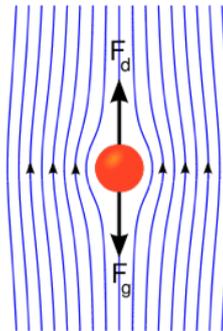


Figure 1.14: Balance of gravity and Stoke's force for a sphere falling through a medium.

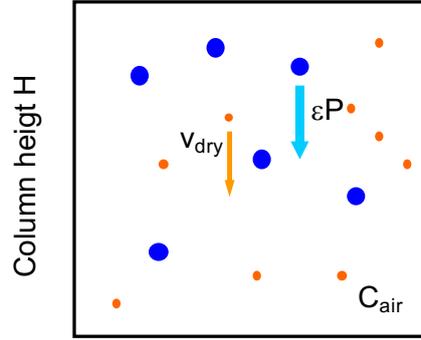


Figure 1.15: Conceptual view of the aerosol depletion in an air column away from its source by dry deposition of aerosol particles (orange dots) and wet deposition by droplets (blue dots).

being deposited faster. In fact, the differential equation above is valid for each size bin and can be written in the case of sedimentation as

$$\frac{dC_{air}(r, t)}{dt} = \frac{\frac{2}{9\eta}g(\rho - \rho_{air})r^2 + \epsilon P}{H} C_{air}(r, t). \quad (1.27)$$

and, thus,

$$C_{air}(r, t) = C_{air,0} e^{-\frac{\frac{2}{9\eta}g(\rho - \rho_{air})r^2 + \epsilon P}{H} t} = C_{air,0} e^{-\frac{t}{\tau}} \quad \text{with } \tau = \frac{H}{\frac{2}{9\eta}g(\rho - \rho_{air})r^2 + \epsilon P} \quad (1.28)$$

This implies that the lifetime τ is smaller for large particles than for smaller particles and that the air column will deplete faster in large particles. Effectively, the size distribution will shift to smaller particles away from the source. Vice versa, if we keep the lifetime τ constant but vary the transport time (thus wind speed) we could say that the faster the transport time, the more large particles will survive the transport over a given distance.

1.4 Atmospheric structure and circulation

The physics of fluid dynamics (see also chapter 9) applies in principle in the same way for the atmosphere as for the ocean. Note, however, that air and, thus, the atmosphere is compressible. This is reflected in the general gas law (the equation of state for air)

$$p = \frac{RT}{V_{mol}} = \frac{\rho RT}{M_{mol}}, \quad (1.29)$$

where the volume of a mol of molecules V_{mol} is inversely related to the pressure p and linearly related to the temperature T (ρ is the density, M_{mol} the molar mass of the gas and $R = 8.314$ J/molK the general gas constant).

In addition, the atmosphere allows for evaporation and condensation of water vapor, thus, for latent heat fluxes that alter the energy distribution in the atmosphere. Finally, the viscosity of air is much lower compared to that of ocean water, affecting the time and spatial scales of phenomena encountered in the atmosphere compared to those in the ocean. For example,

synoptic phenomena, such as cyclones in the atmosphere or large-scale eddies in the ocean, last only on the order of days in the atmosphere while they prevail for months in the ocean. Vice versa their spatial extent in the atmosphere covers thousands of kilometers while they are several hundred km large in the ocean.

We will start with the discussion of the vertical structure of the atmosphere. We can relate the vertical pressure gradient to gravity in the so called hydrostatic balance (see also Chapter 9) by

$$dp = -g\rho dz = -\frac{pgM_{\text{mol}}}{RT} dz. \quad (1.30)$$

Literally, this differential equation implies that any change in the pressure dp over a small depth interval dz is caused by the weight of the air column of thickness dz . From this follows (by separation of variables and integration) the well-known barometric formula

$$p = p_0 e^{-\frac{M_{\text{mol}}gz}{RT}}, \quad (1.31)$$

where we assumed that the temperature does not change with altitude (for a more rigorous derivation, we would also have to integrate over the temperature change in z).

Assuming a constant temperature over the whole depth of the atmosphere is obviously a crude simplification, as everybody knows, who ever climbed a mountain. Typically, the temperature declines by up to 1 °C with every 100 m of altitude. To derive the vertical temperature gradient, we apply the 1st law of thermodynamics, which reflects the energy conservation in the gas ($dU = dQ - pdV$, where $dU = C_v dT$ is the internal energy of the gas which is determined by its temperature, dQ is the heat exchange of the air parcel with its environment and $-pdV$ the work done by expansion of the gas when the pressure is lowered at higher altitudes). Assuming an adiabatic cooling of an air parcel, when lifting it to higher altitudes (adiabatic means no heat exchange with the environment: $dQ = 0$), and no condensation of water vapour, we can write

$$\begin{aligned} dQ_{\text{mol}} &= C_v dT + pdV_{\text{mol}} \\ 0 &= (C_v + R) dT - \frac{RT}{p} dp, \end{aligned} \quad (1.32)$$

where dQ_{mol} is the heat exchange per mol, $C_v = 5/2R$ is the molar heat capacity for constant volume of a gas consisting of two atom molecules (N_2 , O_2) such as air ($C_p = C_v + R$ is the heat capacity of air for constant pressure) and where we used $d(RT) = RdT = d(pV_{\text{mol}}) = V_{\text{mol}} dp + pdV_{\text{mol}}$. Using the hydrostatic balance and the general gas law, it follows the dry adiabatic lapse rate with

$$\frac{dT}{dz} = -\frac{gM_{\text{mol}}}{C_p} = -0.01 \text{ K/m}. \quad (1.33)$$

This is often fulfilled at the lee sides of mountains, where, due to the descending motion of the air mass, no condensation occurs. If a moist air mass is cooled and condensation of water vapor occurs, this condensation provides latent heat to the air parcel (this is equivalent to $dQ \neq 0$) and, thus, the lapse rate is smaller. In the case of the dry adiabatic temperature

gradient we base the deduction of the first law of thermodynamics on one Mol. For the moist adiabatic it is more convenient to base it on volume. Thus, we can write

$$dQ_v = \frac{C_p p}{RT} dT + \frac{M_{\text{mol}} g p}{RT} dz, \quad (1.34)$$

where dQ_v is the heat exchange per volume. With the latent heat provided by the condensation of water vapour per mass L (at 20°C this is approximately 2.46 MJ/kg) and the water density at saturation $\rho_{w,\text{sat}}$ we obtain

$$dQ_v = -L d\rho_{w,\text{sat}}. \quad (1.35)$$

Using this in the equation above leads to

$$\begin{aligned} -L d\rho_{w,\text{sat}} \frac{dT}{dT} &= \frac{C_p p}{RT} dT + \frac{M_{\text{mol}} g p}{RT} dz \\ \frac{dT}{dz} &= -\frac{M p g}{C_p p + R L T \frac{d\rho_{w,\text{sat}}}{dT}} \end{aligned} \quad (1.36)$$

As one can see the relationship depends on pressure and temperature. To give an rough estimate the moist adiabatic temperature gradient in the lower troposphere (between 500 and 1000 hPa and a temperature of 10°C to 20°C) we find -0.005 K/m, ie., roughly half of the dry adiabatic one.

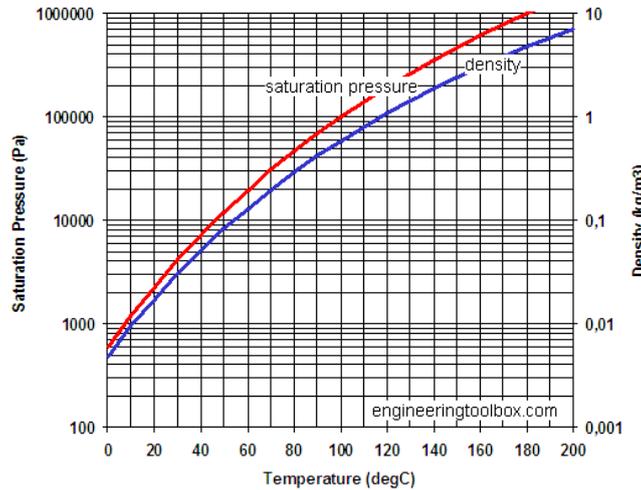


Figure 1.16: Temperature dependency of the saturation water vapor pressure (red) and the water vapor density in air in equilibrium (blue).

Besides, we can be seen in Fig. 1.17 that the vertical temperature structure of the complete atmosphere is not just a linear decline in temperature. The latter only holds for the troposphere and the mesosphere, while both in the stratosphere and the thermosphere the temperatures generally increase with altitude. The reason for this is that the assumption of an adiabatic

cooling breaks down in these layers. Here, heating by the incoming solar radiation exceeds the cooling caused by the expansion of the gas with altitude.

Fig. 1.4 shows that the Earth receives annually significantly more incoming shortwave radiation at the equator than at the poles. Accordingly, the air at the equator warms at the surface and starts to ascend while the air at the poles cools and descends. Because of this general vertical motion, the pole is a constant high pressure area at the surface while the tropics have generally lower pressures at the surface. High in the atmosphere this pressure gradient is reversed. In the case that the Earth would not rotate, these pressure gradients would lead to a surface air flow from the pole to the equator and a return flow aloft. This general scheme is called Hadley circulation. As displayed in Fig. 1.18, the Earth does not show one large Hadley cell in each hemisphere, but is divided into a tropical Hadley cell characterized by the surface trade winds, the so called Ferrell cell in mid latitudes characterized by strong westerlies and cyclonic activity, a polar cell in each hemisphere with generally easterly winds. As we will see this separation is a consequence of the rotation of the Earth. The movement of an air parcel is controlled by the forces acting on it. In case of a rotating sphere, these forces are

1. gravity,
2. forces by pressure gradients,
3. friction at the surface,
4. the centrifugal force accelerating air outwards perpendicular to the rotation axis,
5. the Coriolis force (see also chapter 9), which reflects the different speed the Earth moves at the surface for different latitudes.

To describe what happens in a more pictorial way, we move an air parcel northward from a lower to a higher latitude (without exerting friction at the surface). At its low latitude starting point the air parcel should move with the same speed as the Earth surface from west to east, accordingly an observer standing on the Earth surface would not see a movement of the air parcel. At higher latitudes the Earth surface will move slower than the air parcel (due to the lower radius relative to the rotation axis) and in absence of friction the air parcel would leave the surface behind. An observer at the ground would observe an eastward movement of that air parcel which would require a force in that rotating system, the Coriolis force. In fact, the Earth's rotation is also the reason why the Hadley cell cannot extend all the way to the pole. Moving an air parcel from the equator northward, the air parcel has to conserve angular momentum, if no friction processes occur. Due to the reduction of the radius relative to the rotation axis, the westerly wind of the air parcel would have to increase the more northward the parcel is moved. Thus, the westerly wind speed would become infinite at the pole, which is unphysical. In fact, the westerly air stream becomes unstable (turbulent) when critical wind speeds are reached and dissolves in eddies superimposed on a generally westerly wind direction. The respective Newton's law, describing the forces on an air parcel and its resulting movement on a rotating coordinate sphere, is called the Navier Stokes equation and

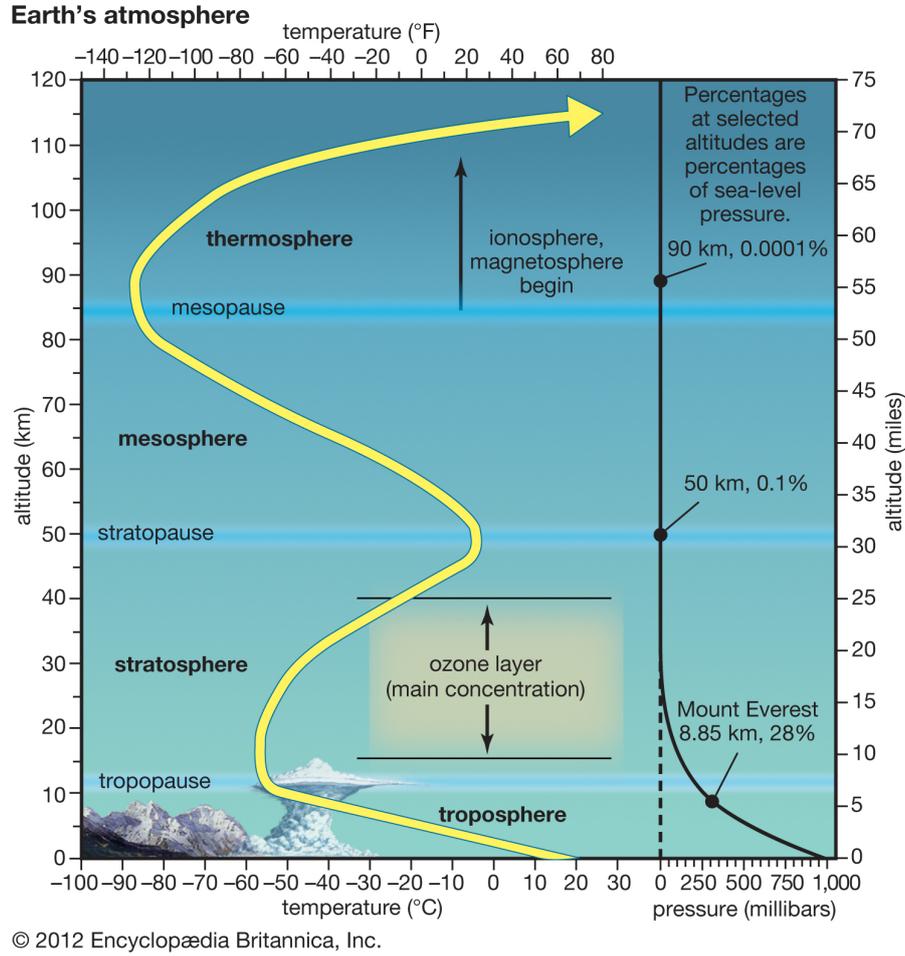


Figure 1.17: General vertical temperature structure of the atmosphere.

can be written as

$$\rho \frac{D\mathbf{v}}{Dt} = \mathbf{F}_{\text{gravity}} + \mathbf{F}_{\text{pressure gradient}} + \mathbf{F}_{\text{friction}} - 2\rho\boldsymbol{\Omega} \times \mathbf{v} - \rho\boldsymbol{\Omega} \times \boldsymbol{\Omega} \times \mathbf{r}, \quad (1.37)$$

where $\boldsymbol{\Omega}$ is the angular rotation of the Earth and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \quad (1.38)$$

is the total (or material) derivative, which considers both the temporal change as well as spatial gradients in a moving fluid such as the atmosphere and u, v, w are the components of the velocity \mathbf{v} in x, y, z direction. Consider e.g. the temperature in the atmosphere in a moving air parcel. This can change at a certain fixed point in the atmosphere because of local heating e.g. by solar radiation (first term on the right side of the total derivative) or because the air is advected to somewhere else where the temperature is higher (last three terms on the right side).

Neglecting the small effect of the centrifugal force and any friction in the vertical, the Navier

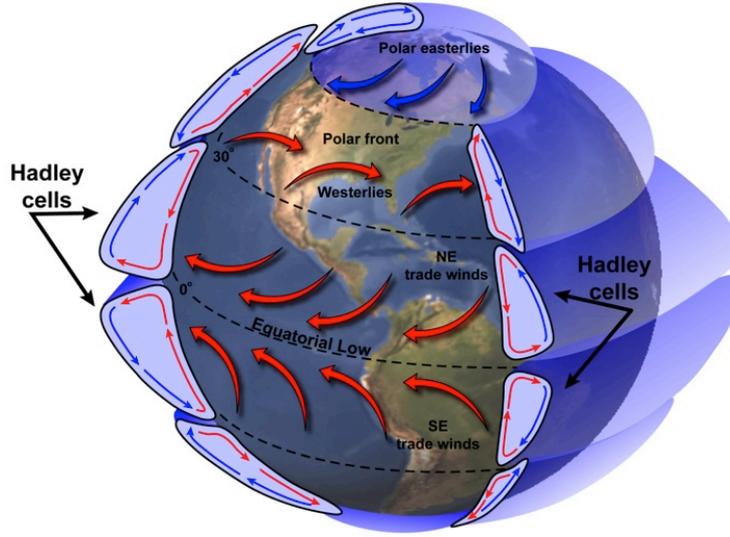


Figure 1.18: Latitudinal structure of the typical circulation zones in the atmosphere (from NASA Earth observatory).

Stokes equation can be written in local x, y, z coordinates as

$$\begin{aligned}\rho \frac{Du}{dt} &= -\frac{\partial p}{\partial x} + F_{fr,x} + 2\rho\Omega v \sin \phi \\ \rho \frac{Dv}{dt} &= -\frac{\partial p}{\partial y} + F_{fr,y} - 2\rho\Omega u \sin \phi \\ \rho \frac{Dw}{dt} &= -g\rho - \frac{\partial p}{\partial z}.\end{aligned}\tag{1.39}$$

This looks complicated but can be solved for some easy but most generally applying cases, when the air parcel has come to equilibrium with its driving forces ($Dv/Dt = 0$, i.e. the air parcel is not accelerated anymore). In that case, the last of the three equations represents only the already well known hydrostatic balance.

Away from the surface (neglecting any frictional forces), the two first equations can be simplified to

$$\begin{aligned}0 &= -\frac{\partial p}{\partial x} + 2\rho\Omega v \sin \phi \\ 0 &= -\frac{\partial p}{\partial y} - 2\rho\Omega u \sin \phi\end{aligned}\tag{1.40}$$

and, thus,

$$\begin{aligned}u_g &= -\frac{\partial p}{\partial y} \frac{1}{2\rho\Omega \sin \phi} = -\frac{1}{\rho f} \frac{\partial p}{\partial y} \\ v_g &= \frac{\partial p}{\partial x} \frac{1}{2\rho\Omega \sin \phi} = \frac{1}{\rho f} \frac{\partial p}{\partial x},\end{aligned}\tag{1.41}$$

where the index g is added to illustrate that this is the geostrophic wind and $f = 2\Omega \sin \phi$ is

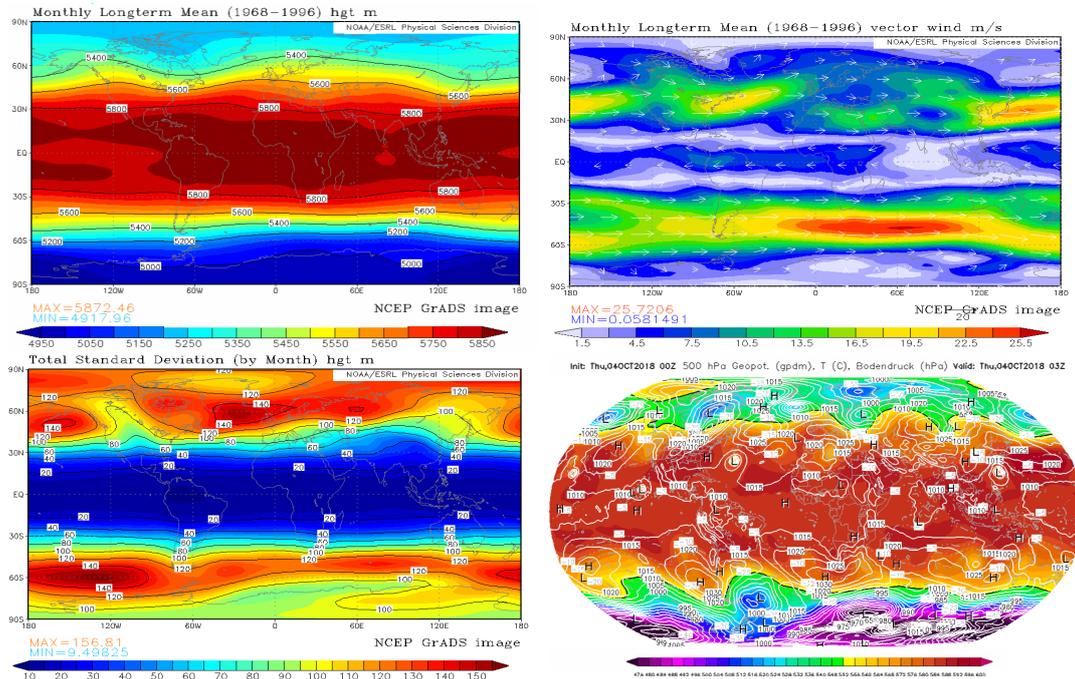


Figure 1.19: Mean annual geopotential height at the 500 hPa level (thickness of the air column from the bottom to a pressure of 500 hPa) (top left) and the mean annual wind speed in the free troposphere at the same pressure level (top right). Annual standard deviation of the geopotential height at the 500 hPa level (bottom left) and the geopotential height at the 500 hPa level (colors) and the surface pressure (contours) for the 26th of October 2008. (bottom right) (from NCEP/NCAR reanalysis project and GFS model of the American Weather Service).

the Coriolis parameter. According to the last two equations the air parcel moves in a direction perpendicular to the pressure gradients, i.e. along the isobars. When looking from the high pressure to the low pressure the air parcel is moving to the right/left on the northern/southern hemisphere. Thus, on the northern/southern hemisphere air circulates around high pressure in clockwise/anticlockwise direction, while the opposite holds true for low pressure centers. This phenomenon is called geostrophic wind and reflects the balance of the pressure gradient force and the Coriolis force. This is the case in the free troposphere (e.g. at the 500 hPa level, i.e. about 5000 m above the surface) in the annual mean (Fig. 1.19). The generally higher pressure in the upper troposphere at low latitudes leads to westerly wind speeds at the 500 hPa level. The westerly wind speeds are highest in the mid latitudes, where also the latitudinal pressure gradients are the highest (Fig. 1.19). Note, however, that in this region the wind direction changes rapidly on a daily basis and these areas are generally characterized by strong cyclonic activity (as our weather report illustrates every day). This is illustrated by the very high standard deviation of the geopotential height in those regions (Fig. 1.19), especially over the North Atlantic and the Southern Ocean. Looking on a daily global pressure chart in Fig. 1.19 it also becomes clear that the pressure distribution is much more complex in middle latitudes than the latitudinal distribution in the annual mean. Thus, the westerly air stream on a global scale is only a net air movement in middle latitudes overprinted by a

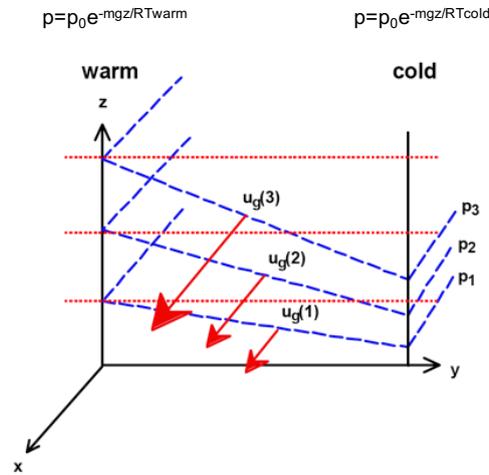


Figure 1.20: Concept of the thermal wind.

very high synoptic activity of the atmospheric circulation.

A specific case of the geostrophic wind is the thermal wind. If we consider two points on the Earth of different temperature the pressure decline with altitude at each point will be different (Fig. 1.20). According to the barometric formula, the pressure declines faster at the cold side. This implies that at a certain fixed altitude the pressure is lower at the cold side and a pressure gradient force acts from the warm to the cold side. On the northern hemisphere this thermally caused pressure gradient lead to a wind speed directed to the right when looking from the higher pressure (warmer side) to the lower pressure (colder side). Due to the continuing decline of pressure with altitude this pressure gradient increases with altitude and so does the thermal wind speed.

To deduce the so-called thermal wind equations, we make use of four equations: both components of the geostrophic wind (Eq. 1.41), the hydrostatic relation (Eq. 1.30), and the ideal gas equation (Eq. 1.29). Using the ideal gas equation in the geostrophic wind components and the hydrostatic relation results in in three equations:

$$\begin{aligned}
 \frac{fv_g}{T} &= R_L \frac{\partial \ln p}{\partial x} \\
 \frac{fu_g}{T} &= -R_L \frac{\partial \ln p}{\partial y} \\
 \frac{g}{T} &= -R_L \frac{\partial \ln p}{\partial z}
 \end{aligned}
 \tag{1.42}$$

with $R_L = R/M_{\text{mol}}$ being the gas constant of air and M_{mol} the mol mass of air. Taking the derivative with respect to z of the first equation and the derivative with respect to x of the third equation and similarly for the second equation the derivative with respect to z and the derivative with respect to y of the third equation we obtain two equations where pressure is

eliminated, the so-called thermal wind equations:

$$\begin{aligned}\frac{\partial v_g}{\partial z} &= \frac{g}{fT} \frac{\partial T}{\partial x} + \frac{v_g}{T} \frac{\partial T}{\partial z} \\ \frac{\partial u_g}{\partial z} &= -\frac{g}{fT} \frac{\partial T}{\partial y} + \frac{u_g}{T} \frac{\partial T}{\partial z}\end{aligned}\quad (1.43)$$

The last term on the right hand side of both components is usually small compared to the first term and is therefore often ignored. The thermal wind vector relation, with its components, describes not really a wind but the change of the geostrophic wind with height. It is a useful tool to check observed wind and temperature fields for consistency, as we find that if the y -component of the geostrophic wind, v_g , increases with height the relationship shows that the temperature must increase to the East ($\partial T/\partial x > 0$). Similarly, for the x -component u_g , which says that if it increases with height we expect a temperature increase to the South ($\partial T/\partial y < 0$).

The phenomenon of thermal wind is most pronounced at temperature fronts as illustrated by weather maps. Here, we observe a very strong wind in direction of the fronts. It also occurs e.g. at the so called polar front that separates warmer mid latitude air masses from the cold polar air. Here, it leads to a westerly air stream that strongly increases with altitude and is focused at high altitude in the sub polar jet with wind speeds above 70 m/s.

As described above the westerly air stream becomes unstable for higher wind speeds and can be disturbed by small variations in the pressure distribution, e.g. caused by mountains or the land/sea contrast in temperature. This disturbance process is called cyclogenesis. A complete description of this process is beyond the scope of this introduction, but qualitatively eddies are induced by an initial zonal disturbance in the pressure distribution. According to the geostrophic wind, air circulates around the high pressure center in Fig. 1.21 in clockwise direction in the northern hemisphere and the low pressure system in anticlockwise direction. This leads to a net air stream equatorwards east of the high and polewards west of the high. The poleward movement is connected to an ascending air motion increasing the surface low. This represents a positive feedback that amplifies the disturbance until strong fronts, high wind speeds and finally eddies develop.

A second relatively easy solution of the Navier-Stokes equation can be found close to the surface, we assume that the friction force is in equilibrium with the Coriolis force and the pressure gradient:

$$\begin{aligned}0 &= -\frac{\partial p}{\partial x} + K\rho \frac{d^2 u}{dz^2} + 2\rho\Omega v \sin \phi \\ 0 &= -\frac{\partial p}{\partial y} + K\rho \frac{d^2 v}{dz^2} - 2\rho\Omega u \sin \phi.\end{aligned}\quad (1.44)$$

Splitting the wind (u, v) into its geostrophic and ageostrophic part, i.e., $u = u_g + u_{ag}$ and $v = v_g + v_{ag}$ and assuming that the geostrophic wind is independent from z , we obtain two

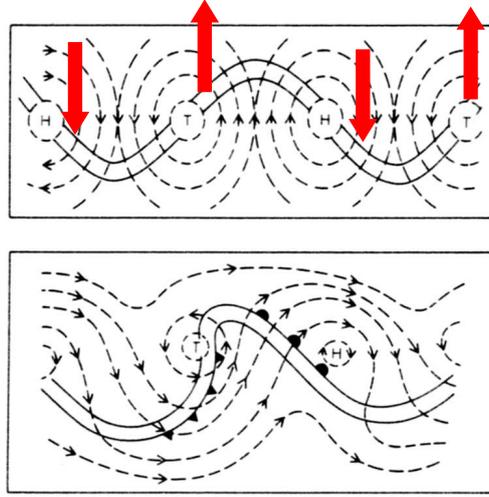


Figure 1.21: Conceptual view of the cyclogenesis in middle latitudes. Figure from *Roedel* [1994].

partial differential equations for the ageostrophic wind $((u_{ag}, v_{ag}))$

$$\begin{aligned} 0 &= K\rho \frac{\partial^2 u_{ag}}{\partial z^2} + 2\rho\Omega v_{ag} \sin \phi \\ 0 &= K\rho \frac{\partial^2 v_{ag}}{\partial z^2} - 2\rho\Omega u_{ag} \sin \phi. \end{aligned} \quad (1.45)$$

Without going into detail, this can be solved using an exponential solution

$$u_{ag} + iv_{ag} = V_0 e^{-\sqrt{\frac{\Omega \sin \phi}{K}} z} e^{-i\sqrt{\frac{\Omega \sin \phi}{K}} z}, \quad (1.46)$$

using the Euler formula for a complex number $c = a + ib$:

$$e^c = e^{a+ib} = e^a \cdot e^{ib} = e^a \cdot (\cos(b) + i \sin(b)).$$

As displayed in Fig. 1.22, this solution represents the atmospheric counterpart to the Ekman spiral in the ocean, which we will later find in Fig. 9.4. Thus, as the wind stress on the ocean leads to a shear in the ocean strata, so does the friction of the surface on the bottom layers of the atmosphere. This leads to a deviation of the wind direction from the geostrophic flow in the free troposphere towards the low pressure side of the geostrophic pressure distribution. With increasing altitude above the surface the direction spirals up and increases in wind speed until it finally matches up with the geostrophic flow at altitudes above 1-2000 m. This altitude dependent differential angle of wind speed is for example used by balloonists to navigate their balloon to some extent by changing their altitude.

1.5 Statistical description – modes of variability

Besides physical equations, researchers utilize statistical relationships to describe atmospheric circulation. One simple example is averaging, e.g. physical quantities such as temperature,

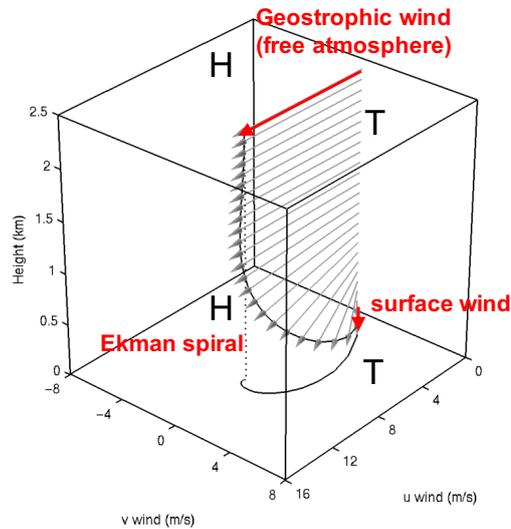


Figure 1.22: Atmospheric Ekman spiral matching the geostrophic flow caused by the horizontal pressure gradients in the free troposphere.

pressure, or precipitation. Averaging implicitly uses a statistical model; specifically, the assumption is that the variable of interest follows the normal (Gaussian) distribution. The approach is descriptive and defines a normal and deviations. For example, climatologies focus on the means and variability around the mean state [von Storch and Zwiers, 1999]. Recently, the paradigm has shifted away from a descriptive approach to a comprehensive analysis of the dynamics, and statistics plays an important role [von Storch and Zwiers, 1999]. The aim of this section is to use El Niño Southern Oscillation (ENSO) as an example to illustrate these paradigms.

Definitions

Before presenting the ENSO and some underlying mechanisms suggested in the literature, one needs to clarify what a mode of variability is, and how it is distinguished from the terms 'pattern' and 'regime', as these terms are often misused in the literature. We follow the definition given by *Stephenson et al.* [2000]:

- **Pattern:** A pattern is a coherent spatial structure with regard to at least one variable of interest. This variable may be a physical variable such as temperature or pressure, but may also be any statistical measure of dependence between two variables, such as the correlation.
- **Regime:** A regime is a preferred state in time in a system of interest. The system considered here is the atmosphere and/or the climate system. In the case that the climate system has more than one stable state, the system is able to flip between those if it is perturbed by either internal (random) fluctuations or external forcing. In this context, the term 'tipping point' is relevant. A tipping point describes any process in which beyond a critical threshold, the rate of the process increases dramatically. In

other words, the system shifts abruptly on time scales much smaller than compared to either previous or new stable state regimes.

- **Mode of variability:** A mode of variability is a physically interpretable, and physically meaningful spatial pattern together with a time series of weights, which describe the influence on the system of interest. The expression is vague, as it is not defined what 'physical meaningful' exactly means. Still, one important assumption is that the system of interest can be split and described in a good approximation by a finite sum of several 'space-time modes'. Several statistical approaches are used to deduce such a description, of which the most prominent is the EOF analysis (Box 1.1). Note however, that these statistical methods do not guarantee that the resulting modes are physical meaningful (Example given in Box 1.1). A hint that the statistically derived mode is physical meaningful is if the time series of the mode has a preferred periodicity.

Box 1.1 EOF analysis and its problems

The EOF analysis decomposes a field, e.g. sea level pressure (slp, p), into a sum of several space-time modes, where the mode is a product of a time series (pc) and a spatial pattern (EOF):

$$p(x, y, t) = \sum_{k=1}^q pc_k(t) EOF_k(x, y) + \epsilon(x, y, t)$$

where $k = 1 \dots q$ is the number of separations and $\epsilon(x, y, t)$ the remaining noise (normally white noise is assumed). Thereby the basis functions are orthogonal to each other and chosen in such a way that the residual variance is minimized. The pairs (EOF and pc) are sorted according to their explained variance of the total variance of the variable. The basis functions are found by computing the eigenvectors of the covariance matrix of the data set (more details in e.g., *von Storch and Zwiers* [1999]).

Clearly, the EOF analysis is a powerful technique to identify modes in a statistical sense but caution must be taken when these modes are physically interpreted. *Dommenget and Latif* [2002] showed examples where EOF analysis failed to identify meaningful modes, and says that real modes of variability shall not depend on the statistical method of detecting them.

The first two EOFs and regression patterns of the tropical Atlantic SST anomalies are presented here. The EOF1 pattern shows a monopole structure for the entire tropical Atlantic whereas is EOF2 signifies a dipole structure around the equator. EOF received attention in the literature (see *Dommenget and Latif* [2002] for discussion) to be a potential physical mode of SST variability on decadal timescales. However, when applying another method like simply regressing the SST fields on indices deduced in the "centers of action" of the dipole structure suggest that the the regions vary independently from each other and thus questioning the hypothesis of a physical meaningful mode of variability in the tropical Atlantic.

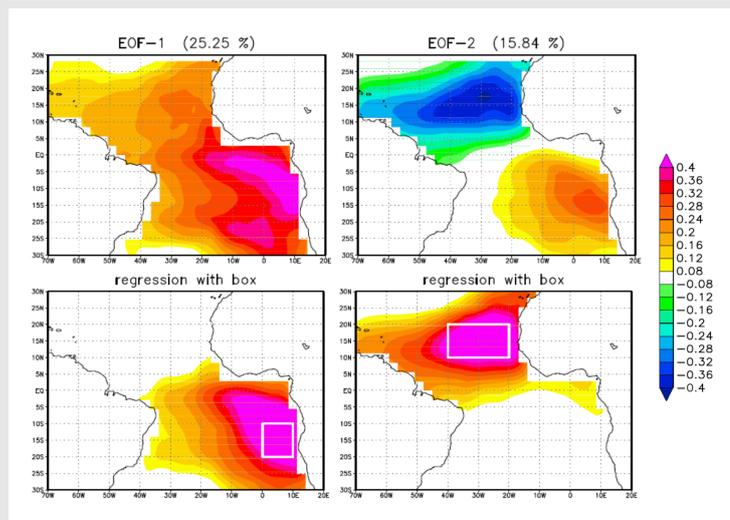


Figure: EOF analysis applied to sea surface temperatures (SST) of the tropical Atlantic (top panels) and (bottom panels) using simple point regression maps where the points are selected in the center of actions (white boxes) are averaged and the index is regression on the SST field. Figure from *Dommenget and Latif* [2002].

ENSO – the phenomenon

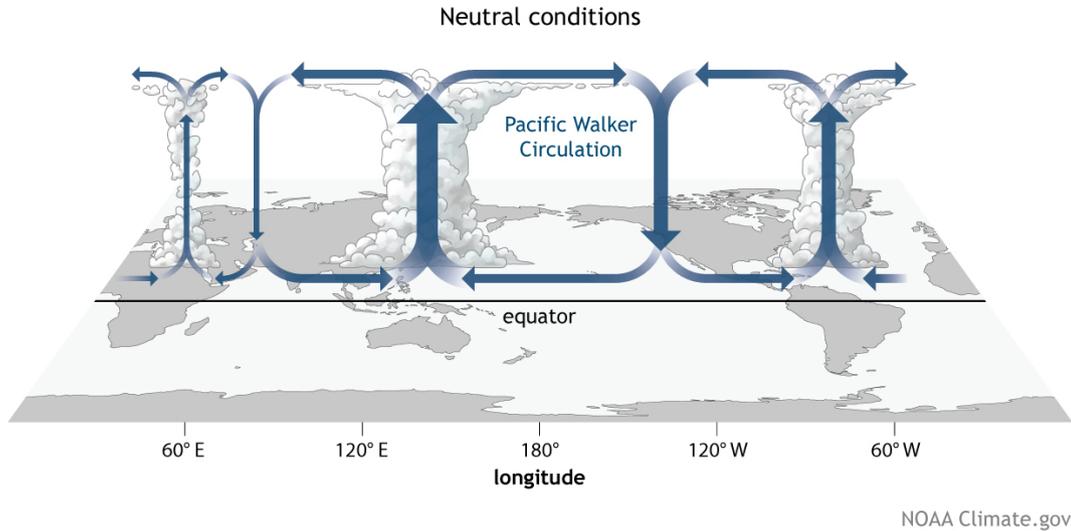


Figure 1.23: Sketch of the Walker Circulation with overturning cells in the Atlantic, the Indian Ocean and the Pacific. Figure from NOAA (<https://www.climate.gov/news-features/blogs/ens0/walker-circulation-ens0-atmospheric-buddy>).

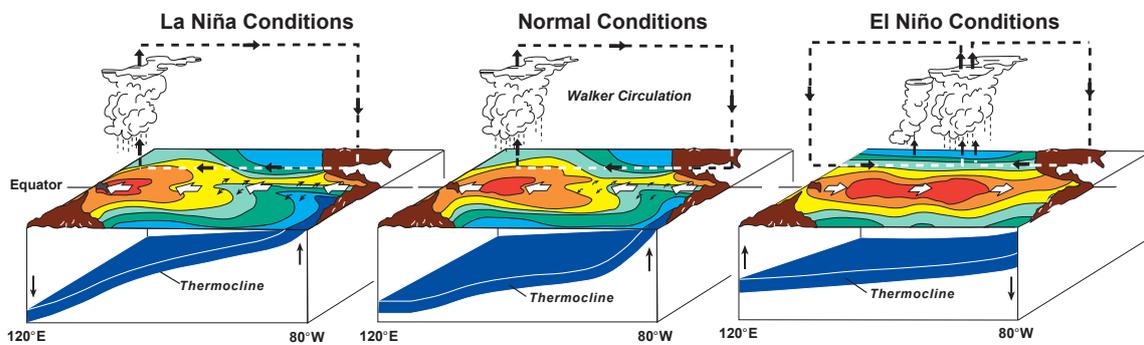


Figure 1.24: Sketch of the three phases of ENSO [Stocker, 2013].

The phenomenon ENSO consists of an atmospheric and oceanic part. The atmospheric part was first identified by *Walker* [1924]; an anti-correlation of the pressure between the eastern and western tropical Pacific. This anti-correlation is part of the Walker Circulation, which spans the entire equatorial region (Fig. 1.23). The reason for this is that the trade winds in the tropical Pacific push surface water masses from the eastern to the western boundary leading to a east west sea surface temperature (SST) gradient. Over warm water (called warm pool region) air masses tend to rise and thus low surface pressure results. The rising air masses in the west, in the warm pool area, lead to intense precipitation, mainly induced by massive thunderstorm systems. These systems are fueled directly by strong solar heating

and indirectly via evaporation from the warm Pacific underneath. Next, condensation heat is released in rising air masses, and this sustains the rising motion. The rising motion in this region reaches the tropopause and returns eastward and descends over the colder east Pacific, leading to high pressure. This closes the zonal overturning cell. Note that such large-scale zonal circulations can be found solely in a sufficiently narrow belt along the equator where the Coriolis force is negligible.

The ocean part reacts on the atmospheric forcing: The trade winds cause upwelling of cold ocean water at the eastern boundary. This cold anomaly moves westward at the ocean surface while warming up by the intense tropical solar radiation, so that finally there is a relatively low SST at the eastern boundary of the tropical Pacific Ocean and a relatively high SST at the western boundary. A subsurface feature of the ocean is the thermocline, a region of sharp temperature change in the vertical, that divides the ocean in two parts: a warm upper part and a cold lower part, respectively. This thermocline reacts to the mean westward surface wind and drives the ocean motions poleward in both hemispheres in the first 50 m. This water needs to be replaced by upwelling waters. This causes a shallower thermocline in the eastern tropical Pacific compared to a deeper one in the warm pool region. Thus, in the western tropical Pacific warmer water is upwelled whereas in the eastern Pacific cold water is transported upward causing the cold SST in front of the western coast of South America.

So, clearly one sees that the atmospheric circulation and the ocean work hand in hand (neutral state in Fig. 1.24). Superimposed to this normal state of the tropical Pacific is an irregular cycle of cooling and warming which affects both the atmospheric and the ocean circulation, leading to Bjerknes hypothesis.

The Bjerknes hypothesis and phases of ENSO

In the seminal study, Bjerknes [1969] suspected that the ENSO dynamics arises through ocean-atmosphere interactions in the tropical Pacific. The essence of Bjerknes's hypothesis is that of two self-sustained, mutually inverse developments of the atmosphere-ocean system in the tropical Pacific [Neelin *et al.*, 1998]: SST anomalies in the Pacific cause the trade winds to strengthen or to weaken and that this in turn changes the ocean circulation in such a way that SST anomalies are generated. In the words of Bjerknes *"A change toward a steeper pressure slope at the base of the Walker Circulation is associated with an increase in the equatorial easterly winds and hence also with an increase in the upwelling and a sharpening of the contrast of surface temperature between the eastern and western equatorial Pacific. This chain reaction shows that an intensifying Walker Circulation also provides for an increase of the east-west temperature contrast that is the cause of the Walker Circulation in the first place. On the other hand, a case can also be made for a trend of decreasing speed of the Walker Circulation ..."*. [Bjerknes, 1969]. To illustrate this, the two phases of ENSO are discussed in detail (sketches of Fig. 1.24):

- During **positive phases**, also called El Niño, there is a positive SST anomaly in the eastern equatorial Pacific. The zonal SST gradient is reduced, and as a consequence, the air pressure difference between the east- and west-side of the tropical Pacific. The

pressure difference leads to weaker-than-normal easterly winds in the tropical Pacific and consequently to a weaker Pacific Walker Circulation. In the case of strong positive SST anomalies the equatorial winds could even become westerly winds. Upwelling is reduced or even prevented in the eastern Pacific. The thermocline rises in the western Pacific and lowers in the eastern part of the tropical Pacific, which finally reduces the 'cold tongue' and hence the SST gradient. These changes in temperature and pressure go along with above average rainfall along the western coast of the South America, and below average rainfall in the area between Southeast Asia and northern Australia. The name El Niño refers to the observations of fishermen in Peru and Ecuador, as during warm phases upwelling of the cold and nutrient rich cold water is suppressed, which drastically reduces their catches. As this was often observed during Christmas time, the fishermen named it El Niño (the boy child; the divine infant in Spanish).

- **Negative phases** of ENSO, also called La Niña (El Niño's sister), are characterized by temperature and pressure gradients with directions as in the neutral phase but considerably larger magnitudes. Air pressure become higher than normal in the east and lower than normal in the west leading to enhanced wind velocity and wind stress over the tropical Pacific. The tropical Pacific enhances upwelling and a shallower thermocline in the eastern part. Lastly, the precipitation is increased over Indonesia.

Bjerknes had pointed out that his hypothesis lacks some important ingredients: (i) the mechanism does not explain how the system is able to switch from a phase with warm SSTs (El Niño) to a neutral phase and then to a phase with anomalous cold SSTs in the eastern tropical Pacific and (ii) it does not predict the characteristic time scales involved in ENSO.

Time scales of ENSO

Besides the patterns of the physical variables discussed so far, ENSO is characterized by a range of specific preferred time scales. Several indices are used to calculate either atmospheric based or oceanic based. Walker suggested the Southern Oscillation Index (SOI). The SOI is for a particular month i

$$\text{SOI} = \frac{P_{Ti} - P_{Di}}{\sigma_m} \quad (1.47)$$

where P_{Ti} and P_{Di} are the standardized surface pressure values of month i at Tahiti and Darwin, respectively and σ_m is the standard deviation of these standardized time series of monthly differences: $\sigma_m = \sqrt{\frac{\sum_{i=1}^N (P_{Ti} - P_{Di})^2}{N}}$. Ocean indices are deduces by averaging over predefined regions called Niño regions:

- Niño 1+2: 0-10°S and 90-80°W
- Niño 3: 5°N-5°S and 150-90°W
- Niño 4: 5°N-5°S and 160°E-150°W
- Niño 3.4: 5°N-5°S and 170-120°W

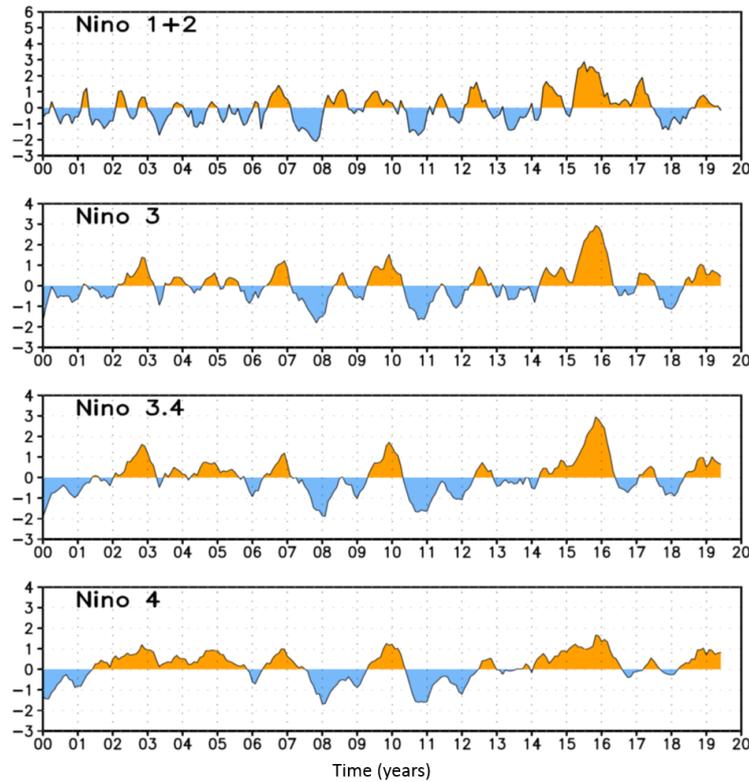


Figure 1.25: Time series of different indices used to characterize ENSO from 1996 to 2015. Source: NOAA, Climate Diagnostics Bulletin.

Note that none of the indices introduced is able to capture the full complexity of ENSO variability. Still, most of the indices try to represent a substantial portion of this variability (Fig. 1.25). All indices agree in the general picture, and identify similar periods of positive and negative indices associated to El Niño and La Niña events. The SOI index shows slightly more deviations compared to the other indices. Already the time series show preferred time scales. A wavelet analysis for the period 1871 to 2002 using monthly data (Fig. 1.26) highlights enhanced power on periods of 3 to 7 years. This is a hint that ENSO is a real mode of variability.

1.6 A selected mechanism: Delayed Oscillator

As stated by *Bjerknes* [1969] the positive feedback mechanism lacks an explanation for the perpetual turnabout from warm to cold states. One of the mechanism proposed by *Schopf and Suarez* [1988], *Battisti* [1988] and *Battisti and Hirst* [1989] is based on linear equatorial ocean dynamics – the delayed oscillator equations. Here, we follow the description presented in *Sarachik and Cane* [2010].

The starting point is the observations, where the main wind anomalies occur in the center

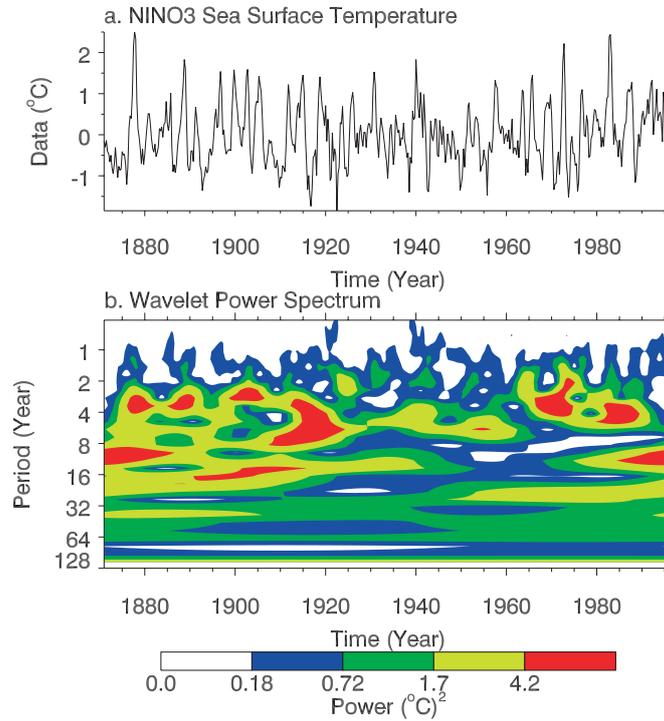


Figure 1.26: Time series of Niño 3 index for the period 1871 to 2002 using monthly data and corresponding wavelet analysis.

of the tropical Pacific, where SST changes concentrate at the eastern part of the tropical Pacific. The surface winds are controlled by the East-West temperature gradient and will depend on the SST changes in the East. The SSTs in the East are controlled by thermocline depth variations and these variations are again driven by surface wind stress. In the El Niño case (warming in the East), the wind anomaly will be westerly and form a Kelvin wave signal in the ocean. This signal leads to a deepening of the thermocline and thus enhance the warming. This happens also in regions (further east), which are not directly affected by the wind changes. In essence, the eastern Pacific is warmed by the central to western wind anomaly. The area of warmth grows and at the same time, the westerly wind anomaly is enhanced due to the growing temperature gradient. This leads to an excess of warm water in the east, which must become balanced by an area of colder water, i.e., shallower than normal thermocline. Schematically the mechanism is summarized for the El Niño case in Fig. 1.27. The dynamics at the equator only allow westward propagating signals that are equatorial Rossby waves. When the western boundary is reached, these Rossby waves are reflected as Kelvin waves, but with opposite sign. The Kelvin waves travel east and reduce the warming in the eastern tropical Pacific, but with a delay given by the travel speed of the two wave kinds.

The transition from a El Niño to a La Niña state is illustrated by the modelling results of *Schopf and Suarez* [1988], showing Hovmoeller diagrams for sea level, a measure for the thermocline (Fig. 1.28). Following *Schopf and Suarez* [1988], Figure 1.28a show the surface

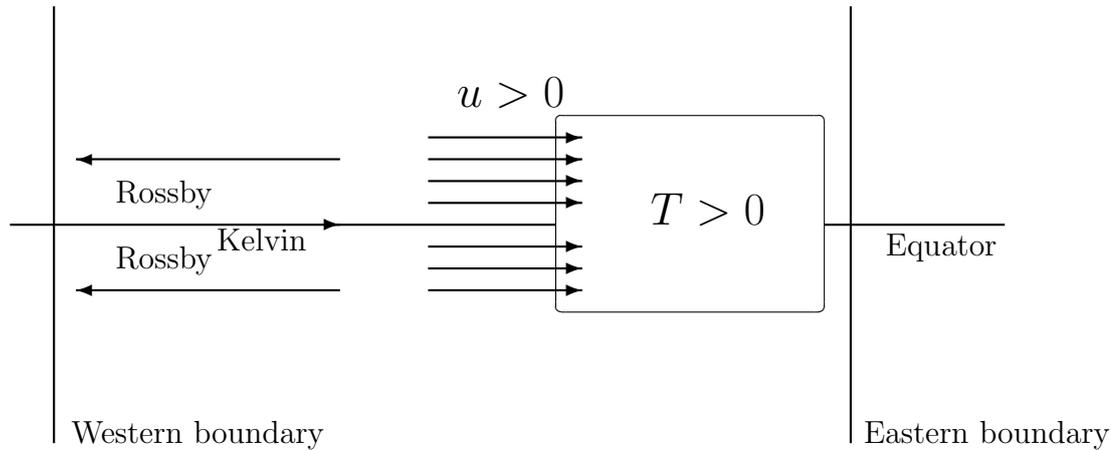


Figure 1.27: Schematic sketch of the growing warming phase of ENSO. Figure modified from *Battisti et al.* [1989].

height anomalies off the equator (note that 80°W on the left and 120°E on the right). A westward propagation is given if we see a tilting to the upper right. Here, we see a sustained height perturbation in the middle of the tropical Pacific, but no propagation from the eastern boundary. Figure 1.28b shows the equatorial height between 120°E and 140°W . The difference in meridional structure of the signal in both panels clearly shows a reflection at the western boundary a continuous phase propagation to the upper right. In Fig. 1.28c the wind anomalies are shown for the equator in the central part of the tropical Pacific. These wind anomalies are highly correlated with the surface height in this region. In Fig. 1.28d the surface height off the equator is plotted but only from the center of the wind anomaly out to the eastern boundary. This shows the strong relationship of the positive wind anomaly. Rossby wave induced negative anomalies on the thermocline propagate westward and are then reflected as Kelvin wave (but with negative signal) at the western boundary, which then moves eastwards (Fig. 1.28e). Note that at the eastern boundary the Kelvin wave is reflected as Rossby waves but spread out of the equator region leading to removal of the signal.

Schopf and Suarez [1988] combined this mechanism in the following simple equation for the temperature in the eastern tropical Pacific T (averaged over the warm area in the case of El Niño):

$$\frac{dT}{dt} = c \cdot T - b \cdot T(t - \tau) - d \cdot T^3, \quad (1.48)$$

τ is the sum of the times of the Rossby wave traveling from east to west and the reflected (at the western boundary) Kelvin wave traveling from west to east. Note that there are many Rossby waves with different speeds involved so that τ is an average (weighted by the individual contributions of each wave). The constant c represents the local growth of the warm area and b represents the effect of the returning Kelvin wave on this warm area in the eastern tropical Pacific. In principle, the total mass of water is conserved so that the amount of water going westward (with Rossby waves) should be balanced with eastward transported water via Kelvin waves. Higher-mode Rossby waves travel very slowly, which leads to a substantial

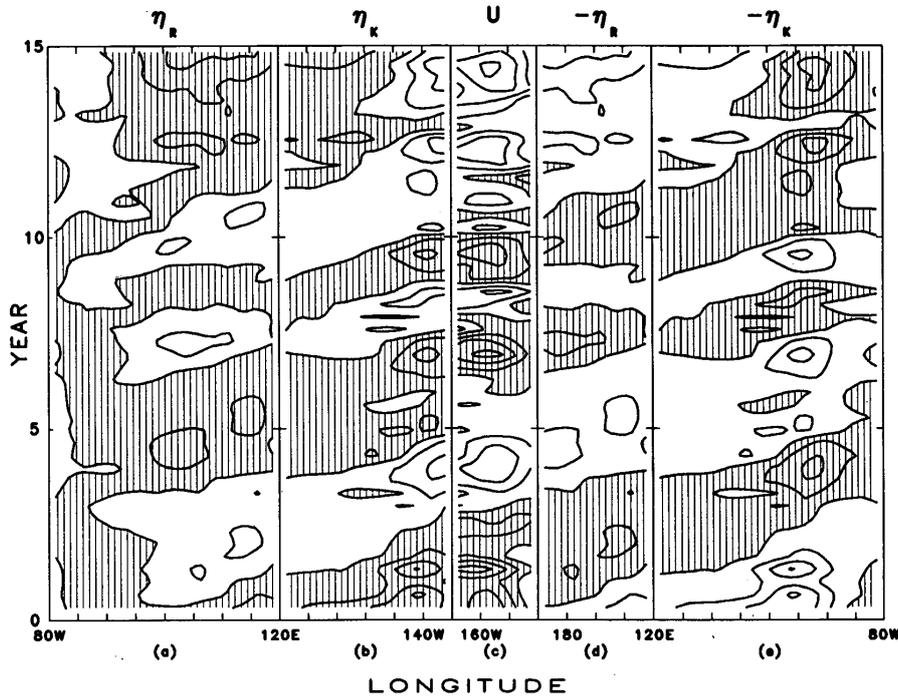


Figure 1.28: Time-longitude behavior of the oscillator: (a) η_R is the dynamic height (inverse to thermocline depth) off the equator (5° to 7°) from 120°E to 120°W , (b) η_K is the dynamic height on the equator from 80°W to 120°E , (c) zonal surface wind on the equator from 180°W to 125°W , (d) η_R from 160°W to 120°E and (e) η_K from 120°E to 80°W . Hatching is used for positive values in (a) to (c) and negative values in (d) and (e). Figure from *Schopf and Suarez* [1988].

dissipation, so that *Schopf and Suarez* [1988] assumed $b < c$. Thus, the last constant d steers the equilibration of the growth of the warm area. The terms on the right hand side of Eq. 1.48 are:

- The first term represents the unstable exponential growth of the warm area.
- The second term damps this growth via the effect of the returning Kelvin wave.
- The third term is the non-linear term needed to stop the unstable growth at some finite amplitude. It is essential for periodic solutions.

For this equation oscillations are found between two fixed points given by $T = \pm\sqrt{(c-b)/d}$, if $b < c$ and $c > 0$. *Schopf and Suarez* [1988] deduced this equation in a heuristic way by what they have seen in Fig. 1.28. *Battisti* [1988] and *Battisti and Hirst* [1989] derived a similar simple delayed-oscillator model from a more complex coupled ocean-atmosphere model. This is beyond this lecture but is presented in *Sarachik and Cane* [2010].

Chapter 2

Stable isotopes in climate and environmental physics

2.1 Introduction

This chapter is dedicated to studies of climate and environmental physics for which stable isotopic composition is relevant. First of all, we clarify the term "isotope". It was suggested by *Soddy* [1914] and literally means "the very same place". This term is easily understood since isotopes are at the same position within the periodic table. Later on, the so called nuclide table was established, in which the different isotopes appear at different locations according to the different numbers of neutrons.

Isotopes \equiv Nuclides of an element with different atomic weights.

The term "isotope" was first introduced in relation to radioactive (instable) nuclides. *Thomson* [1913] could then distinguish the two stable isotopes of neon (^{20}Ne and ^{22}Ne) with his canal ray instrument. This was the proof of Soddy's hypothesis. *Aston* [1920] improved the instrument of Thompson, so that he was able to discover 212 out of 288 naturally occurring isotopes within a short time. His instrument was called "mass spectrograph" today better known under the term "mass spectrometer".

Until now, we know more than 2500 different nuclides but only about 118 different elements. Most of these nuclides are radioactive, only 254 are stable and 34 are primordial nuclides, i.e. with sufficiently long half-lives greater than the Earth's age of 4.5 billion years. Most elements have two or more stable isotopes. However, there are also elements with only one stable isotope such as beryllium, fluorine, sodium, aluminium and phosphorous etc.

2.2 Fractionation of isotopes

If the abundances of isotopes remained constant over time, isotope studies would be rather limited. However, this is not the case. In reality, isotope abundances vary from sample to sample. The question related to these variations is: Which processes are responsible for these changes? Radioactive nuclides are produced and decay and change in this way their abundances. Determinations of radioactive isotopes allow the investigation of questions related to time scales on which processes of interest are occurring. Of similar importance are measurements of stable isotopes and their variations. They are naturally incorporated into the material and allow to retrieve information about these processes. A change of abundance is often called fractionation. Principally, we distinguish two different kinds of fractionations: (1) non-equilibrium fractionation (kinetic fractionation) and (2) equilibrium fractionation.

2.2.1 Non-equilibrium fractionation

Isotope fractionations are often related to dynamic processes which occur rapidly, incompletely or unidirectional (non-reversible). In the following we will investigate a couple of these processes.

Diffusion

Diffusion can alter the isotope abundances independently of their nature (stable or radioactive). This is due to the different masses of isotopes. As an example we look at the diffusion through a tiny opening. Conditions are assumed to guarantee that the free path length of these isotopes is larger than the size of this tiny opening. Based on the kinetic gas theory, all molecules carry the same kinetic energy, E_{kin} , for a given temperature T . Therefore, the different weights of two isotopically different molecules, m_1 and m_2 , lead to the following relation between their velocities v_1 and v_2

$$E_{kin} = \frac{m_1 v_1^2}{2} = \frac{m_2 v_2^2}{2} \quad (2.1)$$

This directly leads to the diffusion law of Graham:

$$\frac{v_1}{v_2} = \sqrt{\frac{m_2}{m_1}} \quad (2.2)$$

This means that the lighter molecule with its higher velocity has a larger probability to pass through the small passage. The remaining gas will be enriched in the abundance of the heavier molecules.

When considering diffusion of gas mixtures, then improved fractionation models need to be used. It can be shown that the fractionation is related to the ratio of diffusion coefficients of involved molecules. Diffusion coefficients themselves are proportional to the inverse square root of the reduced mass.

There are several processes which could be mentioned as examples of kinetic or non-equilibrium fractionation processes. We briefly look at three of them:

Evaporation

Evaporation is a unidirectional process (from the liquid phase to the gas phase). This process is relatively complex, but it is obvious that the molecules containing the lighter isotopes have higher translational velocities and therefore a higher probability to escape from the liquid compared to molecules consisting of heavier isotopes. Hence the abundance of the heavier molecules will be enriched in the remaining liquid.

Kinetic processes

Kinetic processes are often related to slightly different dissociation constants of molecules built from different isotopes. For example: We can imagine that a molecule bond can more

easily be broken when the molecule contains a light isotope since it has a higher vibrational frequency than the same molecule with a heavier isotope. Normally this leads to fractionation effects, in particular when reactions occur incompletely [*Bigeleis*, 1965].

Metabolic effects

Organisms often produce kinetic fractionation. For example, during breathing, the oxygen isotope ^{16}O is preferentially used compared to ^{17}O and ^{18}O . Another important fractionation effect is associated with the plants' photosynthesis. CO_2 molecules are first isotopically separated by diffusion through the small openings of the leaves (stomata) and subsequently during the chemical incorporation that preferentially consumes ^{12}C compared to ^{13}C . This leads to depleted ^{13}C abundances in the plants compared to atmospheric CO_2 gas.

2.2.2 Equilibrium fractionation

Isotopic fractionation can also occur between equilibrium states (for example between different phases). This kind of fractionation is called equilibrium fractionation. Such fractionations are generally small. However, they still can be determined precisely with mass spectrometers or theoretically estimated by using their quantum mechanical energy states. In some cases, equilibrium fractionation is large enough to be estimated from classical thermodynamic methods.

Equilibrium fractionation is largest for light elements in particular hydrogen (H) and still significant for carbon (C), oxygen (O), nitrogen (N), silicon (Si), sulphur (S) and some other elements. Since the atomic masses of these elements are small, the relative mass differences between isotopically different forms of corresponding molecules are large. These differences are the key factor in explaining equilibrium fractionations. However, there are other factors involved. Otherwise it would not be explainable why for instance equilibrium fractionation is larger for molecules containing ^{40}Ca and ^{48}Ca than for molecules containing lighter carbon isotopes. One of the reasons to explain these observations has to do with the fact that such unexpected features occur when covalent bonds are involved. Atoms with such bonding have degrees of freedom for rotation and vibration which are strongly mass dependent and which lead to significant equilibrium fractionations. In contrast, atoms with ionic or metallic bonding behave more like stationary points and are much less dependent on their masses than on their electric charge. Therefore, elements that preferentially use the latter type of bonding show generally lower equilibrium fractionations. As an example for equilibrium fractionation we will look at the water molecule.

Water

Water molecules are very well suited to discuss equilibrium fractionations because hydrogen as well as oxygen positions can be occupied by different isotopes. In every single water molecule either H or D (Deuterium) and ^{16}O , ^{17}O or ^{18}O can be chosen. Therefore, water is a mixture of 9 (12) stable, isotopically different molecules with strongly differing abundances (Table 2.1). These abundances can be calculated from the natural abundances of hydrogen

and oxygen isotopes.



The main point is that the characteristics of these 9 (12) molecules are not identical. The differences are not restricted to mass difference only, but they are obvious in any of the mentioned physical and thermodynamical parameters. The largest deviations are observed when comparing characteristics of normal versus heavy water, which are mainly represented by HH^{16}O and DD^{16}O , respectively (Table 2.2).

2.2.3 Definition of the fractionation factor

The examples given above have shown what is understood by the term "to fractionate". It simply means to separate, to split into different parts or fractions. Therefore, isotopic fractionation means to partition a sample in two or more parts with different ratios of heavy (higher masses) to lighter (smaller masses) isotopes compared to their original isotopic ratio. If one part is enriched then the other part needs to be depleted in its isotope ratios for a two component separation.

Generally, such effects can be represented by the term fractionation factor. It is denoted with the symbol α . This factor characterises the splitting into two phases or two parts of the system A and B in the form

$$\alpha_{A \rightarrow B} = R_B / R_A = \frac{1}{\alpha_{B \rightarrow A}}, \quad (2.3)$$

R denotes the ratio of the heavy to the light isotopes. For example R can either be $^2\text{H}/^1\text{H}$ (often written as D/H), $^{13}\text{C}/^{12}\text{C}$ or $^{18}\text{O}/^{16}\text{O}$. R_A and R_B stand for the same isotope ratio, represent, however, different phases, reservoirs or parts of the system and have normally different values.

Fractionations tend to be small, i.e. the values for α are close to unity. Based on this experimental fact, a new parameter, ϵ was introduced to make the small values handier. ϵ is defined as

$$\epsilon_{A \rightarrow B} = (\alpha_{A \rightarrow B} - 1) \cdot 1000 \text{ ‰}. \quad (2.4)$$

Both parameters, α and ϵ , can either denote equilibrium fractionations or kinetic fractionations.

Oxygen:	^{16}O	^{17}O	^{18}O
Mixing ratio:	997,590 ppm	370 ppm	2,040 ppm
Hydrogen	^1H	^2H	
Mixing ratio:	999,850 ppm	150 ppm	

Table 2.1: Natural isotope ratios of oxygen and hydrogen.

Characteristics		H ₂ O	D ₂ O
Molar mass	g	18.01528	20.02748
Enthalpy, ΔH_v at 25°C	kcal/mol	10.519	10.851
Vapour pressure at 25°C	mbar	31.66	27.38
Melting point	°C	0	3.82
Boiling point	°C	100.00	101.42
Critical temperature	°C	373.99	370.74
Density at 25°C	kg/m ³	997.0	1204.4
Viscosity at 25°C	Millipoise	8.93	11.0
Latent heat (phase changes)	cal/mol	1436.3 at 0°C	1515 at 3.82°C
Toxicity		No	Yes

Table 2.2: Comparison of physical characteristics of normal and heavy water.

Common rules apply for isotope fractionation (of course with some exceptions):

- Heavy isotopes are enriched in the condensed phase
- The fractionation factor $|\alpha|$ is larger for light elements, e.g. $|\alpha|_{2\text{H}} \gg |\alpha|_{18\text{O}}$
- $|\alpha|$ decreases with increasing temperature
- For several isotopes within the same element, the fractionation factor $|\alpha|$ for different isotope pairs is closely proportional to their mass difference, e.g.

$$\frac{\alpha(^{18}\text{O}/^{16}\text{O}) - 1}{\alpha(^{17}\text{O}/^{16}\text{O}) - 1} = 2 \quad (2.5)$$

However, Eq. 2.5 is an approximation. Recent high precision measurements have shown that deviations from this approximation are readily observed. In the case of oxygen, this is done by using a scaled difference between ¹⁷O and ¹⁸O called $\Delta^{17}\text{O}$ following the Eq. 2.6:

$$\Delta^{17}\text{O} = 10^6 \left(\ln \left(\frac{\delta^{17}\text{O}}{10^3} + 1 \right) - 0.528 \ln \left(\frac{\delta^{18}\text{O}}{10^3} + 1 \right) \right) \quad (2.6)$$

The small deviations expressed in $\Delta^{17}\text{O}$ and given in per meg can yield important information about processes such as photolysis in the stratosphere, stratosphere to troposphere exchange, evapotranspiration and fractionations associated with the Dole effect, i.e. difference of isotopic composition between atmospheric and oceanic oxygen.

2.2.4 Delta-notation

The expression of isotope ratios, R , is important and useful to characterise a mixture of components, particularly of radionuclides (strontium, lead etc.). However, the R expression is not preferred to express measurement values of stable isotope ratios. This is due to the fact that the values are generally very small (e.g., D/H \approx 0.00015) and that the main interest lies in the differences of isotopic ratios and not the ratios themselves. The natural variations

of stable isotope ratios are also very small. Therefore, not absolute but relative isotope ratio measurements are given. These differences or deviations are expressed relative to a standard isotope ratio. This leads to the so-called Delta-notation

$$\delta_p = \frac{R_P - R_S}{R_S} 1000 \text{ ‰}, \quad (2.7)$$

where R_P and R_S are the isotope ratios of the sample (probe) and the standard, respectively. R denotes as mentioned above the different isotope ratios. Based on the element of interest, the Delta values are given for δD (or δ^2H), $\delta^{13}C$, $\delta^{18}O$ etc. The factor 1000 converts the values into the permil scale (‰).

Example

The $^{18}O/^{16}O$ -ratio of the SMOW standard is: $R_{SMOW} = 0.0020052$. The $^{18}O/^{16}O$ -ratio of the sample shall be: $R_{SAMPLE} = 0.0019851$

$$\delta^{18}O = \frac{0.0019851 - 0.0020052}{0.0020052} \cdot 1000 \text{ ‰} \approx -10 \text{ ‰}. \quad (2.8)$$

$\delta^{18}O$ deviations can be determined by dedicated isotope ratios mass spectrometers (IRMS) with multiple cup collector systems with a precision of about 0.02‰. Since a couple of years also laser based instruments are available with a similar precision for water and CO_2 isotope determinations. Since they are less expensive they rapidly increase in number.

Standards

Isotope standards are chosen by chance and correspond to zero permil on the corresponding δ -scale. Ideally, standard material is homogenous and close to the isotopic ratio of samples. This material should be available in rather large amounts, be easy to prepare and to distribute. There are organizations that are responsible for the maintenance of standard materials (National Bureau of Standards (USA) and IAEA (International Atomic Energy Agency, Vienna)). One of the mostly used standards for oxygen and hydrogen is SMOW. SMOW means "Standard Mean Ocean Water". It was defined by *Craig* [1961]. For carbon and oxygen isotopes in carbonate Vienna-PDB or in short V-PDB is used. Measurements of nitrogen isotopes are generally referenced to the corresponding ratio of the atmospheric ratio. From time to time, studies are made to determine the absolute ratios of these standard materials. Today's accepted absolute ratios are summarized in Table 2.3.

2.3 Stable isotope in the water cycle

As described above, water is a mixture of 9 molecules with different abundances. The natural abundance variability of $^2H/^1H$ and $^{18}O/^{16}O$ in the water cycle can be demonstrated by the following examples:

$$\begin{aligned} \delta^2H & 0 \text{ ‰ (ocean water) to } -400 \text{ ‰ (snow, Antarctica)} \\ \delta^{18}O & 0 \text{ ‰ (ocean water) to } -50 \text{ ‰ (snow, Antarctica);} \\ & (+40 \text{ ‰ in atmospheric } CO_2). \end{aligned}$$

Internat. standard	$^{13}\text{C}/^{12}\text{C}$	$^{18}\text{O}/^{16}\text{O}$	D/H
Vienna-PDB	11237.2	2067.2	
Vienna-SMOW		2005.2	155.75

Table 2.3: Absolute isotope ratios in parts per million (ppm) for carbon, oxygen and hydrogen for the international Vienna standards [Gonfiantini *et al.*, 1995].

During evaporation and condensation water molecules are fractionated isotopically. This fact is based on the physical characteristics of vapour pressure for H_2^{16}O , which is slightly higher compared to H_2^{18}O . Therefore, the heavier molecule prefers the liquid (more condensed) phase. This consideration allows us to already understand, at least qualitatively, the variations of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values in precipitation. Since ocean water acts as standard its δ -values are defined (δ -values $\equiv 0$), therefore evaporated ocean water, i.e. water vapour will be depleted in heavy isotopes. However, precipitation is enriched in heavy isotopes compared to the water vapour, which eventually will lead to additional depletion of the remaining water vapour. This is sketched in Fig. 2.1. Rayleigh condensation allows us to theoretically calculate the behaviour of water vapour changes along an air trajectory and to roughly understand the distribution of water isotope values as a function of latitude. The Rayleigh model is based on assumptions that are hardly anywhere completely fulfilled and therefore are only approximations of the reality. The assumptions are (i) equilibrium between vapour - condensate, (ii) condensate (rain or snow) is removed immediately after formation and does not exchange with the vapour phase again and (iii) consideration of only isolated air masses, no admixture of other air masses are allowed. However, it is that simple that we can easily develop the basic equations ourselves and understand them (see exercises). The solutions for the change of the oxygen isotopic composition of water vapour for the cloud formation and a raining cloud are given in Fig. 2.2 as a function of the remaining water vapour amount, f . When calculating isotope ratio changes we must be aware of the fact that cooling of the air mass leads also to a volume change. We compare two cases (1) an isobaric and (2) an adiabatic cooling.

Temperature	$\alpha_{v \rightarrow c}(T)$ for $\delta^{18}\text{O}$	$\alpha_{v \rightarrow c}(T)$ for $\delta^2\text{H}$
40°C	1,0082	1,064
20°C	1,0098	1,085
0°C	1,0117	1,112
-10°C	1,0129	1,129
-20°C	1,0141	1,149
-30°C	1,0156	1,173

Table 2.4: Fractionation factor between liquid and vapour phase of water. The values for temperatures below the freezing point are extrapolated!

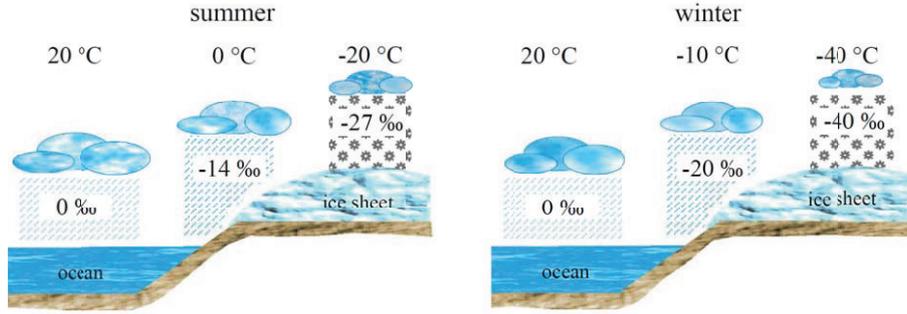


Figure 2.1: For colder temperatures (winter) less water vapour with stronger depleted heavy isotopes remains in the air over a polar ice sheet than during summer [Dansgaard, 1964]. The fractionation factor for water between the liquid and vapour phase is given as a function of temperature (Table 2.4).

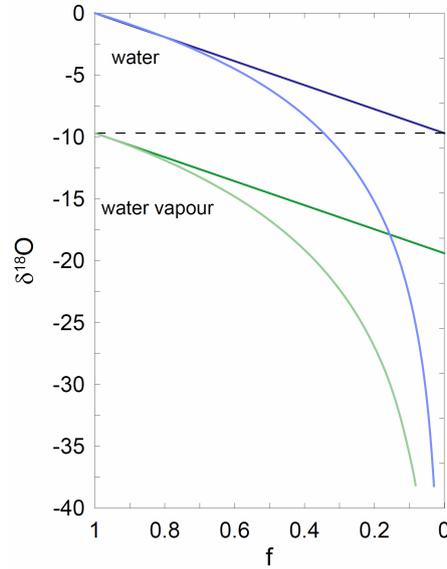


Figure 2.2: $\delta^{18}\text{O}$ values after Rayleigh condensation of the water and water vapour during the formation of a cloud (closed system, straight lines) and during raining from the cloud (open systems, curved lines).

Isobaric $p_1 = p_0$	dry adiabatic $\Delta Q = 0$
$\frac{V_1}{T_1} = \frac{V_0}{T_0}$	$p_1 V_1^\kappa = p_0 V_0^\kappa$ with $\kappa = \frac{c_p}{c_v} \approx \frac{7/2}{5/2} = 1.4$
$V_1 = V_0 \frac{T_1}{T_0} = 0.932 V_0$	$V_1 = V_0 \left(\frac{T_0}{T_1}\right)^{\frac{1}{\kappa-1}} = V_0 \left(\frac{T_0}{T_1}\right)^{2.5} = 1.193 V_0$
$\frac{n_1}{n_0} = \frac{a_s(0^\circ\text{C})}{a_s(20^\circ\text{C})} = 0.280$	$\alpha_{V \rightarrow W, m} = \frac{1.0098 + 1.0117}{2} = 1.0108$
$\frac{n'_1}{n'_0} = \frac{n_1}{n_0} 0.932 = 0.261$	$\frac{n'_1}{n'_0} = \frac{n_1}{n_0} 1.193 = 0.334$
$R_V = \frac{R_0}{\alpha_{V \rightarrow W}(20^\circ\text{C})} = 0.9903 R_0$; $R'_W = R'_V(0^\circ\text{C}) \alpha_{V \rightarrow W}(0^\circ\text{C})$ $R'_V = R_V \left(\frac{n'_1}{n'_0}\right)^{\alpha_{V \rightarrow W, m} - 1}$ $R'_W(0^\circ\text{C}) = 0.9903 R_0 \left(\frac{n'_1}{n'_0}\right)^{0.0108} \cdot 1.0117$	
$R'_W = 0.987 R_0$	$R'_W = 0.990 R_0$
$\delta^{18}\text{O} = -12.5\text{‰}$	$\delta^{18}\text{O} = -9.9\text{‰}$
$\frac{\Delta \delta^{18}\text{O}}{\Delta T} = \frac{-12.5\text{‰}}{-20^\circ\text{C}} = 0.625\text{‰ } ^\circ\text{C}^{-1}$	$\frac{\Delta \delta^{18}\text{O}}{\Delta T} = \frac{-9.9\text{‰}}{-20^\circ\text{C}} = 0.495\text{‰ } ^\circ\text{C}^{-1}$

2.3.1 Comparison with observations

Even this simplified model yields $\frac{d(\delta^{18}\text{O})}{dt}$ values between 0.63 ‰/°C and 0.5 ‰/°C. Obviously, no exclusive relationship between the condensation temperature and the isotopic composition of oxygen, $\delta^{18}\text{O}$, of a single precipitation event exists.

However, *Dansgaard* [1964] has found a rather strong empirical relationship between $\delta^{18}\text{O}$ values of precipitation events integrated over a year and corresponding mean annual surface temperatures. His results are given in Fig. 2.3. The linear regression line for coastal regions yields:

$$\delta^{18}\text{O} = 0.69\overline{T}_a(^{\circ}\text{C}) - 13.6\text{‰}. \quad (2.9)$$

Siegenthaler and Oeschger [1980] have compared monthly precipitation with mean monthly temperatures within the Swiss precipitation network. The coefficient for the relationship $\delta^{18}\text{O}$ vs. temperature is lower, with values in the range of 0.35 – 0.55 ‰/°C. In addition, the correlation is much worse compared to annual correlations. Only 50% – 67% of the variability can be explained by temperature variations.

Continental stations have lower $\delta^{18}\text{O}$ values than coastal stations for equal mean temperatures. Therefore, the relationship between $\delta^{18}\text{O}$ values and mean temperature is region specific:

$$\begin{aligned} \text{Greenland ice sheet: } \delta^{18}\text{O} &= 0.67\overline{T}_a(^{\circ}\text{C}) - 13.7\text{‰}, \\ \text{Antarctic ice sheet: } \delta^{18}\text{O} &= 0.76\overline{T}_a(^{\circ}\text{C}) - 7.6\text{‰}. \end{aligned} \quad (2.10)$$

The altitude effect amounts approximately to: $\frac{d(\delta^{18}\text{O})}{dh} = -1.5$ to -3 ‰/km. (for a "lapse rate" of 6°C per km one would actually expect -4.2 ‰/km).

Temperature	over water		over ice	
	e_s (Pa)	a_s (gm ⁻³)	e_s (Pa)	a_s (gm ⁻³)
25°C	3167.2	23.070		
20°C	2309.0	17.319		
15°C	1713.2	12.850		
10°C	1255.4	9.417		
5°C	908.1	6.811		
0°C	647.1	4.854	647.1	4.854
- 5°C	455.3	3.415	433.7	3.253
-10°C	315.2	2.364	286.1	2.146
-15°C	214.8	1.611	185.6	1.392
-20°C			118.4	0.888
-25°C			74.0	0.555
-30°C			45.46	0.341
-40°C			16.00	0.120

Table 2.5: Water vapour pressure e_s and absolute humidity a_s over water and ice.

2.3.2 Relationship between $\delta^{18}\text{O}$ and $\delta^2\text{H}$

^{18}O and ^2H are generally very well correlated in precipitation and ground water according to

$$\delta^2\text{H} = 8\delta^{18}\text{O} + d, \quad d: \text{Deuterium excess.} \quad (2.11)$$

The factor eight in the linear relation can be understood approximately considering

$$\frac{dR_v}{R_v} = (\alpha - 1) \frac{dn}{n} \quad \text{and} \quad R_v = R_{\text{Standard}}(1 + \delta/1000), \quad \text{see exercises,} \quad (2.12)$$

with numbers for 20°C :

$$\frac{dR_{2\text{H}}R_{18\text{O}}}{R_{2\text{H}}dR_{18\text{O}}} = \frac{1000 + \delta^{18}\text{O}}{1000 + \delta^2\text{H}} \frac{d\delta_{2\text{H}}}{d\delta_{2\text{O}}} \approx \frac{\alpha_{2\text{H}} - 1}{\alpha_{18\text{O}} - 1} \approx \frac{85}{9.8} \approx 8.7. \quad (2.13)$$

The mean observed value is somewhat smaller because ^{18}O is more strongly influenced by the diffusion through the laminar boundary layer than ^2H . Similarly, a large part of the variation in d in Eq. 2.11 is also due to diffusion through the laminar boundary layer. The strength is mainly dependent on kinetic effects during evaporation and condensation. However, evaporation is dominant, particularly during snow events (ice caps, ice sheets) in middle latitudes. The effect becomes stronger, the drier the air above the water/ice surface is (higher water vapour deficit), i.e. the faster the evaporation rate gets. The variations in d can be explained by a simple model (film model) and by different diffusion constants for heavier and lighter molecules based on their isotopic compositions. Formally it can be calculated as follows:

$$\alpha_k = f + (1 - f) \left(\frac{D_{\text{HDO}}}{D_{\text{H}_2\text{O}}} \right)^n, \quad \alpha_T = \alpha_k \alpha \quad (2.14)$$

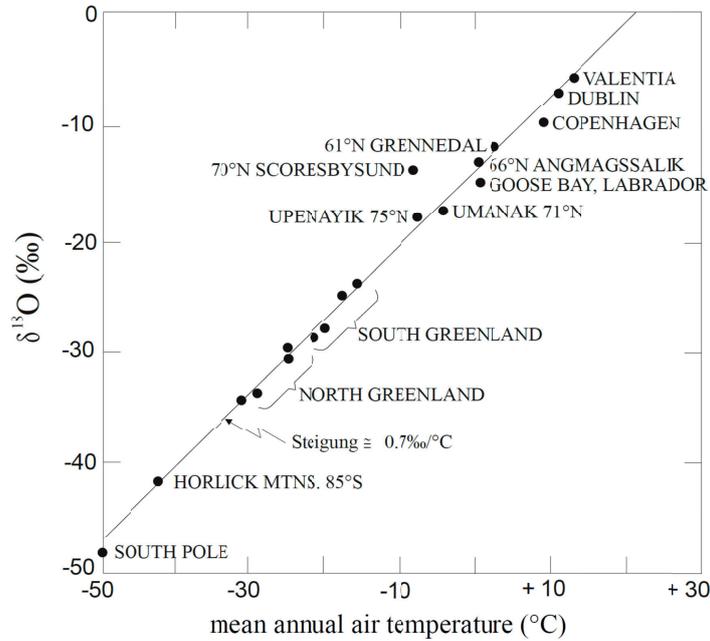


Figure 2.3: Mean of ^{18}O values vs. annual mean surface temperatures [Dansgaard, 1964].

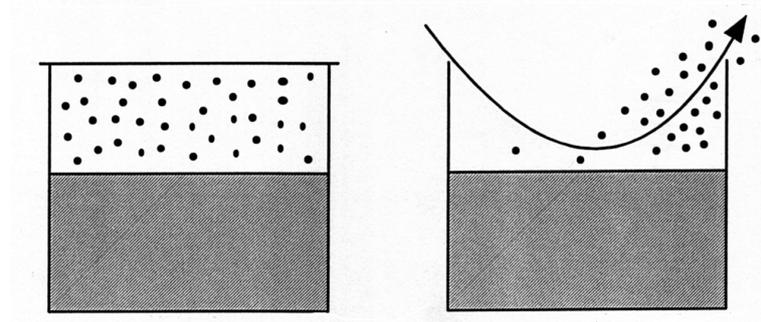


Figure 2.4: Equilibrium versus kinetic fractionation factor between liquid and vapor phase of water. Left: $\delta^{18}\text{O}_{\text{vapour}} = (1/\alpha - 1) \cdot 1000$; right: $\delta^{18}\text{O}_{\text{vapour}} = (1/\alpha_T - 1) \cdot 1000$

where	α_T	Total fractionation factor
	α_k	Kinetic fractionation factor
	α	Equilibrium fractionation factor
	f	Relative humidity (100% = 1!)
	n	Factor of turbulence
	D	Diffusion constants for H_2O and HDO

For Holocene ice values around 8‰ and for ice from the last glacial maximum, values around 4‰ for d were found in Antarctica. It is believed that this difference is due to higher wind speeds of dry air over the water vapour source region.

Chapter 3

Radio nuclides as tracers of environmental processes

3.1 Introduction and overview

Radio nuclides are extremely valuable since they contain time information. Potential candidates are:

Nuclide	T1/2	Decay	Source	Reservoir (application)
³ H (Tritium)	12.3 a	β^-	Bomb tests	(fluorescents)
⁷ Be	53 d	EC, γ	Cosmic rays	Atmosphere
¹⁰ Be	$1.6 \cdot 10^6$ a	β^-	Cosmic rays	(Prod. by CRays)
¹⁴ C	5730 a	β^-	Cosmic rays Bomb tests	(Dating)
⁴⁰ K	$1.28 \cdot 10^9$ a	β^-, β^+, γ	natural	Soil, nutrition
⁶⁰ Co	5.3 a	β^-, γ	artificial	(Industry)
⁸⁵ Kr	10.4 a	β^-	recycling	Atmosphere
⁸⁷ Rb	$4.8 \cdot 10^{10}$ a	β^-	natural	(Dating)
⁹⁰ Sr	28 a	β^-	U-fission	Soil
^{99m} Tc	6 h	γ	artificial	(med. Diagnosis)
¹³¹ I	8 d	β^-, γ	U-fission	(med. Diagnosis)
¹³⁴ Cs	2 a	β^-, γ	n-activation	after Tschernobyl
¹³⁷ Cs	30 a	β^-, γ	U-fission	Soil
²¹⁰ Pb	22 a	β^-, γ	²³⁸ U-decay	(Dating)
²²⁰ Rn	55 s	α	²³² Th-decay	Soil
²²² Rn	3.8 d	α	²³⁸ U-decay	Soil, air
²²⁶ Ra	1600 a	α, γ	²³⁸ U-decay	Soil
²³² Th	$1.4 \cdot 10^{10}$ a	α, γ	natural	Soil
²³⁵ U	$8 \cdot 10^8$ a	α, γ	natural	Soil (air)
²³⁸ U	$4.45 \cdot 10^9$ a	α, γ	natural	Soil
²⁴¹ Am	430 a	α, sf	artificial	(fire alarm)

3.1.1 Overview of reservoirs and processes

Radioactivity in	Statements, conclusions
Air	Propagation, Dilution, Creation
Air - surface occupation	Dry- and wet deposition (section 3.5.1) $v \int C_{\text{air}} dt = C_{\text{soil}}$ (v = deposition velocity) $S = C_{\text{rain}} \rho_{\text{air}} / C_{\text{air}}$ Scavenging ratio
Soil	
Time records	Concentration records
Soil-Depth profile	Penetration condition; $A = A_0 e^{-\frac{z}{z_0}}$
Biosphere	CO ₂ cycle: See section 4.
Air - Soil Plants	Transfer functions (section 3.5)
Grass - Milk, Hay Milk	Transfer functions (section 3.5)
Grass Meat	Transfer functions (section 3.5)
Lake water - Plankton Fish	Transfer functions (section 3.5)
Activity in animal and humans	Transfer functions (section 3.5)
	Dose factors
Aquatic systems	
Ocean, lakes	Mixture, circulation, ageing, sedimentation
Groundwater	Rock-water interaction Mixture, circulation, ageing Paleo waters

3.2 Production of radio nuclides

3.2.1 Production in the atmosphere

Overview

Nuclide	Sources	Sinks, Processes
^{37}Ar	Production in stratosphere through cosmic rays	Radioactive decay, Vertical mixing
^7Be , ^{10}Be , ^{36}Cl , ^3H , (^{14}C)	Production in stratosphere through cosmic rays	Vertical mixing and deposition (+ exchange with the ocean)
^3H , ^{14}C , ^{90}Sr , ^{137}Cs	Atomic bomb tests Northern hemisphere Cosmic radiation	Exchange Strat. \rightarrow Trop. - " - North \rightarrow South - " - with ocean
^{37}Ar	Atomic bomb tests	Mixture in the Northern Hemisphere
^{85}Kr	Nuclear reprocessing in Northern hemisphere	Vertical mixing Exchange North \rightarrow South Horizontal distribution
^{222}Rn	Soil of continents Uranium decay	Vertical mixing from below \rightarrow top, oceanic clean air mixes with air over continents
^{14}C	Emissions from atomic power plants	Distribution on local to regional scales
radioactive noble gases ^{131}I , ^{137}Cs	Emissions from atomic power plants (e.g. Tschernobyl)	Distribution on local to regional scales and deposition

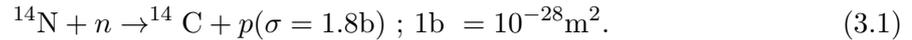
Production through cosmic rays

Through cosmic radiation, in particular through neutrons (as secondary products of nuclide reactions), isotopes are readily produced in the atmosphere. The production rate depends mainly on the intensity of the galactic cosmic radiation (GCR) that is modulated by the magnetic fields of the sun and the Earth. Protons with energies below about 10 GeV are shielded by the solar magnetic field. The solar magnetic field varies with the number of sun spots and leads to a corresponding variation of the intensity of the cosmic radiation and therefore also of the isotope production. The GCR consists mainly of protons (p , 90%), helium nuclides (α , 9% and heavy nuclides (1%). Such a spectrum can only be observed in polar stratosphere regions, at other latitudes the Earth's magnetic field shields charged particles, which are below a certain threshold value, from penetrating. This threshold or cut-off is dependent on the geomagnetic latitude, the geographic length as well as on the angle of incidence of the radiation. Because of these dependencies the composition and the energies of the secondary GCR are dependent on latitudes and height. The increase in the

uppermost part of the atmosphere corresponds to the build-up of secondary radiation, namely neutrons, from the primary cosmic rays. At lower latitudes the neutron flux is decreasing due to absorption and/or reactions of neutrons in the atmosphere. Production rates of nuclides are therefore height- and latitude dependent (Figs. 3.1 and 3.2). In particular the secondary neutrons (n) are of great importance. Their relative contributions to the isotope production in the atmosphere are much larger than those from protons (p), gamma quants (γ) and muons (μ), which are approximately: $n : p : \gamma : \mu = 100 : 10 : 0.5 : 0.003$. The production of nuclides in the atmosphere is mainly based on the following three processes:

- Spallation reactions of high energy particles
- Spallation of neutrons and protons from excited nuclides
- Nuclide reactions initiated by primary and secondary particles

Neutrons have a larger reaction cross section compared to protons. The thermal neutron capture reaction, for example, leads to the production of radiocarbon ^{14}C :



In reality, reaction cross sections are often not well enough constraint. For most nuclides one can safely assume that the production rate in the stratosphere is twice as much as in the troposphere.

Solar cosmic radiation

The solar cosmic rays are also a source of isotope production in the atmosphere, however, to a much smaller degree as GCR. The origins of solar cosmic radiation are the sun energy outbreaks (flares) that happen typically on hourly or daily time scales. The intensity and the energy spectrum vary from flare to flare, which is taken into account in the rigidity factor:

$$J \propto e^{-R/R_0} \text{ with } R_0 \text{ from } 50 \text{ to } 300 \text{ MV} \quad (3.2)$$

Rigidity is a parameter that characterises the stiffness of the particle path. This stiffness is driven by the balance of the Lorentz and Centripetal force in a magnetic field B:

$$qvB = \frac{mv^2}{r}, \quad r = \frac{mv}{qB} = R \frac{1}{cB} \quad (3.3)$$

with R , Rigidity (energy/charge = volts) and c , velocity of light in vacuum.

The proton to alpha particle ratio (p/α) varies between 1 and 30. Since the energies of protons are mostly below the cut-off level, except for polar regions, effects are primarily observed in these regions (polar lights). Solar flare particles lead only rarely to spallation. Nuclides that are produced from neutrons show, however, the well-known 11-year cycle of the sun (e.g. ^{14}C). An especially strong flare took place on the 20.1.2005 (Figs. 3.3 and 3.4).

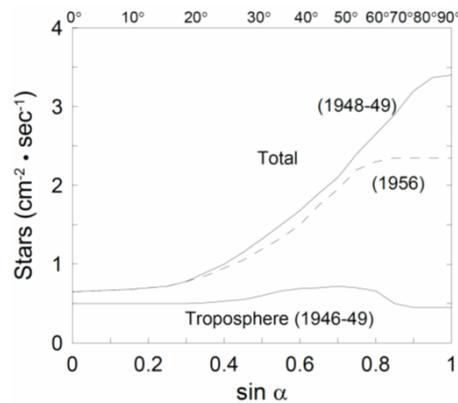


Figure 3.1: Number of nuclear disintegrations (stars) per cm^2 of column and per second in the atmosphere as function of the geomagnetic latitude α . Redrawn from *Lal and Peters* [1967].

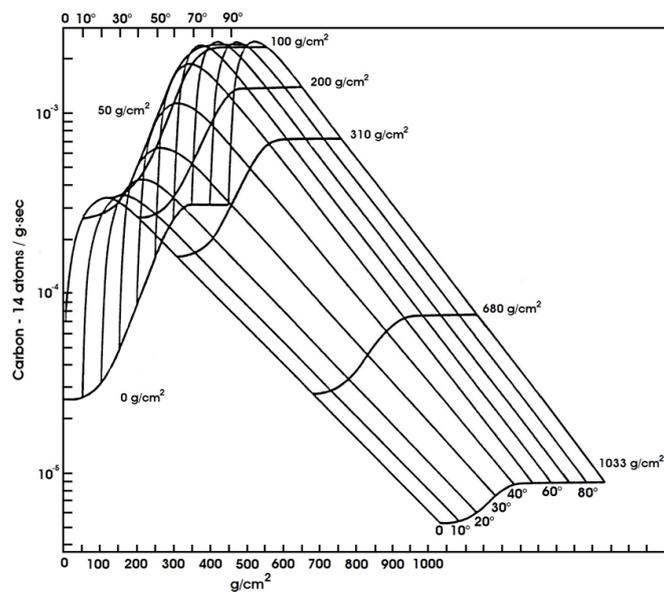


Figure 3.2: The number of ^{14}C nuclides produced per gram of air per second is plotted as a function of atmospheric pressure for various latitudes. The curves have been normalized to a production rate of one neutron/sec cm^2 column at the magnetic poles during sun spot minimum. Redrawn from *Lal and Peters* [1967].

Nuclide	in Troposphere (atoms/cm ² s)	Total (atoms/cm ² s)	Inventory on Earth
¹⁰ Be	$1.5 \cdot 10^{-2}$	$4.5 \cdot 10^{-2}$	260 tons
²⁶ Al	$3.8 \cdot 10^{-5}$	$1.4 \cdot 10^{-4}$	1.1 tons
⁸¹ Kr	$5.2 \cdot 10^{-7}$	$1.2 \cdot 10^{-5}$	8.5 kg
³⁶ Cl	$4 \cdot 10^{-4}$	$1.1 \cdot 10^{-3}$	15 tons
¹⁴ C	1.1	2.5	75 tons
³⁹ Ar	$4.5 \cdot 10^{-3}$	$1.3 \cdot 10^{-2}$	52 kg
³² Si	$5.4 \cdot 10^{-5}$	$1.6 \cdot 10^{-4}$	0.3 kg
³ H	$8.4 \cdot 10^{-2}$	0.25	3.5 kg
⁷ Be	$2.7 \cdot 10^{-2}$	$8.1 \cdot 10^{-2}$	3.2 g

Table 3.1: Selection of production rates of natural nuclides [Turekian and Graustein, 2003].

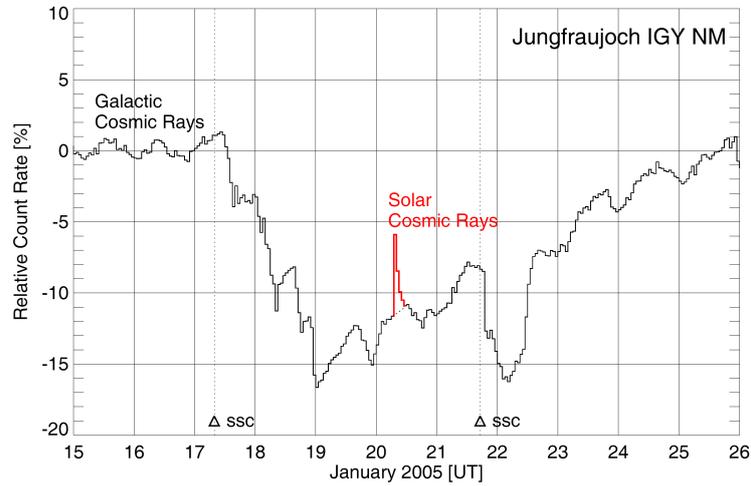


Figure 3.3: Sun eruption as measured by neutron monitors on Jungfraujoch on the January 20, 2005 [Bütikofer et al., 2008].

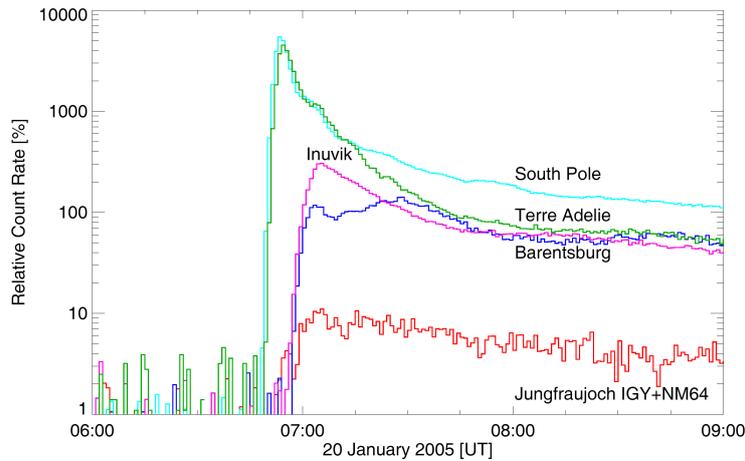


Figure 3.4: Sun eruption on the 20.1. 2005. Comparison of observations at different neutron monitor stations worldwide [Bütikofer et al., 2008].

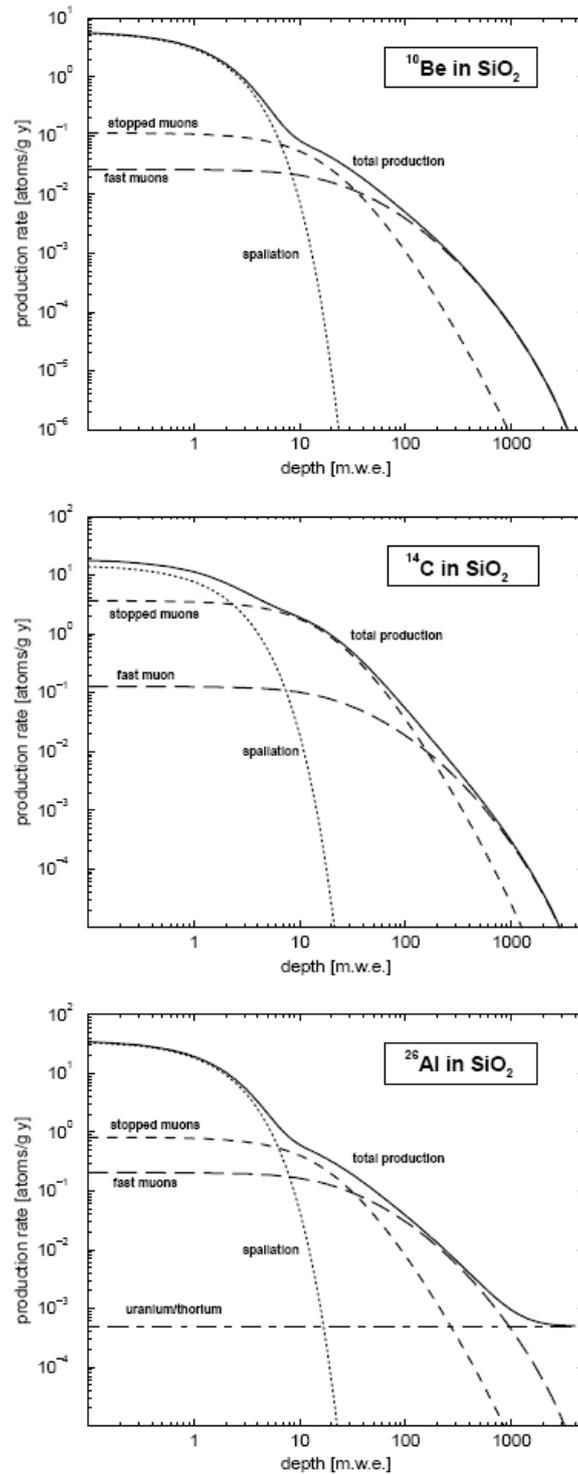


Figure 3.5: Production rates of ^{10}Be , ^{14}C and ^{26}Al in silicium oxide [Heisinger and Nolte, 2000]. Deeper down in the soil only neutrinos are still present from the cosmic ray particles that hardly produce any nuclides due to their small reaction cross section.

3.3 Subsurface production

3.3.1 Secondary particles in the soil

Secondary particles produced by cosmic rays penetrate a couple of meters into the soil and produce there new nuclides. Since the soil composition looks different than that of the atmosphere other nuclides are produced, e.g. more ^{37}Ar , e.g. (reaction $^{40}\text{Ar}(n, 4n)^{37}\text{Ar}$) in the atmosphere versus ($^{40}\text{Ca}(n, \alpha)^{37}\text{Ar}$) in the lithosphere.

3.3.2 Radon and its daughter products

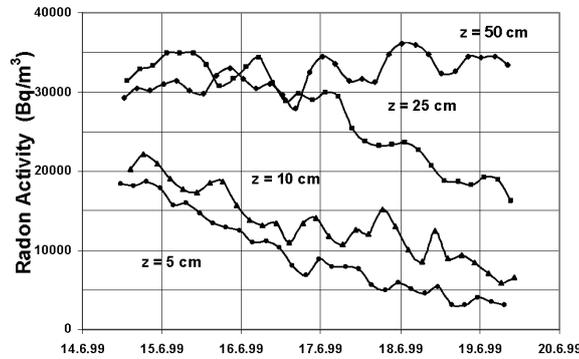


Figure 3.6: Radon activity for four depths during six days [Lehmann *et al.*, 2000].

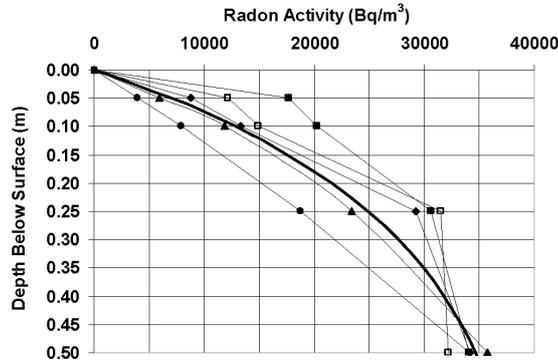


Figure 3.7: Radon profiles for five depths (during midnight from Fig. 3.6) compared to a mean exponential profile [Lehmann *et al.*, 2000].

3.3.3 Uranium and thorium

Decay (and fission) products of uranium and thorium and n -induced reaction products are sketched in the Fig. 3.8 below [Andrews *et al.*, 1989]. This subsurface production of diverse nuclides can be used for dating purposes as well as studying rock-water and water-atmosphere interactions. Particularly worthwhile are noble gas nuclides such as ^3H , ^3He and ^{40}Ar in this respect. Applied to paleo-waters, noble gases can be used for temperature reconstructions [Aeschbach-Hertig *et al.*, 2000].

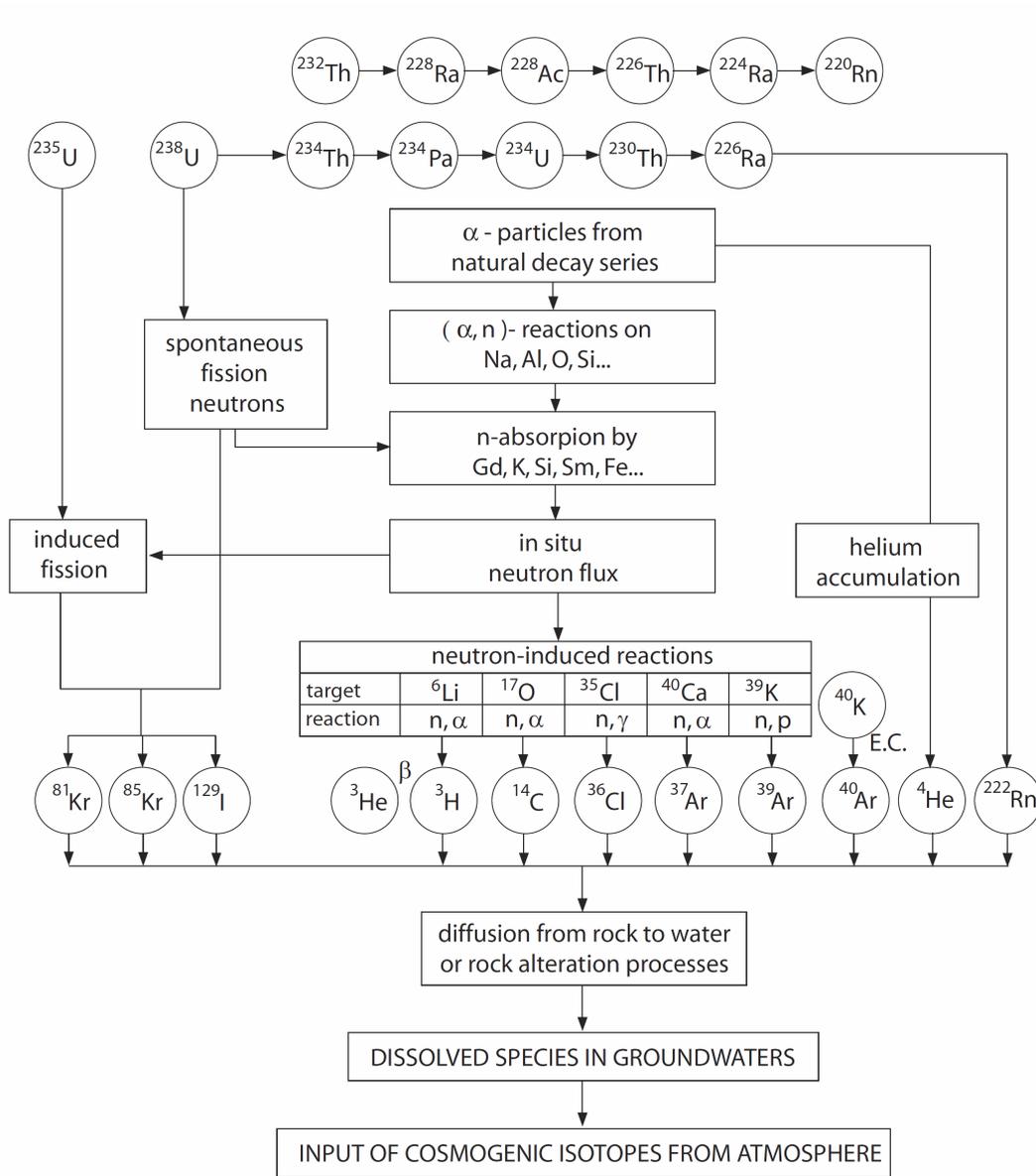


Figure 3.8: Subsurface production paths of stable and radioactive nuclides.

3.3.4 Bomb induced isotopes

Through nuclear bomb tests in particular thermonuclear explosions (H-bombs = fusion bombs since 1953, longest test series: USA/UdSSR 1961/62), large amounts of radioactive nuclides were produced. With the largest bombs, most of these products were injected into the stratosphere from which they were gradually mixed downwards into the troposphere and from there into other natural reservoirs, i.e. the biosphere, the ocean, and the soil.

3.4 Concentration of radioactive nuclides in the atmosphere

3.4.1 Range of activity of radionuclides

Element	Range of radioactivity in air
^{222}Rn	1 - 20 Bq/m ³
^{210}Pb	0.5 - 2 mBq/m ³
^7Be	2 - 7 mBq/m ³
^{37}Ar	0.5 - 4 mBq/m ³
^{85}Kr	1 Bq/m ³
^{137}Cs (Tschernobyl)	up to 1 Bq/m ³
^{137}Cs (today)	0,1 - 10 $\mu\text{Bq/m}^3$
^{14}C	50 mBq/m ³

3.4.2 Temporal variation of ^{90}Sr due to bomb tests

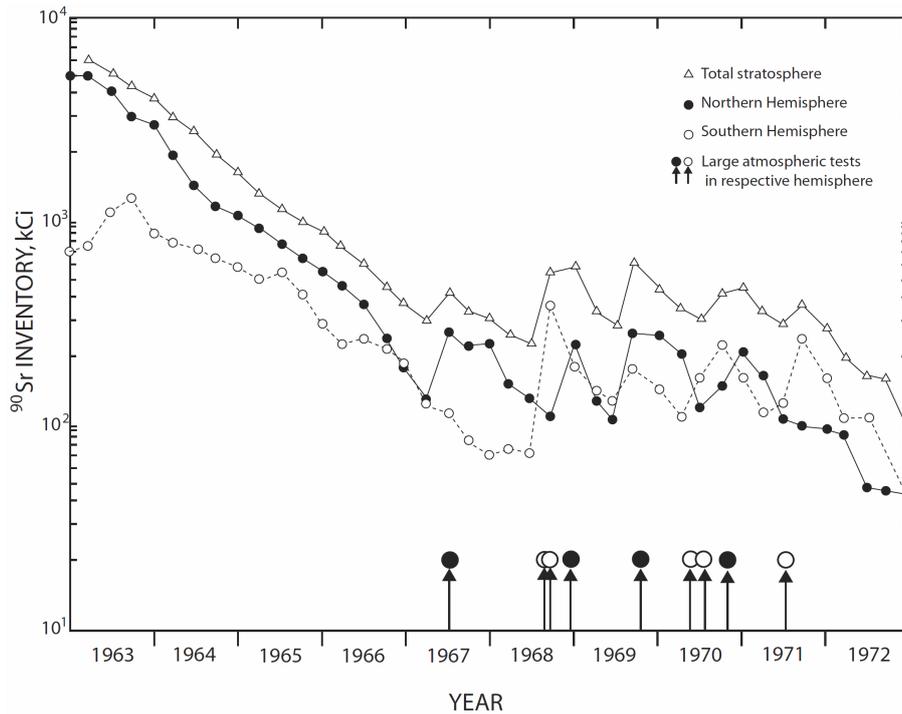


Figure 3.9: ^{90}Sr amount in the stratosphere; NH, SH as well as the total stratosphere. From Reiter [1978].

As mentioned above, bomb tests lead to a huge accumulation of radionuclides in the atmosphere. Gradually they were mixed between the different levels of the atmosphere and they migrated into reservoirs such as the biosphere, the ocean and the lithosphere. Fission products (Examples: ^{90}Sr , $T_{1/2} = 28.6$ a; ^{137}Cs , $T_{1/2} = 30$ a) adsorbed mainly on aerosols and they were deposited as radioactive fallout on the Earth's surface within a few years. Figure 3.9 shows the stratospheric ^{90}Sr amount as function of time.

The following can be learnt:

- Firstly, the inventory in the northern hemispheric (NH) stratosphere is significantly larger than in the southern hemispheric (SH) stratosphere, because most of the large tests in 1961/62 took place in the NH. Within a few years, concentrations were equilibrated between both hemispheres through interhemispheric air exchange.
- From 1963 - 1967 (test-moratorium: no tests) the ^{90}Sr amount decreased in the whole stratosphere exponentially, with an e -folding time of $\tau = 1.2$ a. This decrease cannot be explained by the radioactive decay: $T_{1/2}(^{90}\text{Sr}) = 28$ a, therefore mean life time of $\tau_{ra} = T_{1/2}/\ln(2) = 40$ a. The cause for this decrease is the transport of air from the stratosphere into the troposphere through which ^{90}Sr is rapidly deposited on the surface. The deposition occurs as wet deposition (rain events, snow events) as well as dry deposition in a ratio of 5:1. The decrease rate of the stratospheric reservoir is given by the exchange flux, F , between the stratosphere and the troposphere. F can be estimated from the ^{90}Sr -budget for the stratosphere.

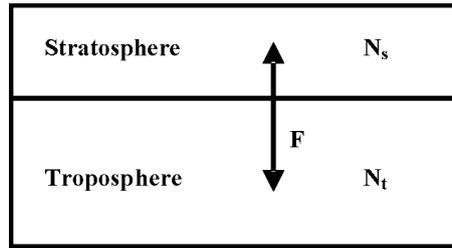


Figure 3.10: Simple box model for the temporal characterization of the strontium content in the atmosphere as measured (Fig. 3.9).

Symbols: N_s, N_t Air mass in the stratosphere and troposphere, respectively; $N_t \approx 4N_s$
 c_s, c_t ^{90}Sr mixing ratio $\frac{\text{amount } ^{90}\text{Sr}}{\text{kg air}}$

Stratospheric ^{90}Sr budget, assumption: both reservoirs are well mixed.

$$\begin{aligned}
 \frac{d}{dt}N_s c_s &= F(c_t - c_s) \approx -F c_s (c_t \approx 0) \\
 N_s \frac{dc_s}{dt} &= -F c_s \\
 \frac{dc_s}{dt} + \frac{F}{N_s} c_s &= \frac{dc_s}{dt} + \frac{1}{\tau_s} c_s = 0 \\
 c_s &= c_{s,0} e^{-\frac{t}{\tau_s}}, \tag{3.4}
 \end{aligned}$$

where

$$\tau_s = \frac{N_s}{F}, \tag{3.5}$$

corresponds to the residence time of air in the stratosphere. The solution of the above mentioned differential equation is proportional to $e^{-t/\tau}$, as observed (Fig. 3.15). From observation

we obtain $\tau = 1.2$ a, therefore the exchange flux F between the stratosphere and the troposphere corresponds to

$$F = \frac{N_s}{\tau_s} \approx N_s \cdot 0.8 \text{ yr}^{-1} = N_t \cdot 0.2 \text{ yr}^{-1}. \quad (3.6)$$

3.4.3 ^{85}Kr emitted from nuclear reprocessing systems

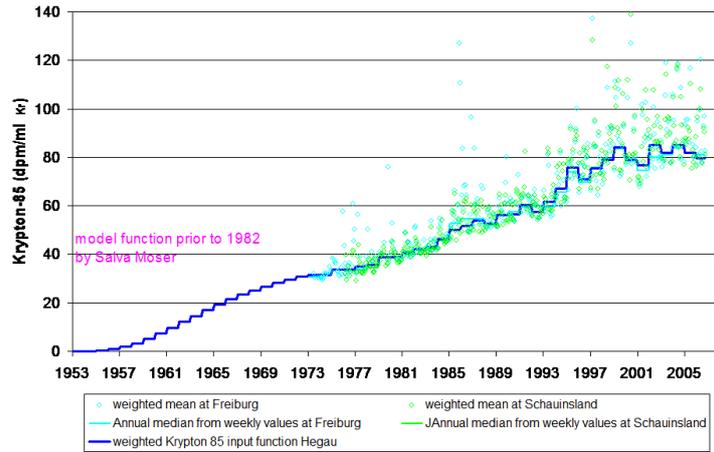


Figure 3.11: ^{85}Kr concentration in air samples taken at Freiburg and Schausinsland.

3.4.4 Temporal variation of isotope abundances

In tree rings, sediments and ice cores temporal evolutions of isotope abundances can be reconstructed. The question is whether such analyses allow retrieving information about changes in the behaviour of natural processes. Possible causes for isotope variations are for example:

- Variable magnetic field of the Earth and sun (temporal and/or spatial variability)
- Variable circulation in the atmosphere and in the oceans.

For example, measured ^{14}C variations on tree rings covering the last 10,000 years are shown in Fig. 3.13. $\Delta^{14}\text{C}$ notation: Eq. 3.22. The small but rapid variations are assigned to a changing magnetic field of the sun that influences the production rate of ^{14}C (magnetic fields lead to a change in charged particle deflection, e.g. protons, which produce secondary particles such as neutrons that are the direct interaction partners for the production of ^{14}C). Not fully understood is the long term trend. It documents the fact that the ^{14}C production was about 10% larger than today. The major reasons assigned to this trend are on the one hand changes in the Earth's geomagnetic field or on the other hand variations in ocean circulation. The blue line displays the mean of many experimentally derived $\Delta^{14}\text{C}$ whereas the light blue lines the one sigma deviations of it.

A comparison between ^{14}C und ^{10}Be variations is given in Fig. 3.14. It was observed that both nuclides show similar trends during the Maunder Minimum (1645 - 1715), a time period

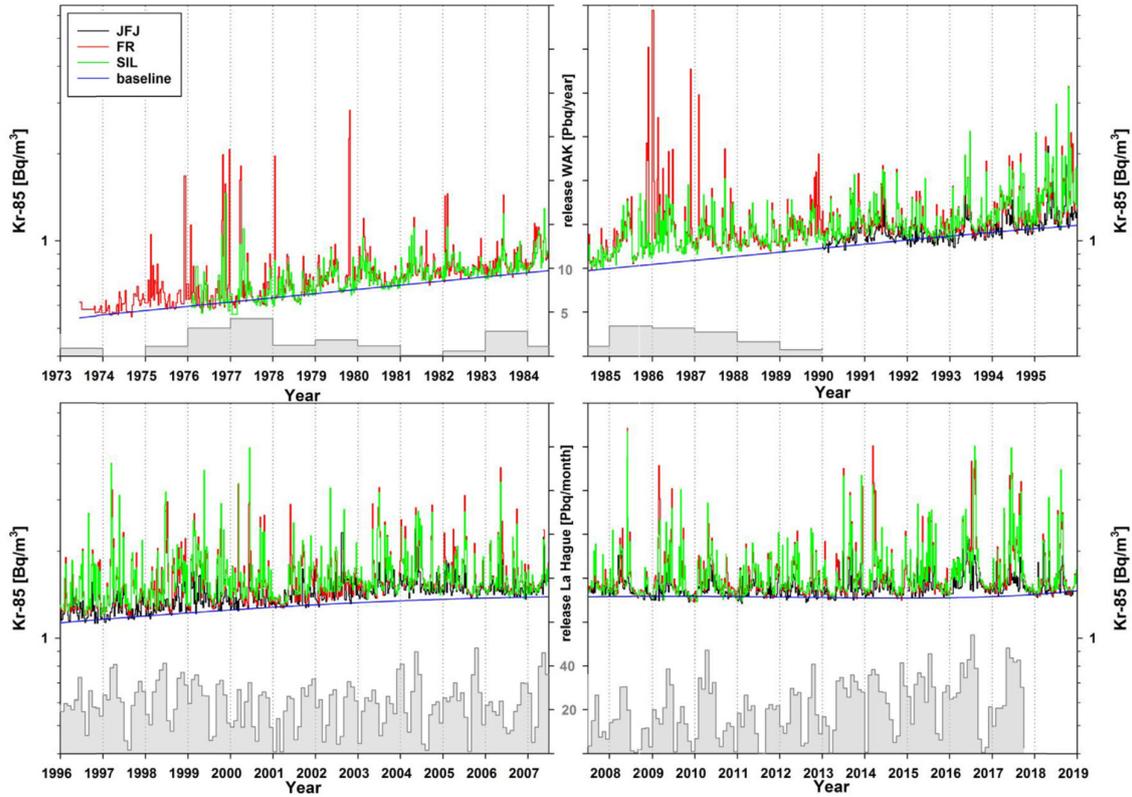


Figure 3.12: Krypton-85 activity concentrations (logarithmic scale) measured in air in weekly samples collected in Freiburg (from 1973), Schauinsland (from 1976) and Jungfrauoch (from 1990). The step graph in the top two panels shows the annual emissions (in PBq per year) from the WAK from 1973 to 1990 to illustrate the influence of local emissions on ^{85}Kr measured in Freiburg. The bottom two panels show the monthly emissions from La Hague (in PBq per month). No emission data were available for La Hague from mid-2017 [Bollhöfer *et al.*, 2019].

of minimal sun spot activity. This similarity is not obvious because a production rate change is immediately reported by ^{10}Be due to its short residence time of only 1 to 2 years in the stratosphere and the fact that aerosols on which ^{10}Be gets attached are washed out within weeks in the troposphere. In contrast, ^{14}C variations are damped since they interact with several gas species of the carbon cycle in the atmosphere. However, when these interactions are taken into account by using carbon cycle models, then calculated ^{14}C concentration agrees quite nicely with corresponding measurements.

Temporal records given in Fig. 3.15 correspond to a much shorter time scale, namely the last decades. Schematically plotted are concentration variations of artificial and natural radionuclides such as ^3H , ^{14}C , ^{39}Ar , ^{85}Kr . These time series are often used for dating purposes of young ground waters or ocean waters. Additionally, they also allow us to calibrate environmental processes during the last decades, i.e. through the penetration of bomb ^{14}C into the Pacific and Atlantic oceans.

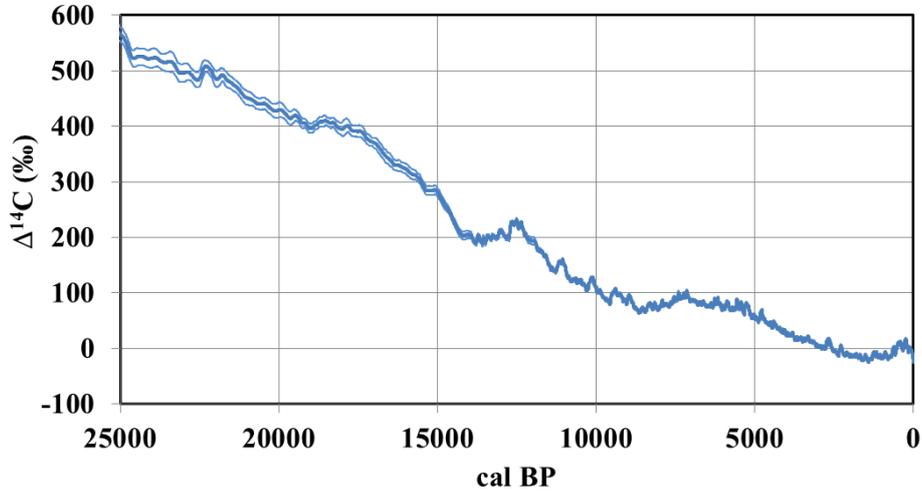


Figure 3.13: Atmospheric $\Delta^{14}\text{C}$ values from tree rings. From *Reimer et al.* [2013], supplementary data.

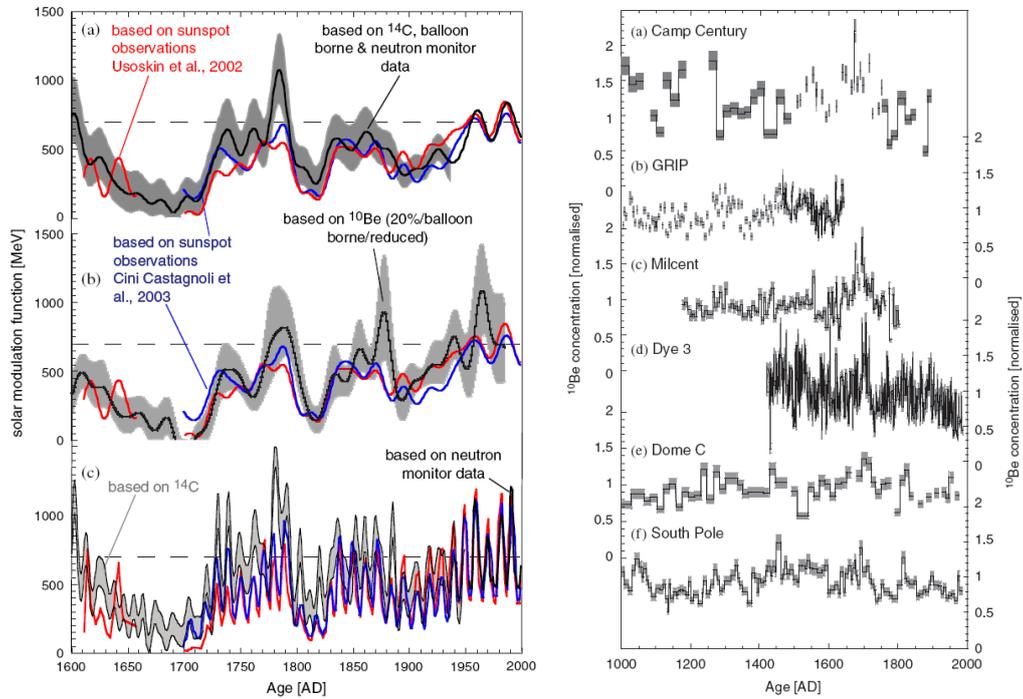


Figure 3.14: Reconstruction of the solar modulation function using different methods (left). ^{10}Be measurements on ice cores (right). From *Muscheler et al.* [2007].

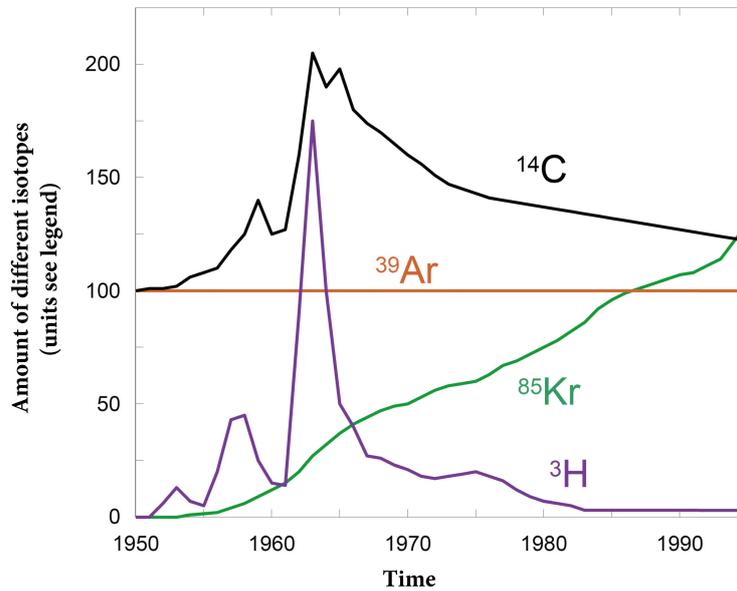


Figure 3.15: Temporal evolution of the activities of different radionuclides in the atmosphere. The vertical scaling, 100 correspond to 1 Bq/m^3 for ^{85}Kr ; 100% modern for ^{39}Ar (corresponding to $1.67 \cdot 10^{-2} \text{ Bq m}^{-3} \text{ air}$), 100 pmC for ^{14}C and 1000 TU for ^3H . pmC: percent modern carbon, TU: Tritium unit, 1 TU = 1 tritium atom per 10^{18} hydrogen atoms = 0.118 Bq/L

3.5 Transfer functions of radio nuclides

3.5.1 Dry and wet deposition

Deposition rates can be investigated by:

- Radon daughters
- Nuclides that are produced by cosmic rays (^7Be , ^{10}Be , ^{36}Cl , etc.)
- Activity emitted by the Tschernobyl and Fukushima reactor incidence
- Chemical tracers

Main parameters of interest in this respect are:

- Particle size
- Meteorological conditions (wind velocity, rain intensity, condition of soil surface, etc.)

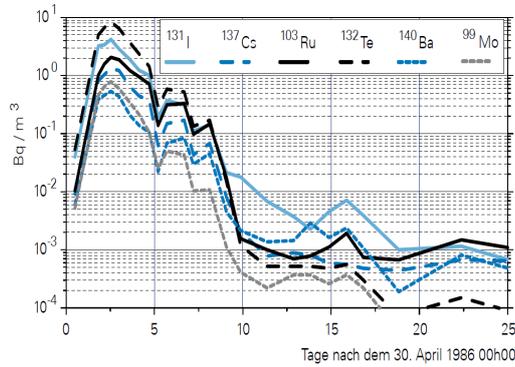


Figure 3.16: Radioactivity on air filter in Switzerland [BAG, 2006].

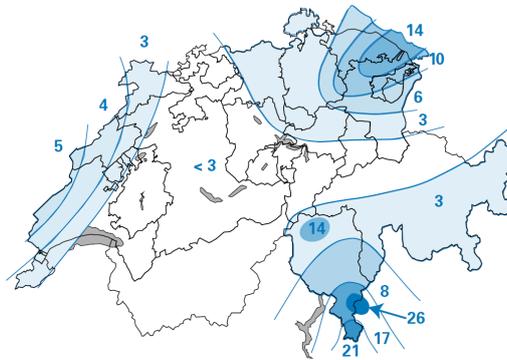


Figure 3.17: Deposition of ^{137}Cs in kBq/m^2 [BAG, 2006].

Dry deposition

$$C_{\text{Soil}} = v \int C_{\text{air}} dt \quad \text{note the unit change: } \text{Bq}/\text{m}^2 = \text{m}/\text{s} \cdot \text{Bq}/\text{m}^3 \cdot \text{s} \quad (3.7)$$

This process could be studied due to the emissions of ^{137}Cs and $^{131}\text{I}_{\text{total}}$ at the Tschernobyl event [Santschi, 1987]. The dry deposition velocities are in the range of 0.03 to 0.15 cm/s with a mean value of 0.05 ± 0.01 cm/s for clear sky conditions. For foggy situations the velocity increases roughly by a factor of four.

Wet deposition

Deposition velocities $^{137}\text{Cs} \approx 10$ cm/s.

Relevant parameters

- Number and size of rain drops.
- Number and size distribution of aerosol particles.
- Falling velocity of rain drops.
- Rain intensity.

Scavenging ratio

$$S = \frac{C_{\text{rain}}}{C_{\text{air}}} \rho_{\text{air}} \quad (3.8)$$

$$S = \frac{\text{Bq/kg}}{\text{Bq/m}^3} \cdot \frac{\text{kg}}{\text{m}^3} = 1 \quad (3.9)$$

Santschi [1987] found: $S_{137\text{Cs}} = 371 \pm 20$
 $S_{103\text{Ru}} = 327 \pm 15$
 $S_{132\text{Te}} = 412 \pm 50$

3.5.2 Penetration of nuclides into the soil, or sedimentation

From the radioactivity deposited onto the surface (e.g. from Tschernobyl) only a small percentage is transported by surface waters (range 1 to 10%). Nevertheless, the radioactivity at the surface is slowly decreasing since nuclides are transported to deeper depths. This depth distribution can be formulated as follows for the first years after Tschernobyl:

$$A = A_0 e^{-z/z_0}. \quad (3.10)$$

The mean penetration depth z_0 corresponds today to about 30 cm for the bomb tests fallout and only a couple of centimeters for Tschernobyl. z_0 is a function of soil composition and the chemical presence of the radionuclide.

- The sedimentation rate depends on particle size and the chemical presence of the radionuclide, see corresponding literature.
- Sediment profiles lead to estimates of sedimentation rates and information about relevant processes within the sediment (e.g. Lake Constance 0.06 – 0.1 g/cm²yr).
- At the depth of 25 mm ¹⁰⁶Ru/¹³⁷Cs and ¹²⁵Sb/¹³⁷Cs ratios were larger compared to the depth range of 14 – 22 mm. Mean ¹⁰⁶Ru and ¹²⁵Sb have higher sedimentation rates than ¹³⁷Cs.

3.5.3 Transfer functions

The transit from a medium 1 to a medium 2 is generally described by a transfer function (*TF*):

$$TF = \frac{A_1}{A_2}. \quad (3.11)$$

Transfer factors were mainly determined based on the bomb fallout and incident at Tschernobyl. Figs. 3.19 to 3.21 can be used to calculate such transfer functions.

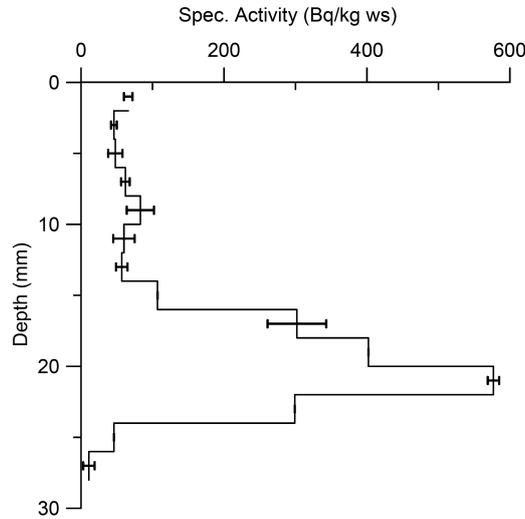


Figure 3.18: Depth distribution of ^{137}Cs radioactivity from the Tschernobyl fallout in Bq/kg wet substance in a sediment core taken in April 1989 from 140 m depth, Lake Constance. Redrawn after *Lindner et al.* [1989].

3.6 Radio nuclides for dating and process studies

For dating, any process with regular time dependence can be used. Mostly used are the decay of radio nuclides and the counting of seasonal variations (tree rings, sediments (varves), ice cores). Other time information will also be discussed in the following. For instance flow velocities, temporal process behaviour and product formation can only be given if time information is available. Since dating methods can be applied in different fields, they are of interdisciplinary nature, in particular radiocarbon dating is well known by several communities. Therefore, a separate section 3.6.2 is dedicated to radiocarbon.

Field	Science Problem	Method (Example)
Geochronology	Age of stars, meteorites, Earth of geological formations	$^{40}\text{K} - ^{40}\text{Ar}$ (Ar/Ar) $^{87}\text{Rb} - ^{87}\text{Sr}$ $^{187}\text{Re} - ^{187}\text{Os}$ $^{238}\text{U} - ^{206}\text{Pb}$ $^{235}\text{U} - ^{207}\text{Pb}$ $^{232}\text{Th} - ^{208}\text{Pb}$
Biology	Age of organic substances, corals, ceramic	^{14}C $^{232}\text{U}/^{230}\text{Th}$ fission track
Ocean	Age of water masses	^3H , ^{85}Kr , ^{39}Ar , ^{14}C
Hydrology	Age of ground waters	^{222}Rn , ^3H , ^{84}Kr , ^{39}Ar , ^{14}C
Glaciology	ice, climate processes	stable isotopes, i.e. ^{81}Kr , ^{36}Cl (?), ^{10}Be

Age calculations are mostly dependent on prior assumptions, for instance assumed constant

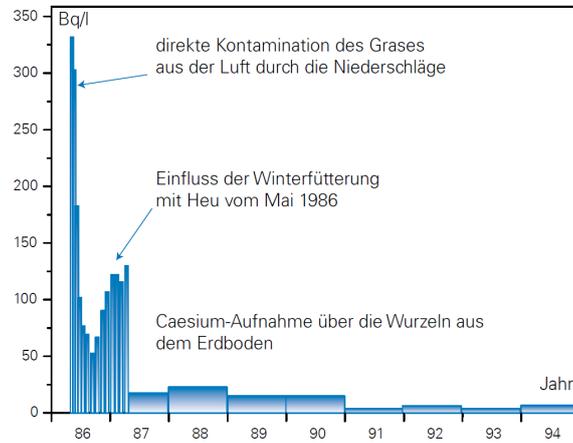


Figure 3.19: Activity measurements of ^{137}Cs in cow milk in the canton of Ticino [BAG, 2006].

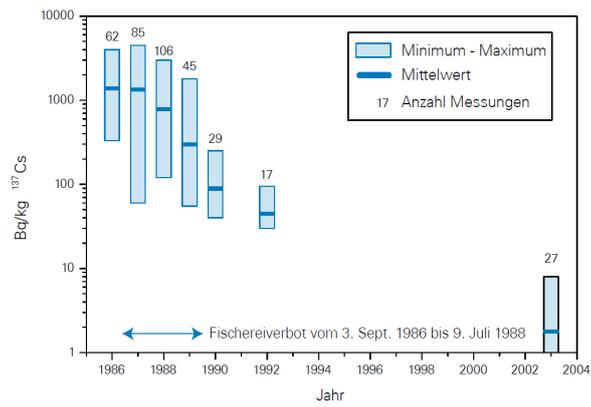


Figure 3.20: Specific activity of ^{137}Cs in fishes from the Luganersee [BAG, 2006].

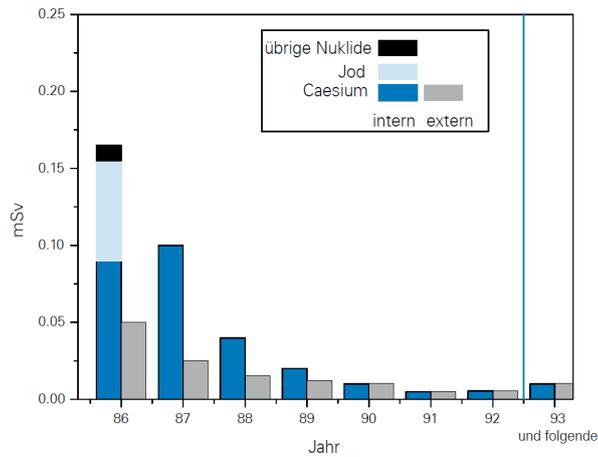


Figure 3.21: Mean dose for Swiss citizens due to Tschernobyl [BAG, 2006].

production rates or stable exchange rates etc. These estimated ages may be in disagreement with the real time scale, therefore, the term model age should be used here. Indication and discussion of boundary conditions used to calculate the model age are as important as the model age value itself. Any result should be handed out with an assigned uncertainty as well as the method used to obtain the result (e.g. dating with ^{14}C or Rb/Sr, etc.). Only with such information in hand does a comparison of results using other methods make sense.

The physical law of radioactive decay is used for instance to retrieve time information. Thereby,

- an initial radioactivity can decrease in a closed system.
- the amount (or activity) of a daughter product can increase in a closed system.
- a radiation source (stable or radioactive) can produce nuclides or generate radiation diseases.

3.6.1 Radioactive decay

An initial activity A_0 is decreasing according to the physical law. If A_0 is known then a single measurement at the time t , $A(t)$ allows to estimate t :

From

$$A(t) = A_0 e^{-\lambda t}, \quad (3.12)$$

it follows that

$$t = \frac{1}{\lambda} \ln \frac{A_0}{A(t)}. \quad (3.13)$$

The radiocarbon (^{14}C) method is best known: t denotes the time during which the initial equilibrium concentration (or activity) produced in the atmosphere through cosmic radiation has decreased to the measured percentage. From this we can already figure out that relative determinations are often used with radioactive methods.

When such an age determination is performed and discussed then the following two points are critical and need to be investigated: a) check whether the initial activity was constant or whether it is allowed to use today's value and b) check whether the system of interest is a closed system or not.

The question regarding the input value is particularly important for nuclides which are produced in the atmosphere by cosmic radiation and which are subsequently incorporated into other Earth systems (e.g. biosphere, ocean etc.). Perturbations during the last couple of decades have mainly been caused by human interactions:

^3H , ^{14}C , ^{32}Si , ^{36}Cl , ^{90}Sr , ^{137}Cs	are bomb products
^{137}Cs	emitted during Tschernobyl incident
^3H , ^{14}C , ^{85}Kr	emitted by nuclear industries

Temporal evolution of input values: See Figs 3.13, 3.14 and 3.15.

Only a few nuclides are still in equilibrium with their production. These are mainly those

that are produced by neutron energies higher than 14.1 MeV (the deuterium – tritium (d,t) reaction in fusion bombs leads to such neutrons). Today, it is assumed that for instance ^{39}Ar and ^{81}Kr concentrations in the atmosphere were not considerably influenced by nuclear tests. On longer time scales the cosmic radiation itself could have varied or transfer processes (transfer functions) from the atmosphere into other reservoirs may have changed their time constants (changing residence times). From investigations of moon samples and from meteorites we know that the cosmic radiation has only changed within a factor of two over the last million years [Geiss *et al.*, 1962]. ^{14}C variations are well known back to about 10,000 years based on a comparison of measured activities with exactly known ages (from tree rings, Fig. 3.13). If we assume a periodic variation of the production rate (e.g. $P(t) = P_0 + P_1 \cos \omega t$), a dampening factor as well as a phase shift for the atmospheric concentration can be calculated ($N = N_0 + N_1 \cos(\omega t + \phi$; [Oeschger, 1980]). Generally the following formula can be used for a noble gas nuclide in the atmosphere.

$$A(t) = \lambda \int_t^\infty P(t - T) e^{-\lambda T} dT. \quad (3.14)$$

When applying Eq. 3.14 the input values are needed. However, if a sample consists of several components (often the case) then we additionally need the age distribution of these components.

As important as the input values is the information about the system itself: e.g. is it a closed system? Is the nuclide of interest conservative, for instance does it follow the water movement or will it be adsorbed somewhere? If nuclides that are dissolved in ground water interact with the surrounding sediment layers, then the age of the water is meaningless for these nuclides. Inactive carbonates or silicates from rock formations can for instance dilute a ^{14}C or ^{32}Si concentration in water. Additionally, radio nuclides can be adsorbed on rock surfaces whereby fractionating the dissolved part or they can be transferred to subsurface dead volumes or gas bubbles and decay there. Furthermore, subsurface production needs to be taken into account since it adds up to the nuclide concentration that was brought in from the atmosphere (it limits or prevents an age determination). A combination of dating methods can be clarifying in regard to questions of conservation or system conditions.

3.6.2 ^{14}C method

Most often used for dating purposes in archaeology and geology are radioactive decay methods. These methods are based on the relationship between the remaining activity of a specific isotope (as part of an initial activity) and the age of the sample. The most important dating isotope is radiocarbon, ^{14}C , with the mass number 14. ^{14}C is produced in the atmosphere through the reaction of nitrogen nuclides with thermal neutrons, which are secondary particles of the cosmic radiation:



After the build-up, ^{14}C in the atmosphere is rapidly oxidised (via ^{14}CO) to $^{14}\text{CO}_2$ and follows the natural carbon cycle (see stable isotopes). Through photosynthesis it is transferred to

plant material. After its incorporation it decays through β -decay into ^{14}N again, with a half life time of $T_{1/2} = 5730$ yr (see below). The production of ^{14}C in a reservoir is mostly negligible, except the atmosphere. Therefore the specific activity (activity per gram carbon) of dead wood etc., is a function of the age T .

The decay law reads (Eq. 3.12)

$$A(T) = A_0 e^{-\lambda T}, \quad (3.16)$$

where A = sample activity, A_0 = initial activity, "100 % modern" and T = sample age.

$$\lambda = \frac{\ln(2)}{T_{1/2}}, \text{ decay constant, for } ^{14}\text{C}: \lambda = 1/8267\text{yr} . \quad (3.17)$$

The age is therefore given by (Eq. 3.13):

$$T = \frac{1}{\lambda} \ln \frac{A_0}{A(T)}. \quad (3.18)$$

Measuring methods

The ^{14}C activity of a sample is generally measured through its radioactive decay. The β^- -particles (electrons) which are produced are for instance detected in a proportional counter. These β^- -particles have relatively small energies ($E_{max} = 158$ keV) and a short range. Therefore the carbon of the sample is converted into a counting gas (e.g. methane or carbon dioxide). Dedicated arrangements lead to a low background of the counter: Lead shielding, in our lab: underground laboratory at a depth of 30 m (shielding of cosmic radiation) built from low radiation concrete. Anti-coincidence to eliminate influential background: energy rich particles from outside are additionally recorded in an outer counter which is placed concentrically to the sample counter in contrast to the actual electron, β^- , produced from ^{14}C decays of the sample that is only detected by the sample counter due to its short range. This anti-coincidence electrical unit eliminates all events which are registered in both counters. ^{14}C activities are also made by liquid scintillation counting.

Besides the radioactive decay counting methods, the acceleration mass spectrometry (AMS) has significantly increased over the last decades to determine the ^{14}C activity. This method requires much less sample material (< 1 mg carbon) since not the decay itself is detected but the ^{14}C atoms themselves. The principle is based on an accelerator (e.g. tandem accelerator) that allows reaching high acceleration energies of several Mega electron volts (MeV). One of the leading AMS laboratories is located at the ETH in Zürich (Institut für Mittelenenergiephysik/PSI). Since 2013 the University of Bern has its own AMS radiocarbon measurement facility located at the Department of Chemistry.

Typical age uncertainties of moderately aged samples ($T < 10,000$ yr) are $\pm 0.5\%$ of the measured activity. From Eq. 3.18 we can calculate the uncertainty as follows:

$$T = -\frac{1}{\lambda} \ln \frac{A}{A_0}, \quad (3.19)$$

differentiation yields

$$\Delta T = -\frac{1}{\lambda} \frac{\Delta A}{A} = 8267(\text{a}) \frac{\Delta A}{A}, \quad (3.20)$$

with $\Delta A/A = 0.005$ we obtain: $\Delta T \approx 41$ a. If a detection limit of 0.5% modern is assumed for ^{14}C measurements ($A/A_0 = 0.005$) then the maximal age that is datable corresponds to $T_{max} = 8267(a) \ln 200 \approx 44,000$ yr.

Conventions

$T_{1/2} = 5730 \pm 40$ yr is the best value for the half life time of ^{14}C . However, for the dating the older "Libby" half life time $T_{1/2} = 5568$ yr is used, which leads to the conventional ^{14}C age. This seems to be an illogical convention but the ^{14}C ages are not absolute: An important assumption of the dating method (constancy of the initial activity) is not exactly fulfilled (Fig. 3.13). Ages are often expressed as "before present (BP)", i.e. before 1950. As standard activity A_0 (100% modern) acts the activity of a given standard material for which the age 0 BP (namely 1950 AD) is assigned. The specific activity of the standard is 13.56 d.p.m. per g C (1 d.p.m. = 1 disintegration per minute), i.e. 226 Bq/kg C.

Isotope fractionation

Due to their different masses, isotopes of the same element behave slightly different for chemical equilibrium, reaction rates and transport processes (diffusion, see stable isotope section). As a consequence, ratio such as $^{13}\text{C}/^{12}\text{C}$ and $^{14}\text{C}/^{12}\text{C}$ (also true without radioactive decay) are not constant in nature. Therefore, the isotope fractionation needs to be accounted for when performing ^{14}C dating. This is achieved with accompanying $^{13}\text{C}/^{12}\text{C}$ ratio measurements.

The ratio $^{14}\text{C}/^{12}\text{C}$ changes twice as much as $^{13}\text{C}/^{12}\text{C}$ since the fractionation is proportional to the mass difference of the considered isotope pair (Eq. 2.5). Isotope fractionation results in age independent changes, which need to be taken into account: Correction by means of normalising the measured $\delta^{13}\text{C}$ values to a unified value of $\delta^{13}\text{C} = -25\text{‰}$ (because wood has a value of $\delta^{13}\text{C} \approx 25\text{‰}$). Therefore for $^{14}\text{R} = ^{14}\text{C}/^{12}\text{C}$:

$$^{14}\text{R}_{corr} = \alpha^{14}\text{R}; \alpha = 1 + \epsilon/1000 = 1 - 2(\delta^{13}\text{C}_{sample} + 25\text{‰})/1000. \quad (3.21)$$

For geophysical applications the following notation is often used:

$$\Delta^{14}\text{C} = \left(\frac{^{14}\text{R}_{corr}}{^{14}\text{R}_0} - 1 \right) 1000\text{‰}, \quad (3.22)$$

where $^{14}\text{R}_0 = 0.95^{14}\text{R}$ (oxalic acid standard, decay corrected to 1950). It corresponds to a specific ^{14}C activity of 13.56 d.p.m./g C (1 d.p.m. = 1 disintegration per minute) = 100% modern.

Basic assumption for the ^{14}C dating method

1. The initial activity A_0 is well-known. This is equivalent to a constant production rate of radiocarbon in the atmosphere as well as to constant geochemical cycles regarding ^{14}C (e.g. constant distribution of carbon between the atmosphere and the ocean).

2. After the initial uptake of ^{14}C , e.g. by photosynthesis of plants, the sample should represent a closed system. This requires no exchange of carbon with its environment.

To 1: The atmospheric ^{14}C concentration (more precisely: the ratio $^{14}\text{C}/\text{C}$) was not constant in the past. Based on dendrochronologically dated annual tree rings it was possible to reconstruct corresponding changes in radiocarbon over about the last 10,000 years with great precision (Fig. 3.13). Dendrochronology: Dating of wood with the help of variable thickness of annual rings in particular of so called "Weiserjahren". Thereby trees of different age groups are overlapped which leads to a chronology.

Before 7,000 to 10,000 years the ^{14}C concentration was about 10% higher than today. This corresponds to a deviation of the ^{14}C age from the true age of about 800 years. (The Libby half life time shows only a deviation of about 3% from the best estimate). Additionally short term fluctuations ("wiggles") of about 20‰ are observed. To convert ^{14}C ages to absolute ages, radiocarbon needs to be calibrated. Calibration curves are available from tree rings for instance (Fig. 3.13). Most probably, these variations are initiated by changes in the geomagnetic field (long term trend) and by alteration of the sun activity (oscillations within 100 to 200 years).

To 2: A strong indication for the second relationship is given by the isotope beryllium 10 ($T_{1/2} = 1.5 \cdot 10^6$ a). ^{10}Be attaches to aerosols and is removed from the stratosphere within approximately one year, from the troposphere it is washed out within a few weeks. Therefore, ^{10}Be is a good proxy for production variations without significant dampening. The measured ^{10}Be concentrations in an ice core from Dye-3, Greenland, vary antiparallel to the 11 yr cycle of the sun spot number. Radiocarbon does not show such variations because the variations are strongly damped due to the dilution within the atmosphere but also through the exchange with other reservoirs, namely biospheric and oceanic CO_2 . The amplitude of $\Delta^{14}\text{C}$ -variations would correspond to only a few permil according to model estimates.

Today's $\Delta^{14}\text{C}$ concentrations, emissions and radiation

- Atmosphere ca. 0 ‰
- Near boiling water reactors $\leq +200$ ‰
- Production by a nuclear power plant (1 GW) ca. $\approx 10^{11}$ Bq/Jahr
- Natural Dose ca. 13 $\mu\text{Sv}/\text{Jahr}$

Production of ^{14}C by bomb tests

Through the large neutron fluxes ^{14}C was produced by the reaction with ^{14}N in the atmosphere but rapidly oxidised to $^{14}\text{CO}_2$. This temporarily initiated a doubling of the atmospheric ^{14}C concentration in the Northern hemisphere (Fig. 3.23). The decrease in radiocarbon was much slower than observed for ^{90}Sr (Fig. 3.9) because the geochemical behaviour is significantly different. ^{90}Sr is attached to aerosols and precipitated within a short time which is mainly

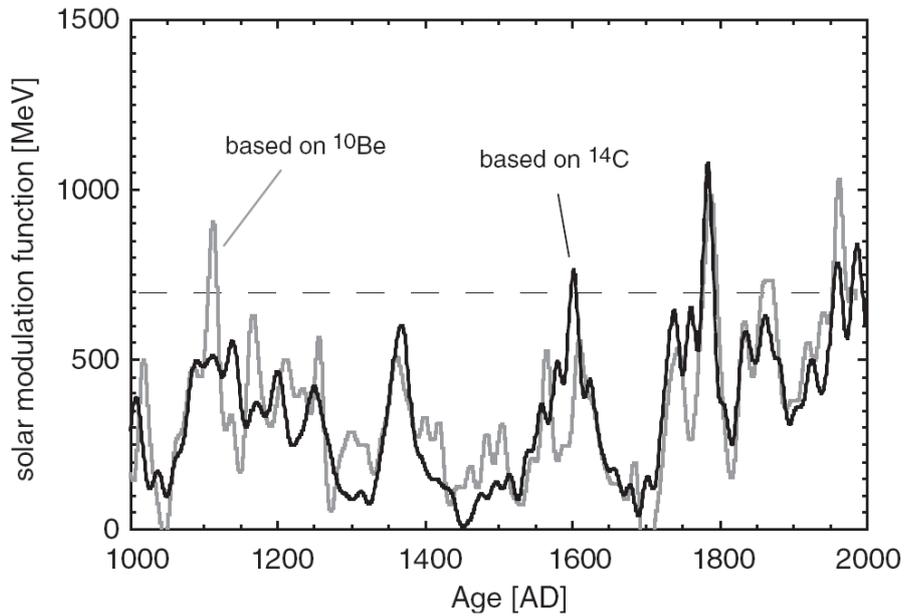


Figure 3.22: Solar modulation function reconstructed based on ^{10}Be and ^{14}C concentrations, respectively [Muscheler *et al.*, 2007].

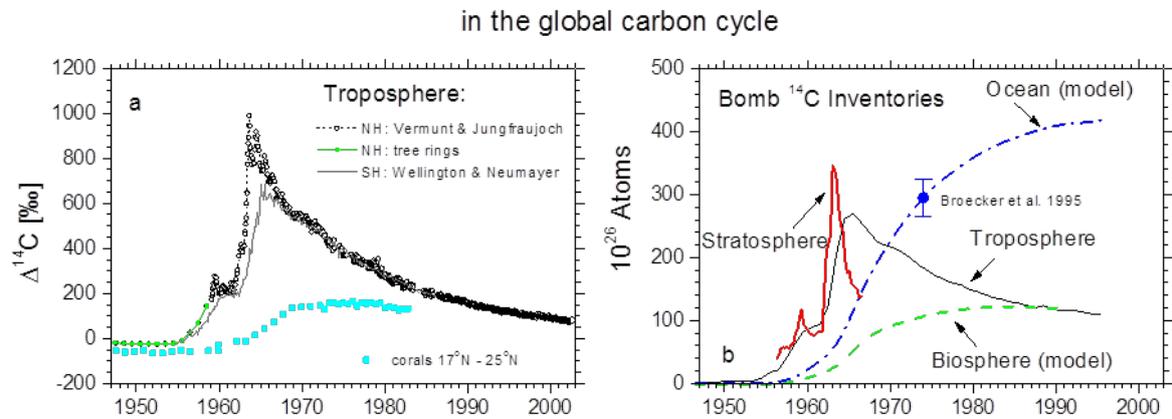


Figure 3.23: $^{14}\text{C}/^{12}\text{C}$ ratio in the troposphere and in ocean surface waters. Note the splitting in two branches in 1965/66: Observations in the Northern and Southern hemisphere; since 1967 practically identical values in both hemispheres [Naegler and Levin, 2006].

governed by the stratosphere troposphere exchange. ^{14}C is incorporated to the carbon cycle and therefore exchanged between the atmosphere, the ocean and the biosphere.

Chapter 4

Biogeochemical Cycles in the Climate System

4.1 Introduction and Motivation

4.1.1 A brief historical perspective

It has become clear that anthropogenic emissions of carbon dioxide (CO_2) and other greenhouse gases such as methane (CH_4) and nitrous oxide (N_2O) cause human induced global warming [*IPCC*, 2007]. This has spurred an ever growing interest in the cycles of carbon and other biogeochemical elements as well as their interaction with the physical climate system. The relevance of research on global warming is emphasized by the 2007 Nobel Prize for Peace awarded to the Intergovernmental Panel on Climate Change (www.ipcc.ch).

The basic greenhouse gas theory was introduced in the beginning of the 19th century by the French mathematician J.-B. Fourier, who suggested that certain gases could absorb long-wave radiation emitted from the surface and the lower atmosphere. The absorbed energy is re-emitted thereby increasing the incoming radiation at the Earth's surface. Without the natural occurring greenhouse gases, the most important is H_2O , our environment would be uncomfortably cold. By the late 1850s, the British physicist (and alpinist – first on Weisshorn) J. Tyndall had analyzed the radiative properties of atmospheric gases and demonstrated that carbon dioxide is among those that strongly absorb infrared radiation.

In 1896, the Nobel Prize-winning Swedish chemist Svante Arrhenius proposed that carbon dioxide emitted into the atmosphere by burning fossil fuels such as coal, oil and natural gas would cause a change in the transparency of the atmosphere that might result in a warming larger than ever experienced in human history. Arrhenius calculated that a doubling of atmospheric CO_2 leads to a warming of about 5 to 6°C. Arrhenius and his colleagues were mainly interested in the role of carbon dioxide for the large glacial-interglacial climate swings.

Today, we know that CO_2 has increased from its preindustrial level of 280 ppm to more than 400 ppm today (Figs 4.1, 4.2). This increase has been demonstrated by C. D. Keeling's measurements of carbon dioxide starting in 1958 at Mauna Loa, Hawaii and by the first CO_2 measurements on air bubbles entrapped in Antarctic ice analyzed in the early 1980s at the Climate and Environmental Physics in Bern [*Neftel et al.*, 1985]. This data set forms perhaps the most important geophysical record of the century. Meanwhile, the CO_2 increase has been confirmed by other laboratories analyzing several different ice cores; the atmospheric concentration is continuously monitored on the order of 50 stations around the globe.

Revelle and Suess [1957] published a paper that made major contributions to our understanding of the carbon cycle. They estimated carbon uptake by the ocean using a model to describe air-sea exchange, carbon chemistry and ocean transport. They calculated that the chemical capacity of seawater to take up anthropogenic CO₂ is substantially less than might be expected by assuming that it would redistribute itself according to the present atmosphere and ocean inventories. Their chemical equilibration model that includes this "Revelle effect" show that between 15 to 20% of the carbon dioxide added to the atmosphere will remain there permanently.

A big advancement was their use of radiocarbon measurements to demonstrate that ocean mixing is slow compared to the rate at which carbon is released by humans. Radiocarbon, which is produced by interactions between cosmic rays and atmospheric nitrogen, enters the ocean by gas exchange and is mixed towards the abyss where its concentration decreases due to radioactive decay. This radioactive clock provides a measure of the surface-to-deep exchange rates and allowed Revelle and Suess to validate transport in their ocean model.

Revelle and Suess emphasized the different time scales operating in the climate system and concluded in their paper:

"Human beings are now carrying out a large-scale geophysical experiment of a kind that could not have happened in the past nor be reproduced in the future. Within a few centuries we are returning to the atmosphere and oceans the concentrated organic carbon stored in sedimentary rocks over hundred of million of years."

Siegenthaler and Oeschger [1978] demonstrated by using their box-diffusion model [*Oeschger et al.*, 1975] that carbon dioxide emissions must fall well below present values and be phased out eventually if the concentration of atmospheric CO₂ is to be stabilized. This is a direct consequence of the limited capacity of the ocean and the land biosphere to absorb fossil fuel CO₂ and that CO₂ is chemically stable in the atmosphere and not destroyed through oxidative processes or deposition on the ground. Siegenthaler and Oeschger made the socio-economic consequences clear:

"With climate models becoming more and more realistic, a maximum permissible atmospheric CO₂ level might be found which should not be exceeded if the atmospheric radiation balance is not to be disturbed in a dangerous way. ... This scenario clearly does not allow us to go on burning fossil fuel at the present growth rate for a long time ... around the turn of the century (i.e., 2000 AD) new technologies would have to take over a substantial part of global energy production."

The United Nations negotiated at the Rio Earth Summit in 1992 as an ultimate aim:

"... to achieve stabilization of greenhouse gas concentrations ... at a level that would prevent dangerous interference with the climate system ... within a time frame sufficient to allow ecosystems to adapt naturally to climate change, to ensure that food production is not threatened and to enable economic development to proceed in a sustainable manner."

The findings of Siegenthaler and Oeschger have meanwhile been confirmed by many analyses: anthropogenic carbon dioxide emissions have to fall well below present values and need

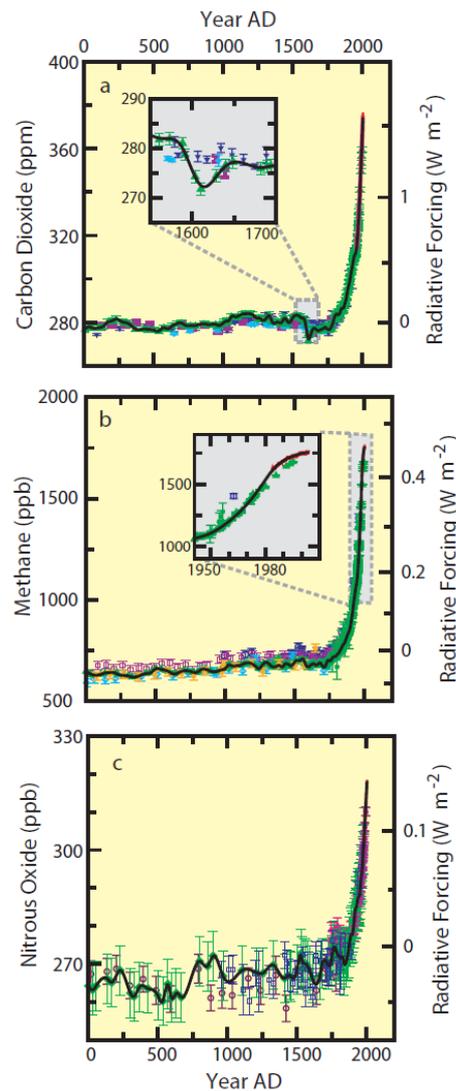


Figure 4.1: The rise in CO_2 , CH_4 , and N_2O . Since 1958 atmospheric measurements of CO_2 have been made on Mauna Loa in Hawaii. To extend the record further back in time, air extracted from Antarctic ice cores has been measured (some of these measurements are carried out here at the University of Bern in the Physics Institute). This figure was produced in 2007 in preparation of the Fourth Assessment Report of IPCC. In 2021, CO_2 has risen to 416 ppm, CH_4 to 1896 ppb, and N_2O to 335 ppb.

eventually be phased out if the greenhouse gas carbon dioxide is to be stabilized in the atmosphere.

Society is still struggling to implement the necessary CO_2 -emission free energy technologies on a large scale and today's carbon dioxide emissions are still growing and are higher than ever.

While our understanding of the functioning of the Earth system and interactions among the physical climate system, biogeochemical cycles, and the socio-economic system has improved, there still remain great scientific challenges. For example, the magnitude of climate-carbon

cycle feedbacks is not well constrained, the impact of ocean acidification, a consequence of CO₂ uptake by the ocean, on the marine ecosystems are poorly understood, the mechanisms driving uptake of excess CO₂ by the land biosphere are not well understood, and a comprehensive quantitative explanation of glacial-interglacial CO₂ variations is still missing.

4.1.2 The key role of biogeochemical cycles

The global cycles of carbon, oxygen, and other elements (N, P, Si, Fe, etc.) are biogeochemical cycles. This implies that biological, geological, chemical and physical processes are important. The cycles of the different elements such as C, H₂O, N and O are closely linked. The carbon cycle plays a key role:

- The carbon cycle influences the composition and fluxes of the marine and terrestrial biosphere. The cycling of carbon and water (and the energy balance at the land surface) are coupled through uptake and release by the biosphere.
- The carbon cycle influences the global radiation balance. CO₂ and CH₄ are ranked as second and third most important greenhouse gases. Both include the element carbon and their concentration has increased over the industrial period.
- A change in the energy balance can potentially lead to irreversible changes in the climate system. For each kg of carbon emitted into the atmosphere by the burning of coal, gas, and oil, around 200 grams will still be airborne after one thousand years and around 100 grams after many millennia. Changes in the radiation balance may lead to a completely altered climate due to non-linearities in the system.
- Uptake of fossil fuel CO₂ leads to a decrease in ocean pH and to ocean acidification that is likely to affect marine ecosystems.
- The carbon cycle influences the distribution of nutrients and oxygen in the ocean. Tracer distribution, ocean circulation and carbon cycle are closely linked. Tracer distributions provide crucial information about how the ocean is operating.

Variations in atmospheric CO₂ and CH₄ and their isotopes and the distribution of nutrient and carbon isotopes as recorded in Antarctic ice, tree rings and marine sediments provide important, quantitative insights into past climates.

4.2 Greenhouse gas increase and climate change

Human activity has led to a significant perturbation of the global carbon cycle and climate during the past 200 years and especially since World War II. The rise in CO₂ and other greenhouse gases has been clearly documented due to the combination of atmospheric data (since 1958) and the analysis of ice cores. Prior to industrialisation, the concentrations of CO₂, CH₄ and N₂O in the atmosphere were about constant over the last millennium. Their concentrations today are the highest ever over the last 800,000 years.

4.3 The main carbon reservoirs

Figure 4.3 and Table 4.1 summarize the size and fluxes of the main carbon reservoirs. The preindustrial carbon cycle is perturbed considerably due to the addition of fossil fuels and deforestation.

The most important carbon reservoirs are characterized as follows:

Atmosphere

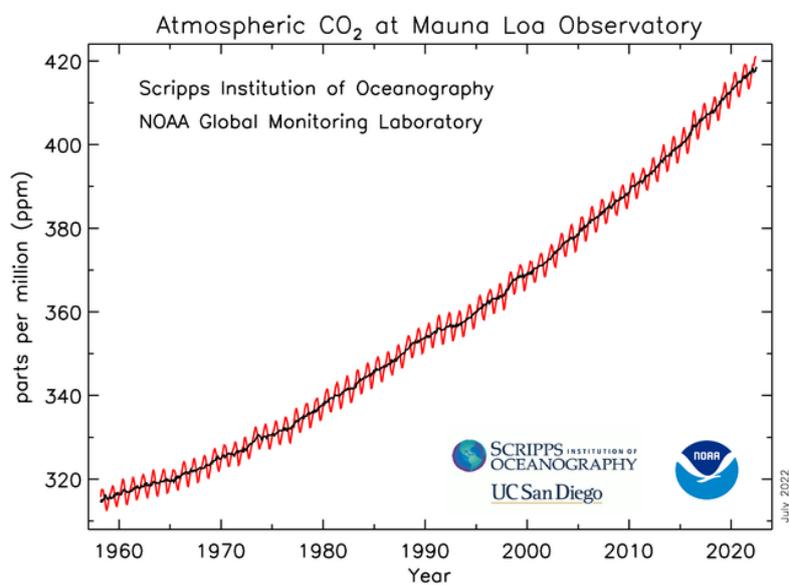


Figure 4.2: CO₂ mixing ratio at Mauna Loa Observatory, Hawaii (www.esrl.noaa.gov).

The main forms of carbon in the atmosphere are carbon dioxide (CO₂) (2000: 370 ppm; 2015: 400 ppm), methane (CH₄, 2000: 1.7 ppm = 1700 ppb; parts per billion volume) and carbon monoxide (CO; 50 ... 100 ppb). The atmospheric CO₂ concentration has a direct effect on the radiative balance and therefore on physical climate variables. Today's concentration is almost 50% higher than during the preindustrial. The CH₄ concentration has more than doubled in the past 200 years. The atmospheric carbon concentration is increasing by approximately 1.5 ppm per year due to the burning of fossil fuels, forest clearing and changes in land use. Exact measurements of atmospheric CO₂ from many stations worldwide show:

- The rise in the mean CO₂ concentration and in the interhemispheric CO₂ gradient over the past decades.
- There are seasonal fluctuations with maxima in spring and minima in autumn that reflect the metabolism of the corresponding hemisphere.
- A considerably stronger seasonality in the northern hemisphere since it accounts for a larger portion of continental landmass and vegetation.

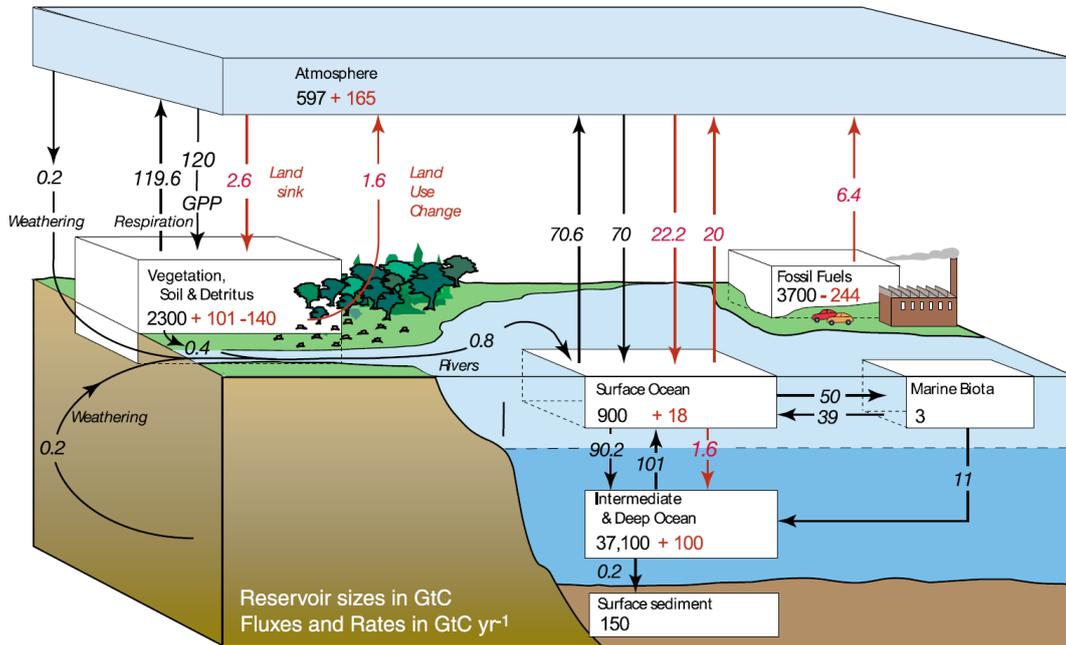


Figure 4.3: The global carbon cycle for the 1990s, showing the main annual fluxes in GtC yr⁻¹ pre-industrial "natural" fluxes in black and "anthropogenic" fluxes in red. The net terrestrial loss of -39 GtC is inferred from cumulative fossil fuel emissions minus atmospheric increase minus ocean storage. The loss of -140 GtC from the "vegetation, soil and detritus" compartment represents the cumulative emissions from land use change [Houghton, 2003], and requires a terrestrial biosphere sink of 101 GtC ([Sabine et al., 2004], given only as ranges of -140 to -80 GtC and 61 to 141 GtC, respectively; other uncertainties given in their Table 1). Gross fluxes generally have uncertainties of more than $\pm 20\%$ but fractional amounts have been retained to achieve overall balance when including estimates in fractions of GtC yr⁻¹ for riverine transport, weathering, deep ocean burial, etc. "GPP" is annual gross (terrestrial) primary production. Atmospheric carbon content and all cumulative fluxes since 1750 are as of end 1994. Figure from Denman et al. [2007].

Biosphere

A simplified formula for organic material is CH_2O , i.e. $\text{C}_6\text{H}_{12}\text{O}_6$ (glucose). The fluxes for the land and marine biosphere are approximately the same, but the average life time differs on account of the variation in the reservoir sizes (see Table 4.1): On the order of 10 years for the living land biosphere (minus soil) and 1 month for the marine biosphere, but many centuries for peat and permafrost soils.

Ocean

The ocean, with a carbon reservoir of 38,000 GtC, is the largest fast (< 1000 years) exchanging carbon reservoir. Carbon is primarily found in the dissolved and inorganic form in the ocean. It is represented as $\sum \text{CO}_2$ or DIC (dissolved inorganic carbon), which comprises the sum of bicarbonate (HCO_3^- ; approx. 90%), carbonate (CO_3^{2-} ; about 10%) and dissolved CO_2 gas (about 0.5%). A small percent of the carbon is in the dissolved organic form, DOC (dissolved

organic carbon).

Sediments/Litosphere

A giant reservoir, but it exchanges carbon very slowly with the ocean on geological timescales. Carbon in sediments is primarily found as CaCO_3 (calcite and aragonite) and MgCO_3 , while approximately 1% is derived from organic matter. It is an important reservoir when considering timescales longer than a few thousand years (i.e. glacial-interglacial transitions).

Box 4.1 Units for carbon and CO_2 .

The amount of carbon in a reservoir or transported by a flux is often expressed in units of ton-carbon or gigaton of carbon ($1 \text{ GtC} = 10^9 \text{ tC} = 10^{15} \text{ g C}$).

Sometimes the mass is related to the CO_2 molecule and the unit is then in tons of CO_2 . It holds $1 \text{ ton-C} = 44/12 \text{ ton-CO}_2$ as the molecular weight of carbon is 12 g/mol and that of CO_2 is 44 g/mol.

Alternatively, the amount is expressed in moles C (mol C). The conversion is $1 \text{ mol C} = 12 \text{ gram of carbon}$.

The atmospheric CO_2 concentration is normally reported as a mixing ratio in ppm (parts per million volume= 1 part CO_2 per 10^6 parts of dry air). An average mixing ratio of 1 ppm is equivalent to 2.12 GtC.

Major global carbon reservoirs and fluxes.			
Reservoirs			10^{15} g C
Atmosphere			
CO ₂	1800	≈ 280 ppm	594
	2015	400 ppm	850
Other Gases	CH ₄ (2015)	1,834 ppm	4
	CO	0.1 ppm	
(Troposphere: 80 %, stratosphere 20 % of atmospheric mass)			
Oceans:	Inorganic C (\sum CO ₂)		38,000
	Dissolved organic matter		700
	Biomass		3
Land biosphere	Living		450-650
	Soil, humus, including frozen soil		3000-4000
Groundwater			450
Sediments/Lithosphere	Inorganic C		60,000,000
	Organic C		12,000,000
Fossil fuels (conventional resources)			5,000
Fluxes			10^{15} g C/yr
Atmosphere - ocean, gross CO ₂ exchange (preind.)			60
Atmosphere - land biota, net photosynthesis/respiration (NPP)			60
Marine photosynthesis (NPP)			50
Net sedimentation in oceans/weathering			0.2
Volcanism			0.1
Fossil fuel combustion (incl. cement production), 2000-2009			7.8
Net land use flux, 2000-2009			1.1
Residence times: $\tau = mass/flux$			
Atmosphere (pre-industrial)	total exchange		4.6 yr
	exchange with ocean only		10 yr
	exchange with biosphere (NPP) only		10 yr
Living land/biosphere: photosynthesis/respiration			11 yr
Marine biosphere: photosynthesis/respiration			0.06 yr
Oceans:	exchange with atmosphere, total flux		630 yr
	sedimentation only		200,000 yr
Atmosphere + biosphere + oceans: sedimentation			212,000 yr

Table 4.1: Global carbon reservoirs and residence times of carbon in the various reservoirs. Main sources: [Denman *et al.*, 2007; Sabine *et al.*, 2004; Tarnocai *et al.*, 2009; Ciais *et al.*, 2013], http://www.esrl.noaa.gov/gmd/ccgg/trends_ch4/#global_data.

Chapter 5

Glacial-interglacial atmospheric CO₂ Variations

5.1 Milankovitch cycles of climate over the past 800,000 years

Atmospheric CO₂ varied between 172 and 300 ppm during the past glacial-interglacial cycles of the last 800,000 years (Fig. 5.1). We discuss first the general relationship between climate and atmospheric CO₂ during the last million years, before addressing the potential causes of the variations in CO₂. Some of the figures and text has been taken from the paleoclimate chapter of the Fourth IPCC Assessment [*IPCC*, 2007]. Paleoclimatic records document a sequence of glacial-interglacial cycles over the past three million years. The typical periodicity of large glacial-interglacial climate swings is around 100 kyr for the period of the past 430 ka, but there is also substantial climate variability on other orbital time scales, in particular the 41,000 kyr period of the variation in the tilt of the axis of the Earth. A minor portion (20% on average) of each glacial-interglacial cycle was spent in the warm interglacial mode, which normally lasted between 10 and 30 kyr. There is evidence for longer interglacial periods between 430 and 740 ka, but these were apparently colder than the typical interglacials of the latest Quaternary. The Holocene, the latest of these interglacial, extends to the present.

Available evidence supports the view that glacial-interglacial cycles are driven by changes in the Earth's orbit around the sun, the so-called Milankovitch cycles (Fig. 5.2). These cycles change the amount of solar energy received at each latitude and in each season, but hardly affect the global annual mean. Many studies suggest that the amount of summer sunshine on northern continent is crucial. A low solar energy flux during northern hemisphere summers results in permanent snow cover and eventually the build-up of glaciers. Growing glacial surface area results in an increase in the reflection of solar radiation back to space and enforce the cooling.

The ice core record indicates that atmospheric CO₂ co-varied with Antarctic temperature over glacial-interglacial cycles, suggesting a close link between natural CO₂ variations and temperature. Variations in CO₂ over the last 420 kyr broadly followed Antarctic temperature, typically by several centuries to a millennium. During the last deglaciation, and likely also the three previous ones, the onset of warming at both high southern and northern latitudes preceded by several thousand years the first signals of significant sea level increase resulting from the melting of the northern ice sheets linked with the rapid warming at high northern latitudes. The global mean sea level rose by around 120 to 140 m during the last transition and was primarily caused by the melting of northern hemisphere ice-sheets that covered large parts of North America and Scandinavia during the Last Glacial Maximum (LGM).

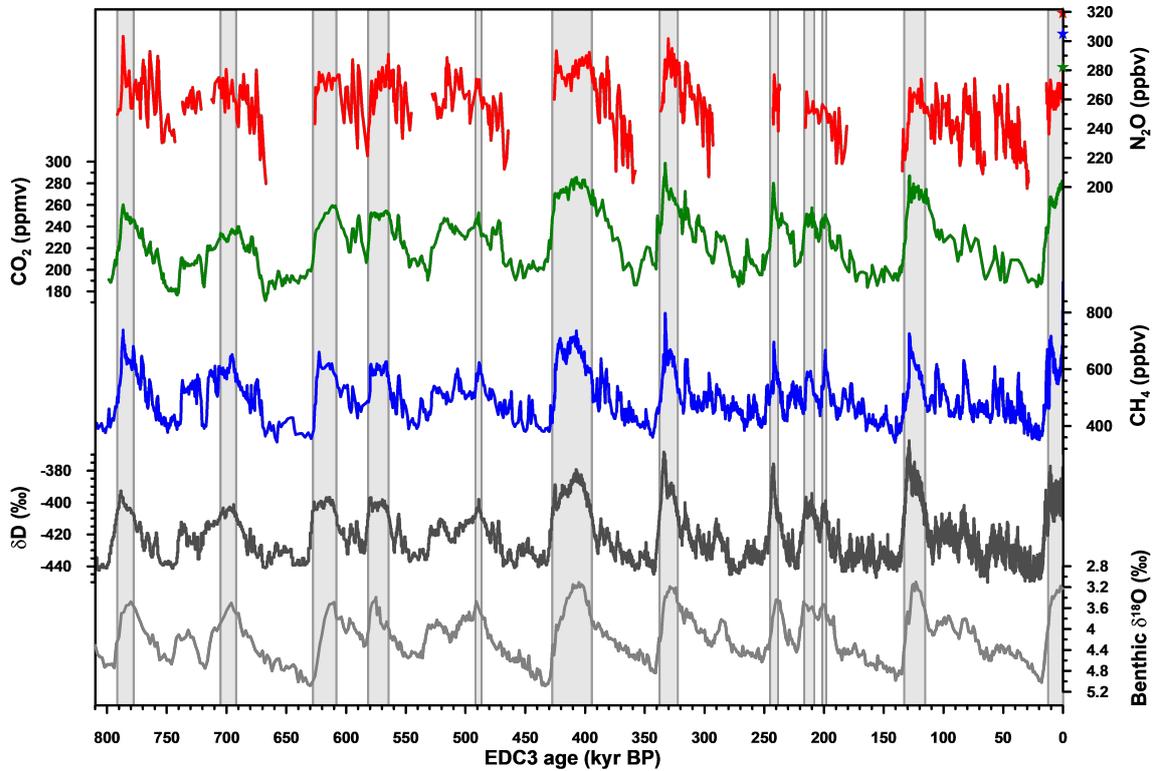


Figure 5.1: Variations of deuterium (δD), a proxy for local temperature, and the atmospheric concentrations of the greenhouse gases carbon dioxide, methane, and nitrous oxide derived from air trapped within ice cores from Antarctica and from recent atmospheric measurements [Jouzel *et al.*, 2007; Loulergue *et al.*, 2008; Lüthi *et al.*, 2008; Schilt *et al.*, 2010; Siegenthaler *et al.*, 2005; Spahni *et al.*, 2005]. The shading indicates interglacial warm periods. Interglacial periods also existed prior to 450,000 yrs, but these were apparently colder than the typical interglacials of the latest Quaternary. The length of the current interglacial is, in the context of the last 800,000 yrs, not unusual. The stack of 57 globally distributed benthic $\delta^{18}O$ marine records, a proxy for global ice volume fluctuations, is displayed in the bottom part of the figure for comparison with the ice core data. Larger ice volume is expressed downwards. The marine record is plotted on its original time scale based on tuning to the orbital parameters. The stars indicate modern concentrations.

Although CO₂ was likely not the primary cause for glacial-interglacial variations, it contributed significantly to the variations in the radiative balance of the surface-lower atmosphere system (Fig. 5.3). The link between Antarctic temperature and CO₂ did not change on the glacial-interglacial time scale during the past 800 ka. This indicates a rather stable coupling between climate and the carbon cycle.

5.2 Mechanisms of natural CO₂ variations

A number of processes contributed to the low CO₂ concentrations during glacial times as discussed in Box 5.1. However, the quantification of contributions from individual mechanisms to the total signal has proven difficult given uncertainties in the proxy records and shortcomings of current models. In the following, we will discuss a few selected mechanisms and ask the following questions:

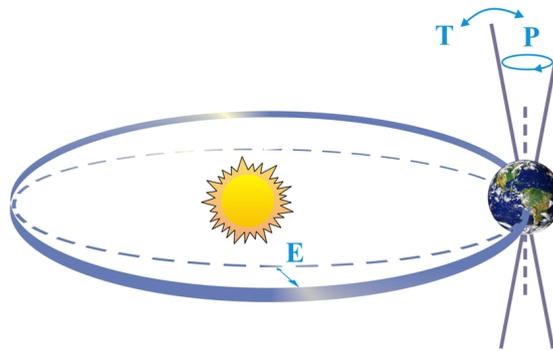


Figure 5.2: Schematic perspective view of the Earth's orbital changes (Milankovitch cycles) that drive the ice age cycles. "T" denotes changes in the tilt (or obliquity) of the Earth axis, "E" denotes changes in the eccentricity of the orbit (due to variations in the minor axis of the ellipse), and "P" denotes precession, i.e., changes in the direction of the axis tilt at a given point of the orbit. Source: [IPCC, 2007].

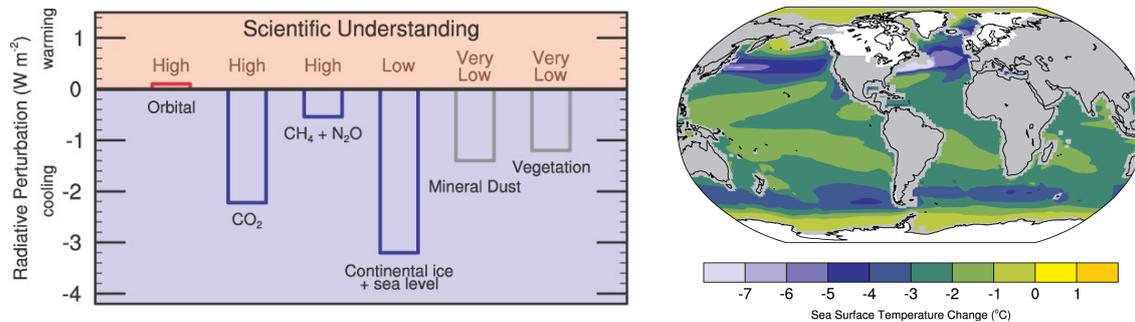


Figure 5.3: The Last Glacial Maximum (LGM) climate (approximately 21 kyrs ago) relative to pre-industrial (1750). left: Global annual mean radiative influences (W m^{-2}) of LGM climate change agents, generally feedbacks on glacial-interglacial cycles, but also specified in most AOGCM simulations for LGM. right: Multi-model average SST change for LGM PMIP-2 simulations by five AOGCMs (CCSM, FGOALS, HadCM, IPSL, and MIROC). Ice extent over continents is shown in white. Source: [IPCC, 2007].

- What regulates the distribution of DIC in the modern ocean and the partition of carbon between the atmosphere and the ocean?
- What is the impact of
 - a) the temperature distribution of the ocean
 - b) the export of organic material out of the surface, and
 - c) the export of calcium carbonate
 on the oceanic distribution of dissolved inorganic carbon (DIC) and on atmospheric CO₂?

The Bern3D model is utilized to answer these questions quantitatively. Our starting point is a dead ocean without any biological activity and a uniform temperature (and salinity) distribution. Solubility of CO₂ is everywhere the same and production and export of organic material and calcium carbonate by biological activities is zero. This implies that the concentration of DIC and other tracers is uniform in the ocean. The carbon inventory in the ocean-atmosphere system is set to the observed inventory and ocean temperature is set

everywhere to 18°C, i.e. the global average surface water temperature of today's ocean. This setup yields a uniform distribution of DIC within the ocean and an atmospheric CO₂ mixing ratio of 560 ppm. This is twice the preindustrial value of 280 ppm simulated with the model for a realistic temperature distribution and export of biogenic material.

5.2.1 Temperature

The temperature (and thus CO₂ solubility) of today's ocean are now implemented in the model. The deep ocean is substantially colder than the average surface water temperature (Fig. 5.4). Cold water stores more carbon than warm waters for a given chemical composition and CO₂ partial pressure. The link between partial pressure and temperature is:

$$p\text{CO}_2(T) = p\text{CO}_2(T_0)e^{0.0423\text{K}^{-1}\cdot(T-T_0)}, \quad (5.1)$$

and the link between pCO₂ and DIC is

$$\xi = \frac{\Delta p\text{CO}_2}{p\text{CO}_{2,o}} / \frac{\Delta \Sigma\text{CO}_2}{\Sigma\text{CO}_{2,o}} \approx 10. \quad (5.2)$$

The cold deep ocean is enriched in carbon relative to the warm surface and atmospheric CO₂ is lower for a realistic temperature distribution than for an ocean with a uniform temperature of 18°C. Atmospheric CO₂ is lowered from 560 to 439 ppm in the model.

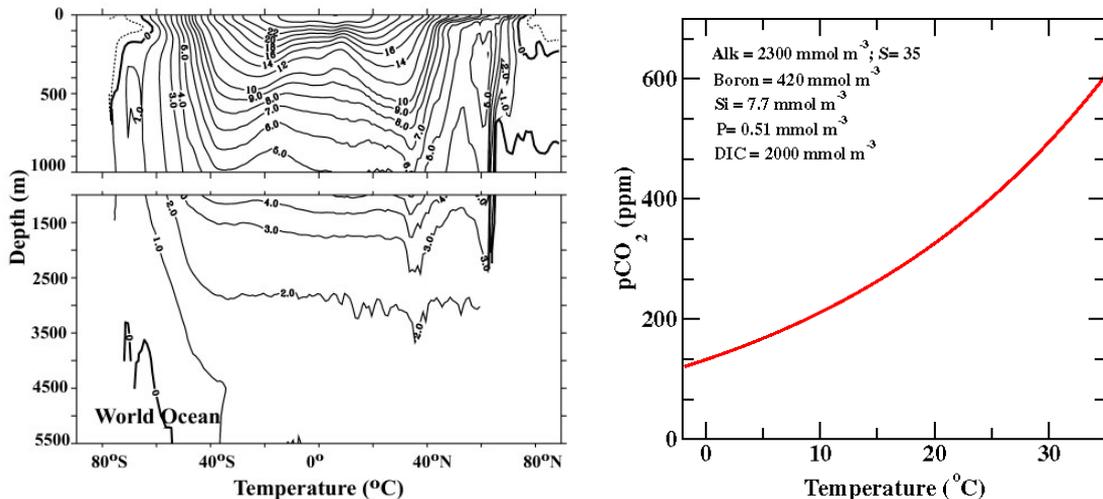


Figure 5.4: Left: zonal mean temperature for the world ocean. Right: The CO₂ partial pressure as a function of water temperature for a chemical composition typical for surface seawater.

5.2.2 Export of organic material

The export of organic matter is mediated by biological activities (Fig. 5.5). Carbon and nutrients (P, N, Fe, ..) are utilized by marine life and dead organic material is export to depth where it is added back again to the inorganic nutrient pool through remineralization by bacteria. The consequence is that nutrient and carbon concentrations are depleted at the

surface and enriched at depth. Adding this cycle of marine organic material to the Bern3D model results in a further lowering of atmospheric CO₂ compared to a dead ocean. An atmospheric CO₂ concentration of 229 ppm is achieved. Even lower atmospheric CO₂ can be obtained when all phosphate (a limiting nutrient) is consumed everywhere in the surface water and DIC lowered in proportion. The biological cycles of nutrients and carbon are strongly correlated and nutrients are (on average) incorporated into organic material according to:

$$\text{Redfield ratios: } P : N : C : O_2 = 1 : 16 : 120 : -170. \quad (5.3)$$

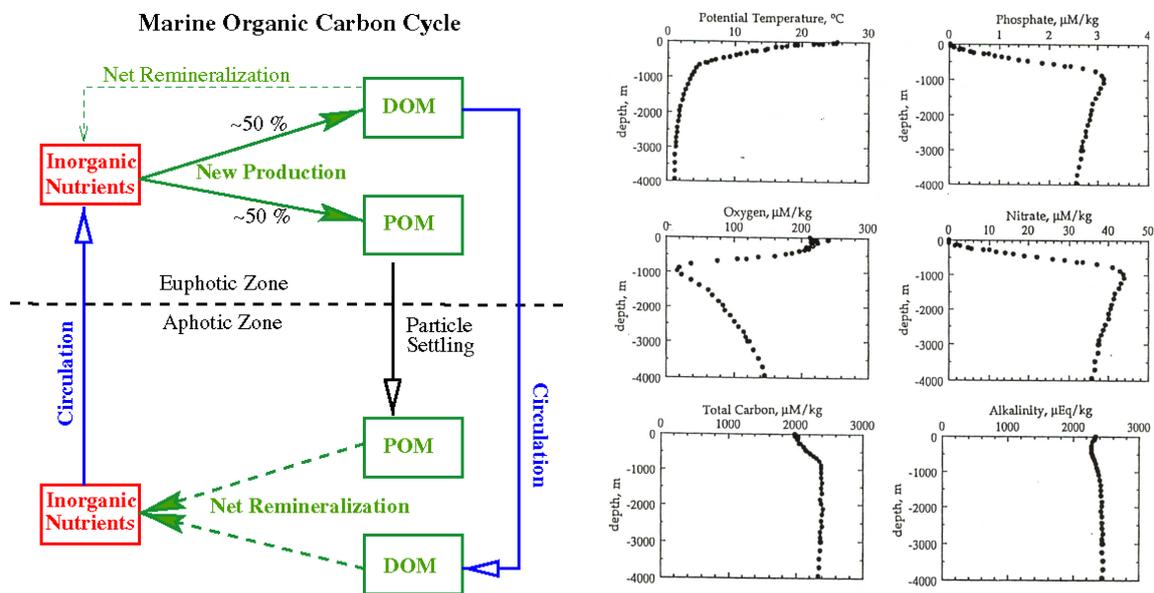


Figure 5.5: Left: A geochemist's view of the marine biological cycle. Inorganic nutrients and carbon are incorporated into dissolved and particulate organic matter (DOM, POM) by marine life. The organic material is exported to depth and remineralized back into inorganic nutrients under the consumption of oxygen. The result is a depletion of nutrients and carbon at the surface and an enrichment at depth. Right: Vertical distribution of temperature, phosphate, oxygen, nitrate, dissolved inorganic carbon, and of alkalinity in the North Pacific.

Box 5.1 Low CO₂ during glacial times.

What caused the low atmospheric CO₂ concentrations during glacial times?

Ice core records show that atmospheric CO₂ varied in the range of 172 to 300 ppm over the glacial-interglacial cycles of the last 800,000 yrs [Lüthi *et al.*, 2008; Petit *et al.*, 1999; Siegenthaler *et al.*, 2005]. The quantitative and mechanistic explanation of these CO₂ variations remains one of the big unsolved questions in climate research. Processes in the atmosphere, ocean, marine sediments, on land, and the dynamics of sea ice and ice sheets must be considered. A number of hypotheses for the low glacial CO₂ concentrations have emerged over the past 20 years, and a rich body of literature is available [Archer *et al.*, 2000; Broecker, 1998; Kohfeld *et al.*, 2005; Sigman and Boyle, 2000; Webb *et al.*, 1997]. Many processes have been identified that could potentially regulate atmospheric CO₂ on glacial-interglacial time scales. However, the existing proxy data with which to test hypothesis are relatively scarce, uncertain, and their interpretation is partly conflicting.

Most explanations propose changes in oceanic processes as the cause for low glacial CO₂. The ocean is by far the largest of the relatively fast (< 1000 yr) exchanging carbon reservoirs, and terrestrial changes cannot explain the low glacial values because terrestrial storage was also low at the Last Glacial Maximum. On glacial-interglacial time scales, atmospheric CO₂ is mainly governed by the interplay between ocean circulation, marine biological activity, ocean-sediment interactions, seawater carbonate chemistry, and air-sea exchange. Upon dissolution in seawater, CO₂ maintains an acid/base equilibrium with bicarbonate and carbonate ions that depends on the acid-titrating capacity of seawater, i.e., alkalinity. Atmospheric CO₂ would be higher if the ocean lacked biological activity. CO₂ is more soluble in colder than in warmer waters; therefore changes in surface and deep ocean temperature have the potential to alter atmospheric CO₂. Most hypotheses focus on the Southern Ocean, where a large volume-fraction of the cold deep-water masses of the world ocean are currently formed, and large amounts of biological nutrients (phosphate and nitrate) upwelled to the surface remain unused. A strong argument for the importance of Southern Hemisphere processes is the co-evolution of Antarctic temperature and atmospheric CO₂.

One family of hypotheses of low glacial CO₂ values invokes an increase or redistribution in the ocean alkalinity as a primary cause. Potential mechanisms are (i) the increase of CaCO₃ weathering on land, (ii) a decrease of coral reef growth in the shallow ocean, or (iii) a change in the export ratio of CaCO₃ and organic material to the deep ocean. These mechanisms require large changes in the deposition pattern of CaCO₃ to explain the full amplitude of the glacial-interglacial CO₂ difference through a mechanism called carbonate compensation [Archer *et al.*, 2000]. The available sediment data do not support a dominant role for carbonate compensation in explaining low glacial CO₂ levels. Furthermore, carbonate compensation may only explain slow CO₂ variation, as its time scale is multi-millennial.

Another family of hypotheses invokes changes in the sinking of marine plankton. Possible mechanisms include (iv) fertilization of phytoplankton growth in the Southern Ocean by increased deposition of iron-containing dust from the atmosphere after being carried by winds from colder, drier continental areas, and a subsequent redistribution of limiting nutrients, (v) an increase in the whole ocean nutrient content, e.g., through input of material exposed on shelves or nitrogen fixation, and (vi) an increase in the ratio between carbon and other nutrients assimilated in organic material, resulting in a higher carbon export per unit of limiting nutrient exported. As with the first family of hypotheses, this family of mechanisms also suffers from the inability to account for the full amplitude of the reconstructed CO₂ variations when constrained by the available information. For example, periods of enhanced biological production and increased dustiness (iron supply) are coincident with 20 to 50 ppm changes. Consistently, model simulations suggest a limited role for iron in regulating past atmospheric CO₂ concentration [Bopp *et al.*, 2002].

Physical processes also likely contributed to the observed CO₂ variations. Possible mechanisms include (vii) changes in ocean temperature (and salinity), (viii) suppression of air-sea gas exchange by sea ice, and (ix) increased stratification in the Southern Ocean. The combined changes in temperature and salinity increased the solubility of CO₂, causing a depletion in atmospheric CO₂ of perhaps 30 ppm. Simulations with general circulation ocean models do not fully support the gas exchange-sea ice hypothesis. One explanation (ix) conceived in the 1980s invokes more stratification, less upwelling of carbon and nutrient-rich waters to the surface of the Southern Ocean, and increased carbon storage at depth during glacial times. The stratification may have caused a depletion of nutrients and carbon at the surface, but proxy evidence for surface nutrient utilization is controversial. Qualitatively, the slow ventilation is consistent with very saline and very cold deep waters reconstructed for the Last Glacial Maximum [Adkins *et al.*, 2002], as well as low glacial stable carbon isotope ratios (¹³C/¹²C) in the deep South Atlantic.

In conclusion, the explanation of glacial-interglacial CO₂ variations remains a difficult attribution problem. It appears likely that a range of mechanisms have acted in concert (e.g. [Köhler *et al.*, 2005]). The future challenge is not only to explain the amplitude of glacial-interglacial CO₂ variations, but the complex temporal evolution of atmospheric CO₂ and climate consistently.

5.2.3 Export of CaCO₃

Many forms of marine life (corals, pteropods, coccolithophorids) produce structures of calcium carbonate (CaCO₃). The production of CaCO₃ and its subsequent export to the deep increases the partial pressure of CO₂ in surface waters. This may seem counterintuitive as the production of CaCO₃ decreases DIC. However, the removal of carbonate ions leads also to a shift in the acid-base state and the alkalinity (Alkalinity = [HCO₃⁻] + 2 [CO₃²⁻]) of the water is decreased, whereas [H₂CO₃*] and thus pCO₂ is increased. Relative changes in partial pressure and alkalinity are

$$\frac{\frac{\Delta p\text{CO}_2}{p\text{CO}_2}}{\frac{\Delta \text{Alk}}{\text{Alk}}} \approx -10. \quad (5.4)$$

Gas exchange does not alter alkalinity, in contrast to the formation and dissolution of CaCO₃. Observations show that alkalinity is slightly depleted at the surface relative to the abyss (Figure 5.5). Invoking the cycling of CaCO₃ in the model brings the atmospheric CO₂ back from 229 to 280 ppm.

Table 5.1 provides an overview of how different processes affect atmospheric CO₂. The take home message is that changes in the ocean's temperature distribution, in the cycling of organic material, and in the cycling of CaCO₃ have all the potential to significantly affect atmospheric CO₂ on time scales from decades to a few millennia.

On time scales larger than 1,000 years, exchange of carbon, nutrients, and alkalinity between ocean waters and the top few centimeters of ocean sediments need to be taken into account to explain CO₂ variations. On geological time scales, changes in the balance between fluxes from weathering on land versus burial fluxes to the lithosphere govern atmospheric CO₂. Weathering, for example of silicate and calcium carbonate rocks, adds alkalinity, nutrients and carbon to the ocean. Burial by consolidation of sediments removes these tracers from the ocean.

Status of/processes in Bern3D model	ΔCO ₂	atmospheric CO ₂
Dead ocean with uniform T of 18°C		560 ppm
+ realistic temperature distribution	-121	439 ppm
+ organic matter cycle	-210	229 ppm
+ CaCO ₃ cycle	+49	278 ppm
All pumps:	282	

Table 5.1: Summary of model results for hypothetical model setups intended to demonstrate the importance of different processes in regulation atmospheric CO₂ on century to millennial time scales. Ocean-sediment interactions acting on multi-millennial time scales would modify these responses.

Chapter 6

The uptake of anthropogenic carbon by the ocean

6.1 Dissolved gases in the ocean

Dissolved gases are found in many areas of climate and environmental research. For many cycles, such as carbon and oxygen, dissolved gases play a role in the ocean. Dissolved trace gases in aerosol drops are often important for atmospheric chemistry reactions. The concentration of dissolved gases in groundwater is dependent on temperature and therefore noble gases in old groundwater are able to provide information of environmental conditions at the time of infiltration.

6.1.1 Solubility of gases

The concentration of a dissolved gas in water is proportional to the concentration of the gas in air. The gas concentration in air and water can be expressed differently, which affects how the solubility of a gas is defined.

Here we will express the concentrations as the following:

$p(i)$	Partial pressure of gas, i , in air (gaseous phase),	unit: atm
c_w	concentration of the dissolved gas in water,	unit: $\frac{\text{ml (gas; STP)}}{\text{l (water)}}$
c_g	concentration of the gas in the gas phase,	unit: $\frac{\text{ml (gas; STP)}}{\text{l (gas)}}$

NB: STP = Standard Temperature and Pressure (0°C, 1 atm)
1 ml (gas; STP) = (1/22,400) mol = $4.46 \cdot 10^{-5}$ mol
Sometimes c_w is expressed as mol/l(water) and c_g in mol/l(gas) ($c_g = n/V = p/(RT)$).

6.1.2 Bunsen's solubility coefficient α and Henry's law

By definition, the Bunsen coefficient is the volume of gas at 273.15 K and standard pressure p_0 , which dissolves in unit volume of a solvent when the partial pressure of the gas equals p_g . It follows:

$$c_w = \alpha \cdot p_g ; \text{unit}[\alpha]: \frac{\text{ml (gas; STP)/ l(water)}}{\text{atm}}, \quad (6.1)$$

α is dependent on the temperature and the salt concentration of water, but only weakly affected by the partial pressure of the gas.

Another formulation is advantageous for certain applications. In this case the units of the concentration in the air and water are the same, thus the solubility is dimensionless. s

represents the volumetric solubility and we write Henry's law as:

$$c_w = s \cdot c_g ; \text{unit}[s]: \frac{\text{ml (gas; STP)/l(water)}}{\text{ml (gas; STP)/l(water)}} \quad (6.2)$$

6.1.3 Relationship between α and s

In agreement with both definitions above:

$$s = \frac{c_w}{c_g} = \alpha \frac{p_g}{c_g} \quad (6.3)$$

Next, we determine the ratio, p_g/c_g . ($p_g; V_g; T$ refer to current environmental conditions and $p_{g,0}; V_{g,0}; T_0$ to STP). From the gas law ($pV = nRT$) we get:

$$\frac{V_{g,0}}{V_g} = \frac{p_g T_0}{p_{g,0} T} ; \text{unit}: \frac{\text{m}^3(\text{gas; STP})}{\text{m}^3(\text{gas})} \quad (6.4)$$

Setting this equal to c_g after conversion of units yields:

$$c_g = \frac{V_{g,0}}{V_g} \frac{1000\text{ml}(\text{gas; STP})}{\text{l}(\text{gas})} = \frac{p_g T_0}{p_{g,0} T} \frac{1000\text{ml}(\text{gas; STP})}{\text{l}(\text{gas})} \quad (6.5)$$

and the following ratio:

$$\frac{p_g}{c_g} = \frac{p_{g,0} T}{T_0} \frac{1}{1000} \frac{\text{l}(\text{gas})}{\text{ml}(\text{gas; STP})} \quad (6.6)$$

and solubility and its coefficients are linked as follows:

$$s = \alpha \frac{T}{273.15 \text{ K}} \frac{1}{1000} \frac{(\text{atm})}{\text{ml}(\text{gas; STP})/\text{l}(\text{gas})} \quad (6.7)$$

From Table 6.1 we can make the following conclusions:

- The solubility of noble gases increases with increasing molecular mass.
- In general, the solubility decreases with increasing temperature. This effect is larger for the heavier noble gases (and CO_2) compared to the lighter gases. The volumetric solubility for CO_2 is close to 1, while for the other gases listed it ranges between 10^{-2} and 10^{-1} .

Example:

We would like to estimate the ratio of oxygen in the ocean and the atmosphere. We apply Henry's law and approximate the atmosphere and the ocean as well-mixed boxes. We define an atmospheric scaling height, h_a , and the corresponding volume, V_a , so that the following holds true: The mass within this volume is equal to the mass of the atmosphere under constant pressure (1 atm, sea level). We get: $V_a = A \cdot h_a$, where A = Earth's surface area and $h_a \approx 8 \text{ km}$ = equivalent height of the atmosphere (e -folding scale of air pressure).

The following holds for the amount of oxygen in the atmosphere, N_a , and in the ocean, N_{oc} :

$$N_a = V_a \cdot c_g, \quad N_{oc} = V_{oc} \cdot s \cdot c_g \quad (6.8)$$

Therefore:

$$\begin{aligned} \frac{N_{oc}}{N_a} &= s \frac{V_{oc}}{V_a} \\ V_{oc} &= A_{oc} h_{oc} \text{ with } A_{oc} = 0.7A, \quad h_{oc} \approx 4 \text{ km}, \quad s = 0.039 \text{ at } 0^\circ\text{C} \\ \frac{N_{oc}}{N_a} &= s \frac{0.7A h_{oc}}{Ah_a} \approx 0.039 \frac{0.7 \cdot 4 \text{ km}}{8 \text{ km}} = 0.014 \end{aligned} \quad (6.9)$$

Note that the deep ocean is cold.

The amount of oxygen that would be dissolved in the ocean would be equal to $\sim 1.4\%$ of the atmospheric amount if ocean water was saturated with respect to O₂ everywhere. Due to the loss of oxygen by respiration at depth (decomposition of sinking organic particles), saturation is about 50%. Therefore, the oceanic O₂ concentration is about 0.7% of the atmospheric amount.

Gas	relative molecular mass	Bunsen's solubility, α $\frac{\text{ml}(\text{gas, STP})}{\text{l}(\text{water}) \cdot \text{atm}}$		
		0°C	24°C	24°C
He	4	7.8	7.4	$0.80 \cdot 10^{-2}$
Ne	20	10.1	8.6	$0.94 \cdot 10^{-2}$
N ₂	28	18.3	11.8	$1.28 \cdot 10^{-2}$
O ₂	32	38.7	23.7	$2.58 \cdot 10^{-2}$
Ar	40	42.1	26.0	$2.83 \cdot 10^{-2}$
Kr	84	85.6	46.2	$5.03 \cdot 10^{-2}$
Xe	131	192	99	$10.8 \cdot 10^{-2}$
CO ₂	44	1437	666	0.725

Table 6.1: Solubility in seawater (Salinity 35‰).

6.2 Uptake of anthropogenic CO₂ by the Ocean

In conjunction with the strengthening of the greenhouse effect it is necessary to estimate the future CO₂ concentrations for different emission scenarios. In order to calculate the oceanic flux, models are needed that take at least the following processes into account:

- Gas exchange between the ocean and atmosphere;
- carbonate chemistry in ocean water;
- vertical mixing of the ocean.

6.2.1 CO₂ exchange: Atmosphere – ocean

The net gas exchange from the atmosphere into the ocean is proportional to the concentration gradient between the air and surface waters. The CO₂ concentration in surface waters is expressed as the partial pressure (pCO₂; units are 1 μatm, where 1 μatm ≈ 1 ppm). The equation for the gas exchange flux per unit area in the ocean:

$$F = w \cdot \alpha(p\text{CO}_{2,\text{a}} - p\text{CO}_{2,\text{oc}}) = k(p\text{CO}_{2,\text{a}} - p\text{CO}_{2,\text{oc}}) \quad (6.10)$$

where w = transfer coefficient (with respect to the gas concentration in the liquid phase), α = Bunsen's solubility coefficient, $p\text{CO}_{2,\text{a}}$ = partial pressure of CO₂ in the atmosphere, $p\text{CO}_{2,\text{oc}}$ = effective partial pressure of CO₂ in water. k is the gas transfer velocity with regard to the gas phase.

6.2.2 Carbonate chemistry

Dissolved CO₂ reacts with H₂O to form carbonic acid, H₂CO₃, which quickly disassociates to form HCO₃⁻ and CO₃²⁻ (bicarbonate and carbonate ions). The mass law equations are as follows:



The concentration of H₂CO₃ is very small ($2 \cdot 10^{-3}$ of the CO₂ concentration) and is difficult to measure. Therefore the total concentration of dissolved CO₂ is calculated, which is defined as: $[\text{H}_2\text{CO}_3^*] = [\text{CO}_2] + [\text{H}_2\text{CO}_3] \sim [\text{CO}_2]$ ($[\text{X}]$ = concentration of X).

The concentrations are related by corresponding equilibrium constants:

First dissociation constant of carbonic acid:

$$K_1 = \frac{[\text{H}^+] \cdot [\text{HCO}_3^-]}{[\text{H}_2\text{CO}_3^*]}, \quad (6.13)$$

$$0^\circ : \quad 6.33 \cdot 10^{-7} \text{ mol/kg}$$

$$30^\circ : \quad 10.47 \cdot 10^{-7} \text{ mol/kg.}$$

Second dissociation constant of carbonic acid:

$$K_2 = \frac{[\text{H}^+] \cdot [\text{CO}_3^{2-}]}{[\text{HCO}_3^-]}, \quad (6.14)$$

$$0^\circ : \quad 6.53 \cdot 10^{-10} \text{ mol/kg}$$

$$30^\circ : \quad 10.78 \cdot 10^{-10} \text{ mol/kg.}$$

Carbonic acid (H₂CO₃) is a very weak acid: the extent of dissociation, i.e. the relative frequency of the various species, is dependent on the concentration of the hydrogen ions, H⁺, which is expressed as pH:

$$\text{pH} = -\log_{10} a_{\text{H}} \quad (6.15)$$

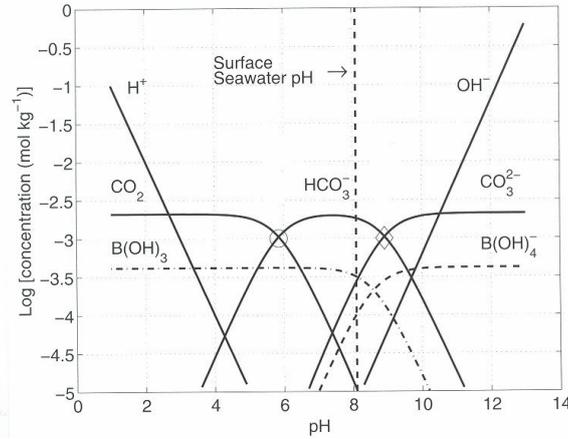


Figure 6.1: Concentrations in the carbonate system versus pH for sea water. (Source: C.M. Turley, PML, UK)

a_{H} is the concentration (activity) of the hydrogen ions. Pure water has a $\text{pH} \approx 7$, seawater: $\text{pH} \approx 8$. Solving Eq. 6.13 we find: $[\text{CO}_2] / [\text{HCO}_3^-] = [H^+] / K_1$. The higher the H^+ concentration (and lower pH), the higher the $[\text{CO}_2] / [\text{HCO}_3^-]$ ratio. The same reasoning applies for the case where $[\text{HCO}_3^-] / [\text{CO}_3^{2-}] = [H^+] / K_2$.

In an acidic environment ($\text{pH} < 7$), dissolved CO_2 comprises the largest component of $\sum \text{CO}_2$, while CO_3^{2-} is the main component under basic conditions. In oceanic surface waters the following ratio is found: $[\text{CO}_2] : [\text{HCO}_3^-] : [\text{CO}_3^{2-}] \approx 0.5 : 90 : 10$.

If water absorbs CO_2 then pCO_2 also rises, the pH drops and the $\text{CO}_3^{2-} / \text{CO}_2$ and $\text{HCO}_3^- / \text{CO}_2$ ratios become smaller (Fig. 6.1). Thus the increase in $\sum \text{CO}_2$ is not proportional to pCO_2 . The relative increase in $\sum \text{CO}_2$ is smaller than that of pCO_2 . The buffer factor (ξ) is a mathematical description of this relationship:

$$\xi = \frac{\Delta \text{pCO}_2}{\text{pCO}_{2,0}} / \frac{\Delta \sum \text{CO}_2}{\sum \text{CO}_{2,0}}, \quad (6.16)$$

$\Delta \text{pCO}_2 = \text{pCO}_2 - \text{pCO}_{2,0}$ and $\Delta \sum \text{CO}_2 = \sum \text{CO}_2 - \sum \text{CO}_{2,0}$, where $\text{pCO}_{2,0}$ and $\sum \text{CO}_{2,0}$ represent preindustrial values. ξ is determined from the carbonate system's equilibrium constants, the disassociation constant of water as well as from other acids and bases found in water.

Typical values:	surface seawater (warm)	$\sum \text{CO}_2 = 1980 \mu\text{mol/kg}$, $\xi = 9$
	surface seawater (cold)	$\sum \text{CO}_2 = 2160 \mu\text{mol/kg}$, $\xi = 14$
	mean ocean value	$\xi = 10$

As of 2019, the amount of CO_2 in the atmosphere has increased by approximately 50%. If steady state is assumed between the surface ocean and the atmosphere, $\sum \text{CO}_2$ in surface waters increased by approximately 5% ($100 \mu\text{mol kg}^{-1}$).

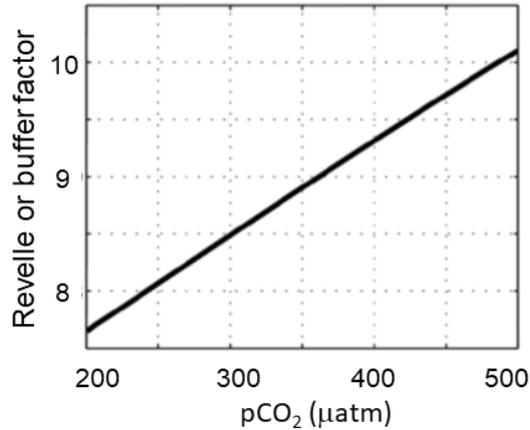


Figure 6.2: The carbon chemistry is non-linear. The Revelle Factor increases with increasing atmospheric CO₂. This means that the ocean uptake capacity is less on a percentage basis for high CO₂ emissions. Here the Revelle factor is shown for a temperature of 25°C, salinity of 35 psu, and total alkalinity of 2.3 mol/m³.

6.2.3 CO₂ uptake by the ocean on long timescales

Question: Which fraction of anthropogenic CO₂ will be taken up by the ocean after the attainment of a new steady state?

Preindustrial amount of inorganic carbon in the atmosphere and ocean:

$$N_{a,0} = 594 \text{ GtC (at 280 ppm)}; N_{oc,0} = 37,400 \text{ GtC} = 63 \cdot N_{a,0}$$

New steady state: x = relative increase in the atmosphere, $\Delta N_a = x \cdot N_{a,0}$.

Increase in the amount of $\sum \text{CO}_2$ in the ocean (ocean and atmosphere are assumed to be well mixed):

$$\frac{\Delta N_{oc}}{N_{oc,0}} = \frac{1}{\xi} \frac{\Delta N_a}{N_{a,0}} = \frac{1}{\xi} \frac{\Delta p\text{CO}_2}{p\text{CO}_{2,0}} = \frac{x}{\xi} \quad (6.17)$$

Fraction of the additional amount that remains in the atmosphere:

$$\frac{\Delta N_a}{\Delta N_a + \Delta N_{oc}} = \frac{x N_{a,0}}{x N_{a,0} + (x/\xi) N_{oc,0}} = \frac{1}{1 + \frac{N_{oc,0}}{\xi N_{a,0}}} = \frac{1}{1 + 6.3} = 14\% \quad (6.18)$$

The $\sum \text{CO}_2$ reservoir in the ocean is 60 times larger than the atmospheric reservoir, but the oceanic capacity to absorb CO₂ at steady state is only 6 times larger than that of the atmosphere. Due to the long mixing time of the deep ocean (100 to 1,000 yr) the deep ocean takes up CO₂ very slowly and a new steady state will be reached only after approximately 1000 years; thus a significant portion of CO₂ remains in the atmosphere for a long time.

This airborne fraction will be reduced on millennial time scales to about 7% by interactions with calcium carbonate sediments. The perturbation is expected to be completely removed on time scales of 100,000 years by feedbacks between CO₂ and silicate rock weathering or when natural variations in CO₂ are superimposed onto the system.

6.2.4 Tracers reveal how fast the ocean is ventilated

A number of transient or radioactive tracers such as naturally-produced radiocarbon, bomb-produced radiocarbon or chlorofluorocarbons reveal the time scales of ocean ventilation. Here, we briefly address the penetration of bomb radiocarbon into the ocean.

The atomic bomb tests in the fifties and early sixties almost doubled the atmospheric $^{14}\text{C}:^{12}\text{C}$ ratio (The total amount of C in the system was not influenced by the bomb tests as $^{14}\text{C}:^{12}\text{C} \approx 10^{-12}$). The bomb signal is clearly detectable in the upper layer of the ocean. Thus, the upper few hundred meters of the ocean exchange with the surface ocean within decades. Radioactive decay of ^{14}C is slow (mean life time of 8267 yr) and the amount of bomb-produced radiocarbon can be assumed to be constant. Bomb-produced $^{14}\text{CO}_2$ penetrates the ocean in a very similar way as anthropogenic CO₂. The $^{14}\text{C}:^{12}\text{C}$ radiocarbon distribution was systematically mapped in the seventies and early nineties.

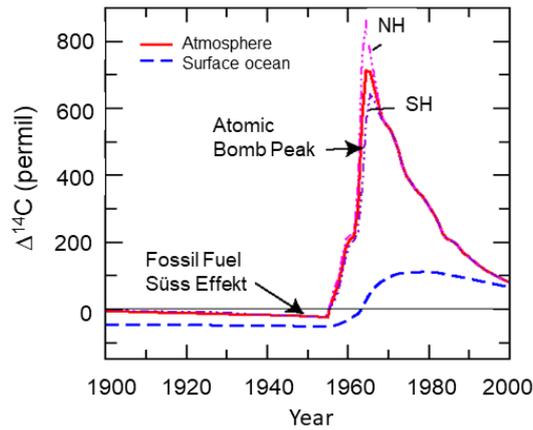


Figure 6.3: Observed signal of atomic bomb tests. The atmospheric radiocarbon ratio decreased after test ban treaties have been set in place in 1963 as the signal enters the ocean and the land biosphere.

The penetration of a perturbation into the ocean can be quantitatively expressed with the mean penetration depth, z_{pen} . Let us assume that the concentration of a tracer (e.g. bomb radiocarbon) decreases exponentially with depth z :

$$C(z) = C_{\text{surf}} e^{-\frac{z}{z_{\text{pen}}}} \quad (6.19)$$

The inventory per unit area (column inventory), I , is then:

$$I = \int_0^{\infty} C_{\text{surf}} e^{-\frac{z}{z_{\text{pen}}}} dz = C_{\text{surf}} \cdot z_{\text{pen}}. \quad (6.20)$$

We define the mean penetration depth, also for non-exponential profiles, by:

$$z_{\text{pen}} = \frac{I}{C_{\text{surf}}} \quad (6.21)$$

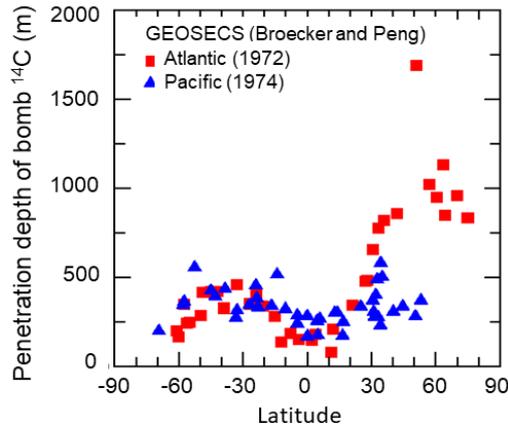


Figure 6.4: Observation-based penetration depth of bomb-produced radiocarbon in the Atlantic and Pacific Ocean at the time of the GEOSECS cruises.

The penetration of bomb-produced ^{14}C until the mid seventies was determined from data to be around 340 m. This yields a lower bound for the penetration of anthropogenic carbon as the increase in atmospheric CO_2 occurred over much longer period than the increase in the $^{14}\text{C}/^{12}\text{C}$ ratio.

GEOSECS EXPEDITION	Mean Penetration Depth of Bomb ^{14}C (m)
Atlantic (1972/3)	414
Pacific (1973-4)	311
Indian (1977-8)	309
World Ocean (1972-8)	344

Table 6.2: Average penetration depth of bomb-produced ^{14}C for different basins

6.2.5 Oceanic CO_2 uptake: A simple model

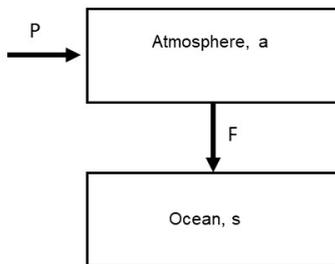


Figure 6.5: A 2-box model.

enters the ocean (net flux F). We want to calculate what the perturbation is to both reservoirs.

A portion of anthropogenic CO_2 is absorbed by the ocean. We can construct a two box model to consider the case.

The uptake of CO_2 is controlled by (1) gas exchange between the atmosphere and the ocean, (2) mixing of surface waters into the deep (more difficult to represent).

To simplify, we assume that the ocean box is well-mixed, with a thickness that is equivalent to the penetration depth of anthropogenic CO_2 . Based on model calculations, we assume this depth to be 500 m (Amount of carbon: $N_{\text{oc},0} = 5,000 \text{ GtC}$). CO_2 is produced due to human activities (production rate P) and emitted to the atmosphere; a portion

Terms:	$N_a = N_{a,0} + n_a$	Amount in the atmosphere; $n_a(t)$ = perturbation
		$N_{a,0}$ Preindustrial atmospheric content
	$N_{oc} = N_{oc,0} + n_{oc}$	Amount in the ocean
	F	net air-to-sea flux

Budgets:

$$\frac{dN_a}{dt} = \frac{dn_a}{dt} = P - F, \quad (6.22)$$

$$\frac{dN_{oc}}{dt} = \frac{dn_{oc}}{dt} = F, \quad (6.23)$$

Air-sea exchange and atmospheric residence time

To solve the two differential equations (Eq. 6.22 and 6.23), we need to express the net air-to-sea flux F as a function of the atmospheric, n_a , and oceanic perturbation, n_{oc} .

Net gas exchange flux (k = gas exchange coefficient, A_{oc} = ocean area):

$$F = A_{oc} \cdot k (\text{pCO}_{2,a} - \text{pCO}_{2,oc}) = A_{oc} \cdot k (\Delta\text{pCO}_{2,a} - \Delta\text{pCO}_{2,oc}). \quad (6.24)$$

The right hand side of the equation (after the second equal sign) is valid because the preindustrial state was approximately static: $\text{pCO}_{2,a,0} = \text{pCO}_{2,oc,0} = \text{pCO}_{2,0}$ (also $F_{0,\text{net}} = 0$). For the perturbation, the following is valid:

$$\frac{\Delta\text{pCO}_{2,a,0}}{\text{pCO}_{2,0}} = \frac{n_a}{N_{a,0}}; \quad \frac{\Delta\text{pCO}_{2,oc}}{\text{pCO}_{2,0}} = \xi \frac{n_{oc}}{N_{oc,0}} \quad (6.25)$$

Also:

$$F = A_{oc} \cdot k \left(\frac{n_a}{N_{a,0}} \text{pCO}_{2,0} - \xi \frac{n_{oc}}{N_{oc,0}} \text{pCO}_{2,0} \right) = F_0 \left(\frac{n_a}{N_{a,0}} - \xi \frac{n_{oc}}{N_{oc,0}} \right). \quad (6.26)$$

$F_0 = A_{oc} \cdot k \text{pCO}_{2,0}$ is the preindustrial gross gas exchange flux, where

$$\tau_{a,oc} = \frac{N_{a,0}}{F_0} \approx 7.5 \text{ yr} \quad (6.27)$$

$$F = \frac{1}{\tau_{a,oc}} \left(n_a - \xi \frac{N_{a,0}}{N_{oc,0}} n_{oc} \right) \quad (6.28)$$

$\tau_{a,oc}$ is the average lifetime of a CO₂ molecule in the atmosphere with respect to exchange with the ocean at preindustrial time.

Airborne fraction

We can solve now the system of the differential equations (Eq. 6.22 and 6.23) numerically for any emission P . In the following, an analytical solution is presented for exponentially increasing emissions, roughly approximating the real emission history. The production rate, $P(t)$, increases exponentially with time:

$$P(t) = P_1 \cdot e^{\mu t}, \text{ where } \mu \approx 1/(30\text{yr}) \quad (6.29)$$

The same time dependence is also valid for the perturbation: $n_a = n_{a1}e^{\mu t}$, $n_{oc} = n_{oc1}e^{\mu t}$, because the differential equations are linear. Substituting in Eq. 6.23 Eq. 6.28 yields:

$$\begin{aligned} \frac{dn_{oc}}{dt} = F : \quad \mu n_{oc} &= \frac{1}{\tau_{a,oc}} \left(n_a - \xi \frac{N_{a,0}}{N_{oc,0}} n_{oc} \right), \\ \text{or:} \quad n_{oc} &= \frac{n_a}{\tau_{a,oc}\mu + \frac{N_{a,0}}{N_{oc,0}}\xi} \end{aligned} \quad (6.30)$$

Numerical:

$$\begin{aligned} N_{a,0} = 594 \text{ GtC}, N_{oc,0} = 5000 \text{ GtC}, \mu = 1/(30\text{yr}), \tau_{a,oc} = 1/(7.5\text{yr}), \xi = 10 \\ \text{leads to } n_{oc} = 0.70n_a. \end{aligned} \quad (6.31)$$

Relationship to production: The sum of the perturbations is equal to the total amount of CO₂ produced (= cumulative production):

$$P_{cum} = \int_{-\infty}^t P dt : n_a + n_{oc} = P_{cum}, \quad n_a(1 + 0.7) = P_{cum} \quad (6.32)$$

It follows

$$n_a = 0.59P_{cum}, \quad n_{oc} = 0.41P_{cum} \quad (6.33)$$

Our simple model predicts an airborne fraction of anthropogenic CO₂ equal to 59%.

Chapter 7

Oxygen in the climate system and links between O₂ and CO₂

7.1 Observed links between atmospheric O₂ and CO₂

The cycles of oxygen and carbon are coupled through photosynthesis and respiration. During photosynthesis (respiration) for every mol CO₂ assimilated (respired) on average about 1.1 mol of O₂ is released (consumed). Correspondingly, for every mol CO₂ released by autotrophic or heterotrophic respiration and by fire, about 1.1 mol of O₂ is consumed. The land biosphere at present is no, or only a small, net producer of O₂ since photosynthesis, respiration and fire fluxes balance one another over the course of a year. The natural oxygen cycle is stationary in an unperturbed state. Large-scale forest clearing has a relatively weak influence on the

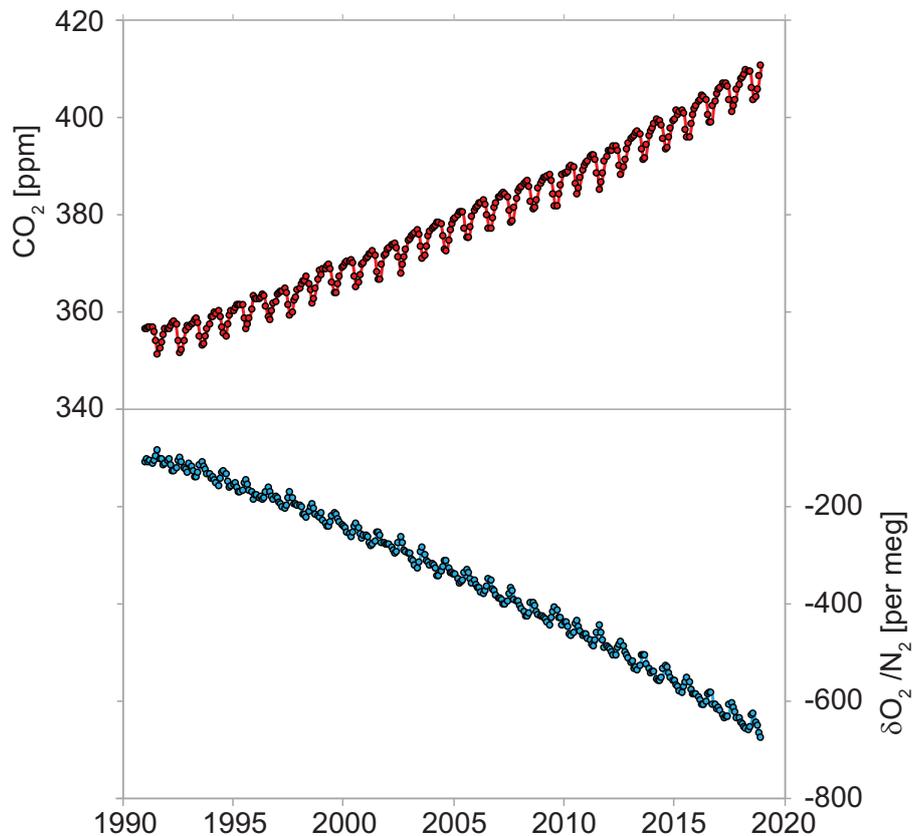


Figure 7.1: Anticorrelated seasonal fluctuations in O₂ and CO₂ are superimposed on the anthropogenic rise in CO₂ and the associated decrease in O₂. Data are from Scripps Institute of Oceanography, La Jolla, USA. For methods see *Keeling et al.* [1996].

oxygen cycle, but has a very strong influence on the carbon cycle (i.e. clearing and burning). The seasonal variation in atmospheric O₂ is phase-shifted by 180° relative to CO₂. A 10 ppm seasonal variation in CO₂ results in a seasonal variation in O₂ of ≈ 11 ppm (out of a total of 210,000 ppm in the atmosphere, since 21% of air is in the form of oxygen). It is challenging to measure such small signals and reliable measurements are available only since the end of the 1980ies. These measurements demonstrate that the O₂ content of the atmosphere is declining due to fossil fuel burning and provide information on the magnitude of the terrestrial carbon sink.

7.2 Overturning time of atmospheric O₂ in the biosphere

The amount of O₂ in the atmosphere is estimated to be $3.8 \cdot 10^{19}$ mol O₂. The amount of O₂ in the ocean is about 2 orders of magnitude smaller, approximately $3 \cdot 10^{17}$ mol O₂. The concentration of O₂ is much smaller in the ocean because O₂ is not very soluble in ocean water. The gross O₂ production is estimated from the GPP of the land and marine biosphere and by assuming that the assimilation of one mol of CO₂ corresponds to the release of about one mol of O₂.

GPP (land)	2 NPP	$1.2 \cdot 10^{17}$ gC/a	$1 \cdot 10^{16}$ molC/a
GPP (marine)		$1.2 \cdot 10^{17}$ gC/a	$1 \cdot 10^{16}$ molC/a
Total			$2 \cdot 10^{16}$ molC/a

The average residence time of oxygen in the atmosphere with respect to exchange with the marine and terrestrial biosphere is calculated as follows:

$$\tau = \frac{\text{Inventory}}{\text{GPP (land + marine)}} = \frac{3.8 \cdot 10^{19} \text{ mol}}{2 \cdot 10^{16} \text{ mol/yr}} \approx 2000 \text{ yr} \quad (7.1)$$

Atmospheric oxygen and oxygen isotopes exchange with the land biosphere through photosynthesis and respiration on a time scale of 2,000 yr.

7.3 How much O₂ can be consumed by fossil fuel burning?

What happens when the total organic carbon (biosphere + fossil fuels) are oxidized? Does this have an effect on the atmospheric oxygen content? Lets compare the inventories:

Vegetation + Soil	3000 GtC	$2.6 \cdot 10^{17}$ molC
Fossil fuels	5000 GtC	$4.6 \cdot 10^{17}$ molC
Total C		$7.2 \cdot 10^{17}$ molC
Atmospheric O ₂		$3.8 \cdot 10^{19}$ mol O ₂

The atmosphere contains approximately 50 times more O₂ than the amount of oxidizable carbon in the biosphere and estimated conventional fossil fuel reserves. If all of this carbon were burned, the O₂ concentration would decrease by ≈ 1/50, i.e. the percent of O₂ in dry air would decrease from 20.95% to 20.55%. Therefore the destruction of biomass and oxidation of conventional fossil fuel resources does not really impact the atmospheric oxygen content.

7.4 How fast can the atmospheric O₂ content vary naturally?

Most of the O₂ was created due to the production of organic matter by photosynthesis. But free O₂ only remains when organic material is not oxidized; for every mol of O₂ that remains free, 1 mol C_{org} (organic carbon) must enter the sediment pool unoxidized.

The flux of organic carbon to ocean sediments is estimated to be around 1 to 2 · 10¹³ mol yr⁻¹

Regeneration time of atmospheric oxygen (net production):

The number of mol O₂ that is produced (net) corresponds to the global sedimentation of organic material and we obtain for the regeneration time:

$$\tau_{\text{net}} = \frac{\text{Inventory}}{\text{Netflux}} = \frac{3.8 \cdot 10^{19} \text{ mol}}{1.5 \cdot 10^{13} \text{ mol/yr}} \approx 310^6 \text{ yr.} \quad (7.2)$$

To accurately calculate the net balance the O₂ sinks must be accounted for, i.e. the oxidation of Fe, S and old sedimentary C_{org}. Therefore the net flux is smaller and τ(O₂) larger than 10⁶ yr. On the timescale of 10⁶ – 10⁷ years the atmospheric oxygen content is approximately constant.

7.5 Origin of free oxygen

Earth's atmosphere was originally reducing, and therefore did not contain O₂. Initially it most likely consisted of CH₄, NH₃, N₂, H₂, H₂O and noble gases, later also CO₂. At first O₂ accumulated due to photodissociation of water vapour by UV radiation (λ < 200 nm):



Hydrogen gas (H₂) escaped from Earth's gravitational field into the universe. The free oxygen was quickly bound due to the oxidation of Fe and S.

Compared to photodissociation, later much more O₂ was created due to the production of organic matter by photosynthesis. But free O₂ only remains when organic material is not oxidized; for every mol of O₂ that remained free, 1 mol C_{org} (organic carbon) must have entered the sediment pool unoxidized.

7.5.1 Development of atmospheric oxygen

- Photosynthesis has been occurring for more than 3 · 10⁹ yr. Initially, photosynthesis took only place in the ocean since UV radiation was life threatening on land due to the lack of an ozone layer. The partial pressure of O₂ remained low for a long time.
- Clues from sediments: Pyrite (FeS₂) could only form in the absence of O₂ (otherwise sulphate, SO₄ is formed).
- In total approximately 10²¹ mol O₂ have been produced due to photosynthesis. A large portion has been utilized to oxidize iron (to Fe₂O₃) and sulphur (to sulphate, SO₄²⁻). Only a few percent are found as free O₂ in the atmosphere (3.8 · 10¹⁹ mol). Corresponding

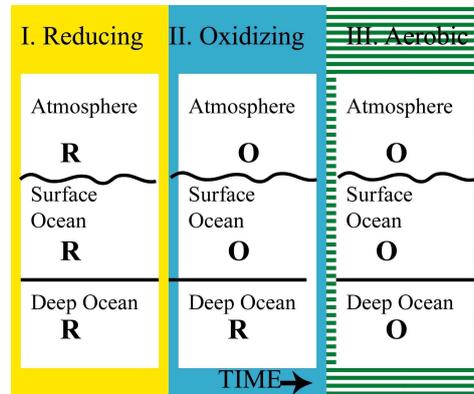


Figure 7.2: Conceptual explanation of the origin of oxygen in the atmosphere. The initial production occurred in the protected ocean surface and then entered the atmosphere via gas exchange and the deep ocean via mixing. R. Reducing environment; O: Oxidizing environment

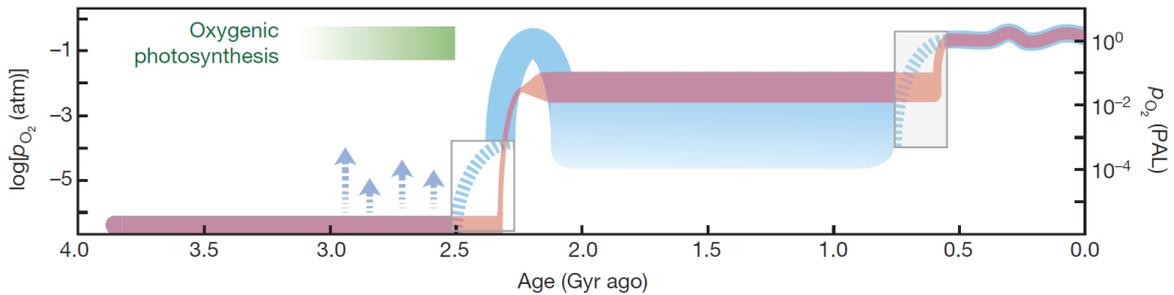


Figure 7.3: Evolution of atmospheric oxygen content through time. The faded red curve shows a "classical, two-step" view of atmospheric evolution, while the blue curve shows the emerging model (pO_2 , atmospheric partial pressure of O_2). Right axis, pO_2 relative to the present atmospheric level (PAL); left axis, $\log pO_2$ [Lyons *et al.*, 2014].

to the amount of O_2 produced by photosynthesis, 10^{21} mol of organic material (C_{org}) must be stored in sediments.

- It appears that the atmospheric partial pressure of O_2 has remained constant over the past 60 million years (since the Tertiary).

Sediment deposits that contain iron oxides appeared approximately two billion years ago. Phytoplankton remains are also recognizable in sediments shortly after this time, indicating that the O_2 content of the atmosphere rose relatively fast.

7.6 The anthropogenic CO_2 perturbation and O_2 changes

The concentration of CO_2 has risen from a pre-industrial value of 280 ppm (prior to 1800) to today's value of approximately 414 ppm, an increase of over 30%. This increase can be attributed to the burning of fossil fuels (coal, oil, gas), as well as the clearing and burning of forests and other land use change activities. Between 1860 and World War I, as well as

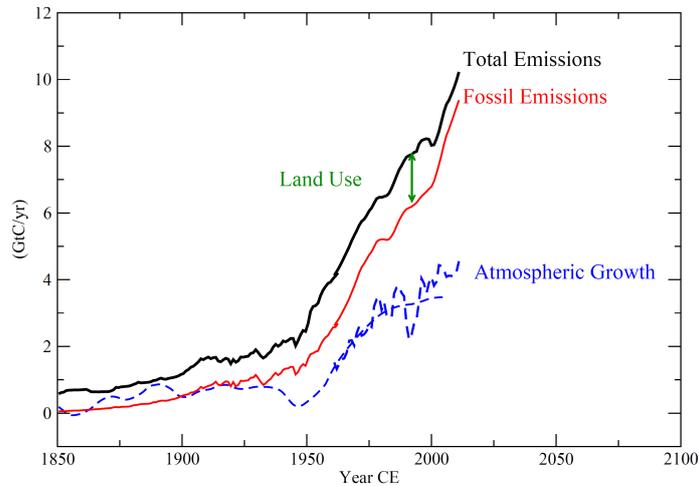


Figure 7.4: Emissions due to the burning of coal, oil and gas, deforestation and land use exceed the increase in the atmospheric inventory by about a factor of two. 5-yr running means are shown after 1961 for atmospheric growth. (updated from [Joos *et al.*, 1999])

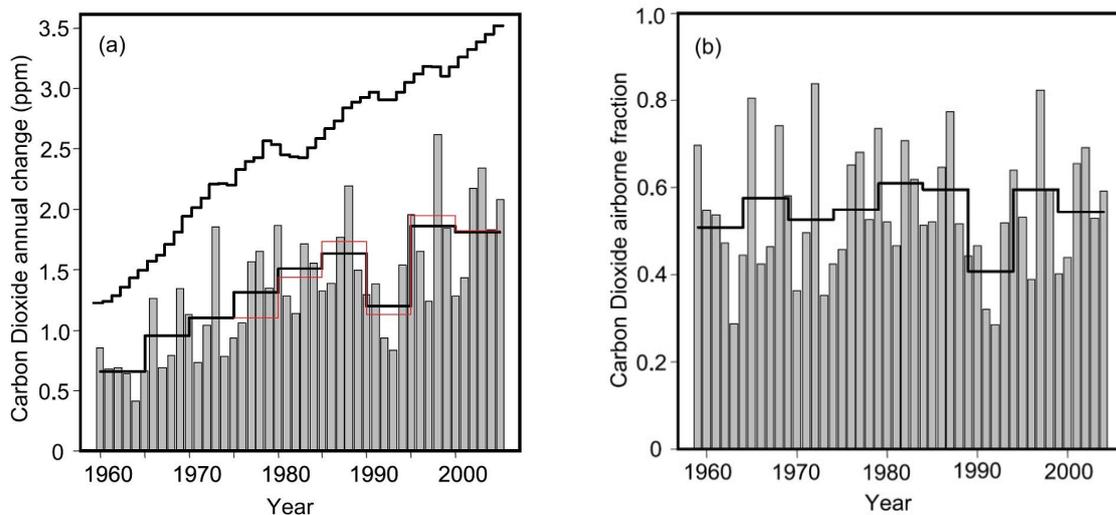


Figure 7.5: Left: Comparison of annual fossil fuel emissions (black line) and annual (bar) and five-year mean changes in the atmospheric CO₂ inventory. Right: Fraction of fossil emissions that remain in the atmosphere (source: [Denman *et al.*, 2007])

after World War II until the present, CO₂ emissions due to fossil fuel burning have risen exponentially. This continuous growth has been interrupted during the two world wars, the Great Depression, the oil crisis in 1973 and the period of higher oil prices since 1980. Today the world-wide emissions of fossil fuels are over 10 GtC/yr. Deforestation in the tropics and land use (e.g. ploughing leads to accelerated oxidation of the soils) are also major sources of CO₂. Estimates of carbon emissions due to anthropogenic land use range between 0.5 – 2.7 GtC/yr for recent decades. It is difficult to estimate land use emissions due to the large spatial and temporal heterogeneity and variability of the biosphere. Open questions include: Which areas

are affected? What is the carbon density per unit area? What is the effect of regrowth on cleared areas?

Currently the total carbon emissions are around 10 GtC/yr. If all of the emitted CO₂ stayed in the atmosphere, the atmospheric concentration of CO₂ would rise by approximately 5 ppm per year (1 ppm = 2.12 GtC). The observed increase is only about half as large. The ocean and the land biosphere have taken up the other half. The 50% that stays in the atmosphere is referred to as the airborne fraction. The estimated oceanic CO₂ absorption today is about 2 GtC/yr and is somewhat smaller than the difference between emissions and the atmospheric growth. This implies that the land biosphere is responsible for taking up a large amount of carbon.

The airborne fraction exhibits considerable interannual variability. This may be caused by variations in the strength of land use emissions and variations in the global air-to-land and air-to-ocean net carbon fluxes. The air-to-biosphere net flux exhibits much larger interannual variability than the air-to-sea net flux. Usually more carbon is released during so called El Niño/Southern Oscillation (ENSO) events when human land use activities and dry climatic conditions cause large release of carbon by forest fires. An example are the Indonesian fires in 1998.

The carbon budget: Oceanic models are used to calculate the strength of the oceanic sink. The land sink is then calculated as a difference from the atmospheric balance equation. The equation is as follows:

$$\frac{dN_a}{dt} = E - F_{a,oc,net} - F_{a,b,net}, \quad (7.4)$$

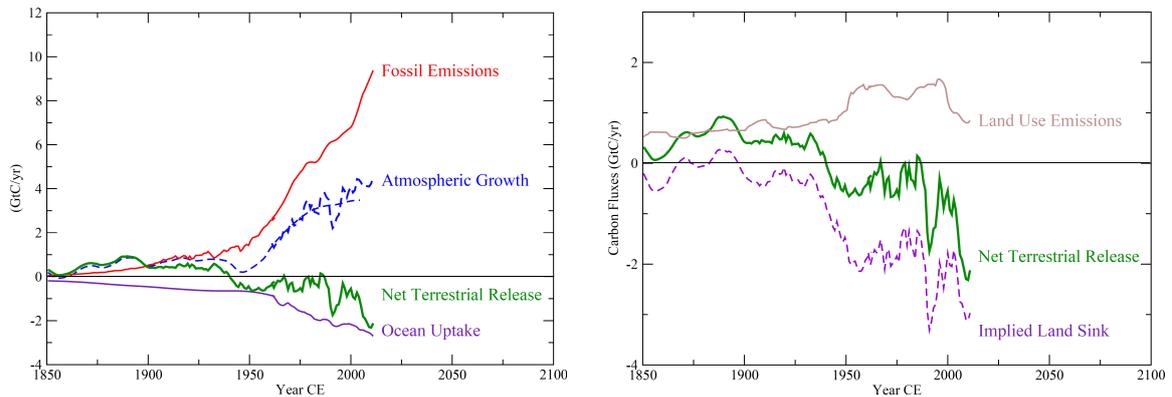


Figure 7.6: Anthropogenically caused fluxes between the various carbon reservoirs. Right panel: The fossil fuel emissions are deduced from the consumption and supply rate of oil, coal and gas, while the increase in the atmospheric C inventory is directly measured. The ocean fluxes are calculated with a model. The terrestrial release, $F_{a,b,net}$, is computed as difference using the atmospheric mass balance equation. Left panel: The difference between the release of CO₂ due to deforestation and land use and the net flux into the biosphere implies a global terrestrial carbon sink (i.e. $F_{sink} > F_{def}$). The magnitude of this sink is poorly constrained because the land use emissions are not well known. Possible causes for this implied sink are fertilization of plants by CO₂ or nitrogen, the regrowth of forests in the northern hemisphere and climate change. Data represent 5-yr running means after 1961 and decadal smoothed values before (updated from [Joos *et al.*, 1999]).

where $F_{a,i,net}$ signifies the net carbon flux from the atmosphere into reservoir i and N_a represents the atmospheric inventory. The total emissions, E , are equal to the sum of the change in the carbon inventory of the atmosphere, ocean and the land biosphere.

The net flux in the biosphere is the difference between the release of CO₂ due to deforestation and land use, F_{def} , and an implied carbon sink, F_{sink} :

$$\frac{dN_b}{dt} = F_{a,b,net} = F_{sink} - F_{def}. \quad (7.5)$$

The carbon and oxygen budget: The net fluxes in the ocean and the biosphere can also be estimated from the measured decrease in atmospheric oxygen content and the rise in atmospheric CO₂. The burning of fossil fuels results in an increase in CO₂ and a decrease in O₂ in the atmosphere. The oceanic uptake of CO₂ does not result in any change in atmospheric oxygen content, while for every mol of carbon stored by the land biosphere, 1.1 mol of O₂ is released. If the outgassing of oxygen from the ocean due to the warming of surface waters is taken into account, the budget for the 1990s is as shown in Table 7.1.

Measuring changes in atmospheric O₂ is a formidable challenge as changes are small compared to the atmospheric O₂ content. The ratio O₂/N₂ is measured to achieve the required precision. Changes in the O₂/N₂ ratio are governed by changes in the O₂ concentration as the atmospheric N₂ concentration is almost constant. Usually, measurements are expressed relative to a standard in units of "per meg":

$$\delta \left(\frac{O_2}{N_2} \right) = \left(\frac{(O_2/N_2)_{sample}}{(O_2/N_2)_{reference}} - 1 \right) \cdot 10^6 \text{ (per meg)} \quad (7.6)$$

4.8 per meg are equivalent to 1 ppm; the oxygen content of the air is 20.95 % and $100/20.95 = 4.8$. We write a budget equation for atmospheric O₂ that corresponds to the budget equation for atmospheric CO₂. The following budget equation is approximately valid for *decadal* time scales:

$$\frac{dO_2}{dt} = -\alpha_f E + \alpha_b F_{a,b,net} \quad (7.7)$$

E und $F_{a,b,net}$ denote fossil emissions and the net carbon flux into the land biosphere. α_f and α_b are stoichiometric ratios that links O₂ consumption and CO₂ release for the process of burning fossil fuel, photosynthesis and organic matter respiration on land. Possible net air-sea fluxes of oxygen in response to changes in ocean circulation, temperature and the marine biological carbon cycle are neglected in the above equation. Rearranging yields the net air-to-land biosphere flux:

$$F_{a,b,net} = \frac{1}{\alpha_b} \left(\frac{dO_2}{dt} + \alpha_f E \right) \quad (7.8)$$

The net air-to-sea flux can now be determined from the atmospheric CO₂ budget and we get:

$$F_{a,oc,net} = E - \frac{dN_a}{dt} - F_{a,b,net}. \quad (7.9)$$

Box 7.1 Increase in CO₂ by human activities?**Is the increase in atmospheric carbon dioxide during the industrial period caused by human activities?**

Yes, there are a number of observational constraints that clearly show that the recent CO₂ increase is man made:

- During each decade since 1900, CO₂ emitted by fossil fuel burning is larger than the atmospheric growth. Since the beginning of industrialization, emissions from land use activities and fossil emissions are larger than the atmospheric increase. The observed atmospheric CO₂ increase accounts only for 55% of the carbon released by human activities since 1959. The rest has been taken up by the ocean and the land biosphere.
- The ratio of heavy (¹³C, ¹⁴C) to light (¹²C) isotopes has changed in a way that is only consistent with a fossil carbon source and a source from the land biosphere. ¹³C is depleted in organic carbon as the light isotope ¹²C is preferentially assimilated during photosynthesis. ¹⁴C is not present in old, fossil carbon as the ¹⁴C that was initially present is decayed with a mean life time of 8200 years. Both, the ¹³C/¹²C and the ¹⁴C/¹²C ratio have decrease in the atmosphere mainly as a result of the release of isotopically light fossil carbon. A release of carbon from the ocean or from volcanic activities, rocks or geothermal sources would have increased the ¹³C/¹²C isotopic ratio.
- The ratio of oxygen to nitrogen in the atmosphere has declined as expected from oxidation of fossil or organic material.
- Measurements of the partial pressure of CO₂ in surface water and of carbon and other tracers within the ocean yield an uptake of carbon by the ocean (and not a release).
- The North-South interhemispheric gradient has increased in parallel with fossil emissions, as the fossil sources are mainly over land in the Northern Hemisphere.
- CO₂ concentration varied only within 20 ppm during the past ten thousand years before the onset of the industrialization and only within 10 ppm during the last millennium. This suggests that the carbon cycle was close to equilibrium before the industrialization started.

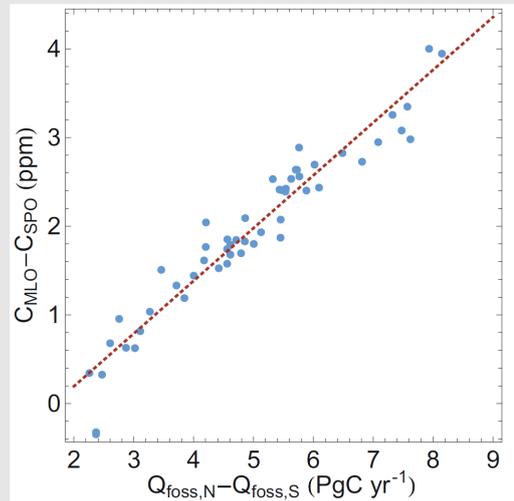


Figure: North-South Gradient in atmospheric CO₂ versus fossil fuel emissions. Blue points: Annually averaged CO₂ concentration difference between the station Mauna Loa in the Northern Hemisphere and the station South Pole in the Southern Hemisphere versus the difference in fossil fuel combustion CO₂ emissions between the hemispheres. Dark red dashed line: regression line fitted to the data points ([Ciais *et al.*, 2013])

Chapter 8

Ocean: Vertical structure, water masses and sea level rise

8.1 Inventories and fluxes

The flux of water in the climate system is of key importance for the transport of energy and the definition of regional climate. Here we follow the approach introduced in section 1 and consider the four most important inventories for mobile water on Earth (Fig. 8.1). With mobile we mean that this water can be transported between these inventories within time scales shorter than 10^6 years. This excludes the large masses of water which are enclosed in rock and sediments.

Figure 8.1 shows the inventory and fluxes of mobile water in the climate system. The numbers indicate that the ocean is by far the largest reservoir with a mean residence time of the water of over 3000 years. The mean residence time in the atmosphere, on the other hand, is only about 11 days, because the inventory is much smaller relative to the substantial exchange fluxes.

Evaporation and precipitation are strongly dependent on the geographical location. On very large scales, excess of precipitation is found in the tropics and middle to high latitudes while the subtropics exhibit excess in evaporation. The fresh water balance can also be determined for entire ocean basins. Here, the salt content of the transported ocean water plays a crucial role (see section 8.3). For example, estimates of the net water balance of the Atlantic Ocean

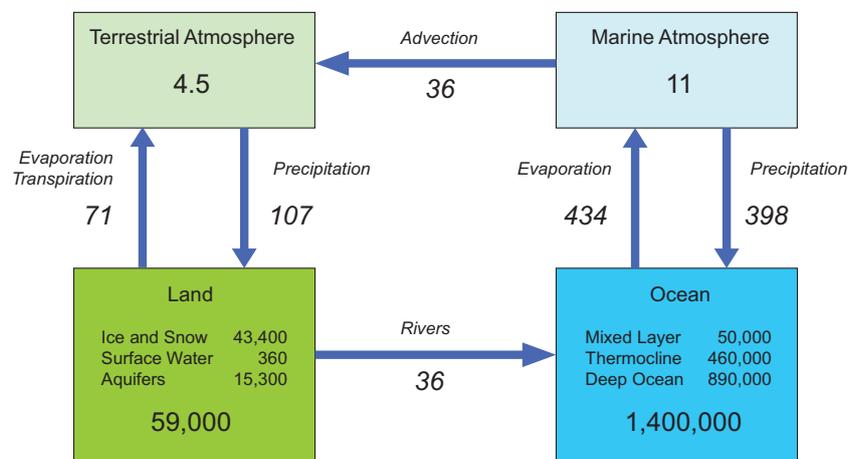


Figure 8.1: Inventories (in 10^{15} kg) and fluxes (*italics*, in 10^{15} kg/yr) of water in the climate system according to estimates by *Chahine* [1992].

basin give values that range from 0.13 to 0.45 Sv (1 Sverdrup = 1 Sv = $10^6 \text{ m}^3/\text{s} = 31.61 \cdot 10^{15} \text{ kg/yr}$). Hence, the atmospheric branch of the hydrological cycle must export water from the Atlantic to the Pacific/Indian ocean basins. This freshwater loss of the Atlantic Ocean has profound consequences for the water mass structure and general circulation of the world ocean. As a result of this freshwater transfer, on ocean basin scales the Atlantic is the most salty ocean, while the Pacific Ocean becomes relatively fresher.

8.2 Density of ocean water

Sea water contains a range of dissolved ions which affect the density of sea water. In the global mean, 1 kg of sea water contains about 19.4 grams of Cl^- , 10.8 g Na^+ , 2.7 g SO_4^- , 1.3 g Mg^{2+} , and many other ions which add up to a global mean "salt" content of 34.73 g/kg water. This is called salinity and reported in "practical salinity units", psu or 1.

In addition to the classical thermodynamic state variables of temperature and pressure, the density of ocean water also strongly depends on salinity. For ocean water at the surface (i.e., pressure $p = 1000 \text{ mbar}$) the density can be approximated by

$$\rho(T, S) = \rho_0 (1 - \alpha(T - T_0) + \beta(S - S_0) + \gamma(T - T_0)^2), \quad (8.1)$$

where the coefficients are given by

$$\begin{aligned} \rho_0 &= 1028 \text{ kgm}^{-3}, & \alpha &= 5.4128 \cdot 10^{-5} \text{ K}^{-1}, \\ T_0 &= 0^\circ\text{C}, & \beta &= 7.623 \cdot 10^{-4}, \\ S_0 &= 35, & \gamma &= -5.0804 \cdot 10^{-6} \text{ K}^{-2}. \end{aligned} \quad (8.2)$$

These coefficients were determined by a least-square fit to the full equation of state for the range of -2 to 30°C and 33 to 38 psu, which are the representative ranges for ocean sea water. The deviations in this range are smaller than 0.2 kg m^{-3} . The T - S diagram using the full equation of state for the ocean surface is shown in Fig. 8.2. Generally, the density increases with increasing salinity and decreasing temperature. The highest density is achieved at the freezing point if the salinity is larger than about 24.7.

The freezing point of sea water depends on salinity, approximately as

$$T_f(S) = (-0.0575S + 1.710525 \cdot 10^{-3}S^{1.5} - 2.155 \cdot 10^{-4}S^2)^\circ\text{C}. \quad (8.3)$$

Up to now we have considered only density changes at the surface. However, as one goes down in the water column, the pressure rises rapidly and the compressibility of the water starts to play a role. The compression and expansion of sea water has an effect on temperature by warming when compressing and cooling when expanding (just like a gas). This process is called adiabatic if no extra heat is added or taken away from the water parcel. Adiabatic changes of temperature must be accounted for when comparing water from different depths of the ocean. This is done by calculating the potential temperature referred to a selected depth level, e.g., the surface, or 2000 m depth. Potential temperature is about 0.1°C colder than in situ temperature for every 1000 m depth. From this, the potential density can be

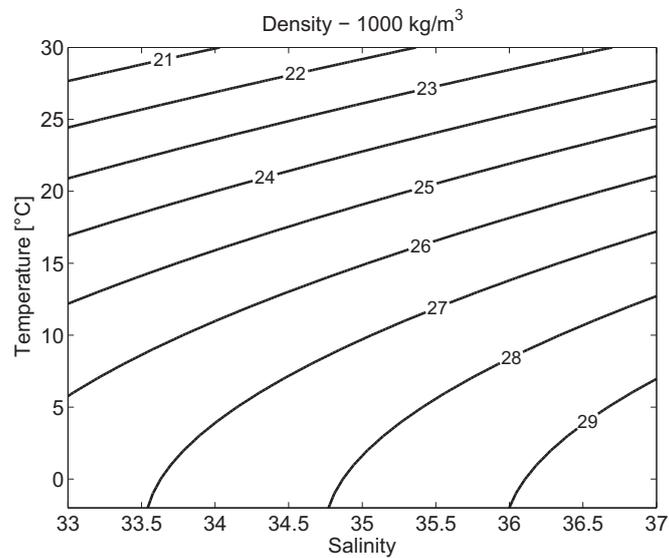


Figure 8.2: Density minus 1000 kg/m^3 using the full equation of state for the typical range of ocean water at the surface in the T-S diagram. Note the curvature of the isopycnals (= lines of constant density) that is due to the important non-linearity of the equation of state. The density was calculated using the Matlab routine developed by Phil Morgan (CSIRO, 1992).

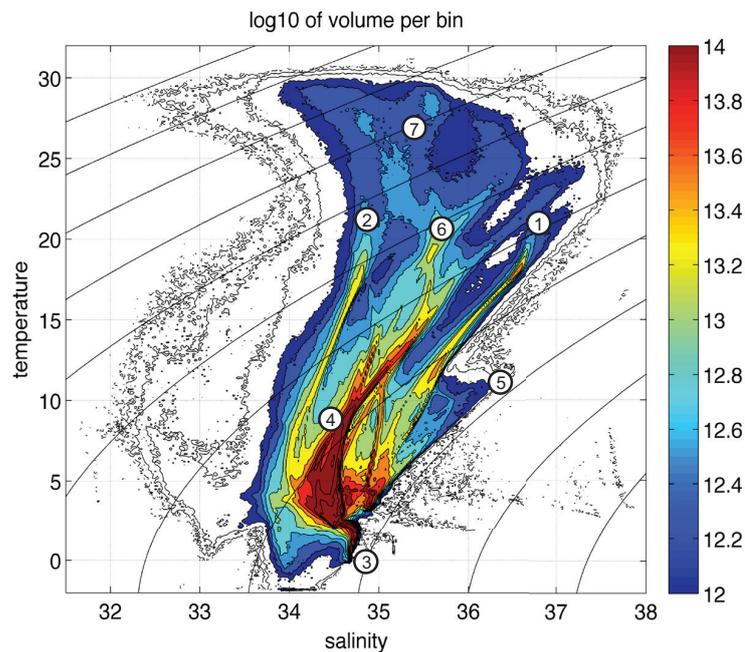


Figure 8.3: Global upper ocean ($< 2000 \text{ m}$) volumetric water mass distribution (in log of m^3 per $0.1^\circ\text{C} \times 0.025 \text{ psu}$ bin) based on measurements by Argo floats. Mixing of water masses occurs along the "ridges" visible in the contour patterns. The circled numbers denote the geographical location of the major water masses: 1 North Atlantic, 2 North Pacific, 3 Antarctic Circumpolar, 4 Subantarctic, 5 Mediterranean outflow, 6 South Pacific, 7 Tropics. Figure from *Speer and Forget* [2013].

determined. The details of this are, however, beyond the scope of this course.

A census of ocean water is shown in Fig. 8.3. It is based on millions of measurements by autonomous floats and provides a distribution of water masses that are characterized by their temperature T and salinity S . Mixing of water occurs along the ridges that are visible in this figure.

An earlier analysis of deep water clearly shows the convergence of water masses in the deep Southern Ocean (Fig. 8.4). Roughly three mixing lines are visible which exhibit the three major pathways of water in the deep ocean and which meet at $T \approx 2^\circ\text{C}$ and $S \approx 34.7$ where most of the water resides in the T - S diagram. The three mixing lines point to the areas of water formation around Antarctica ($T \approx -2^\circ$), the North Atlantic ($S \approx 35$) and the North Pacific ($S \approx 34.4$).

8.3 Vertical structure and water masses

Density in the ocean generally increases with depth; this results in a stable stratification (Fig. 8.5). This is primarily due to temperature which decreases in most locations with depth. This is not the case in high latitudes where changes in salinity are more dominant in determining density (see Fig. 8.2). The warmest temperatures are measured at the surface in

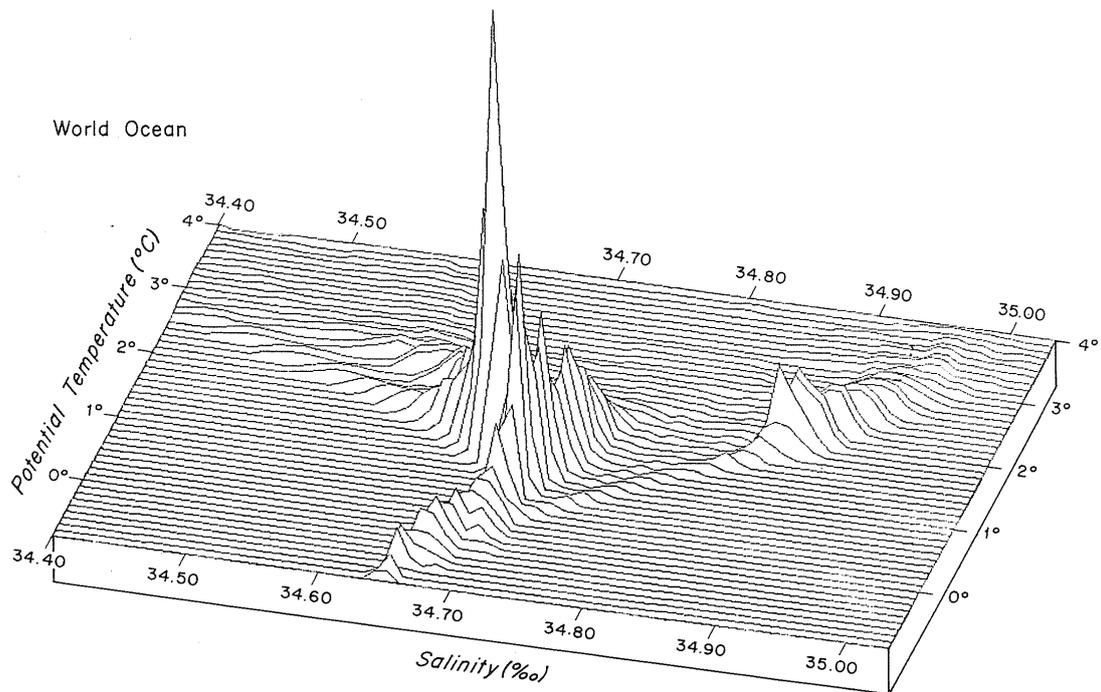


Figure 8.4: Relative occurrence of deep water masses colder than 4°C in the world ocean. Roughly there are two mixing lines with positive and one with negative slope in the T - S diagram meeting at the centre. This dominating water mass is joined by another mixing line which defines the coldest water. The outer endpoints of the mixing lines characterise waters from Antarctica ($T = -1.8^\circ\text{C}$), from the deep North Atlantic ($S = 35$) and the deep North Pacific ($S = 34.4$). Figure from *Worthington* [1981]

the low latitudes due to solar radiation and the positive energy budget there. Therefore, there is a substantial flux of heat into the ocean in these areas. In order to maintain the vertical temperature profile in the presence of a downward heat flux due to diffusion, an upward heat flux caused by slow advection must be postulated. This will be the basis of a simple but very instructive 1-dimensional model explaining the global mean vertical structure of the ocean.

Munk [1966] has used vertical profiles of measured temperature and radiocarbon concentrations (Fig. 8.6) in the Pacific Ocean to estimate the large-scale vertical diffusivity and upwelling velocity. He assumed a one-dimensional advective-diffusive water column in which the conservation equation for temperature is given by:

$$\frac{\partial T}{\partial t} + w \frac{\partial T}{\partial z} = K_v \frac{\partial^2 T}{\partial z^2} \quad , \quad (8.4)$$

where z is the vertical coordinate, positive upward, T is temperature, K_v is a vertical eddy diffusivity, and w is the advection velocity. The general stationary solution of Eq. 8.4 is

$$T(z) = T_0 + (T_1 - T_0)e^{z/D} \quad , \quad (8.5)$$

where $D = K_v/w$ is a characteristic depth scale which can be estimated by a least-square fit of the observed temperature profile to Eq. 8.5. This yields $D \approx 940$ m.

In order to determine the diffusivity and vertical advection velocity separately, a second, independent tracer needs to be considered. Salinity would not work because on large scales it is mixed by the same processes as temperature. Radiocarbon, however, provides independent information because, in addition to the mixing, also natural decay of the ^{14}C atom occurs with a half-life of $\tau = 5730$ yr. This decay process represents a sink term which needs to be considered in the conservation equation for radiocarbon concentration C :

$$\frac{\partial C}{\partial t} + w \frac{\partial C}{\partial z} = K_v \frac{\partial^2 C}{\partial z^2} - \lambda C, \quad (8.6)$$

and $\lambda = \ln 2/\tau = 1/(8267 \text{ yr})$ is the decay constant of radiocarbon. The stationary solution of Eq. 8.6 is given by

$$C(z) = C_0 e^{\mu z/D}, \quad (8.7)$$

where

$$\mu = \frac{1}{2} \left(1 + \sqrt{1 + \frac{4\lambda K_v}{w^2}} \right) \quad . \quad (8.8)$$

The non-dimensional factor μ is estimated by a least-square fit to observed radiocarbon profiles such as those given in Fig. 8.6.

By solving Eq. 8.8, Munk obtained $K_v/w^2 \approx 200 \text{ yr}$. The two results permit the estimation of both the vertical eddy diffusivity and the upwelling velocity:

$$K_v = 1.3 \cdot 10^{-4} \text{ m}^2\text{s}^{-1} \quad , \quad w = 4.4 \text{ myr}^{-1} \quad . \quad (8.9)$$

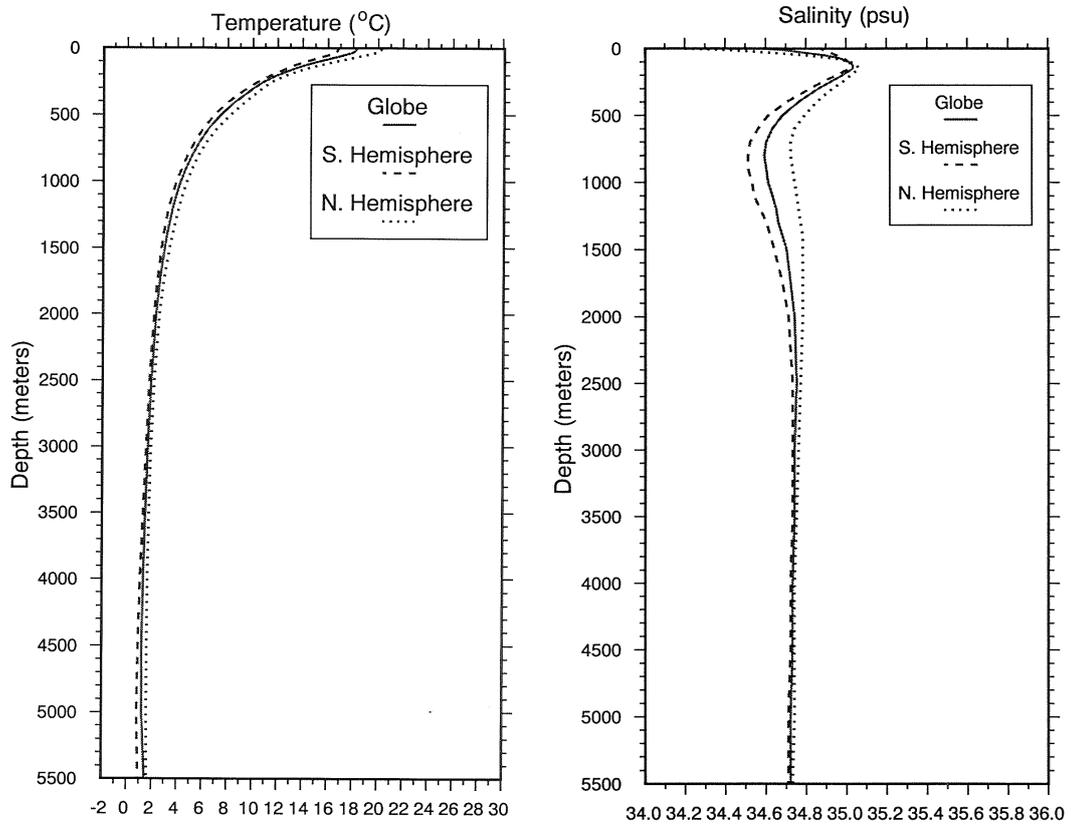


Figure 8.5: Global mean profiles of temperature and salinity. Figure from *Levitus and Boyer* [1994]

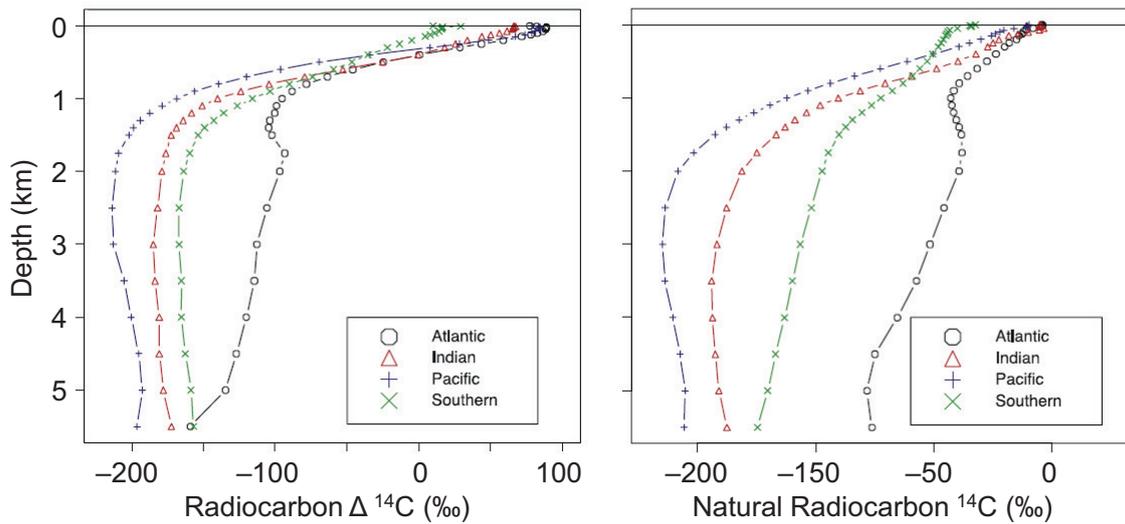


Figure 8.6: Vertical profiles of measured radiocarbon (left) which includes the contribution from anthropogenic radiocarbon from bomb tests and derived natural radiocarbon (right) for the major ocean basins. Figure from *Key et al.* [2004].

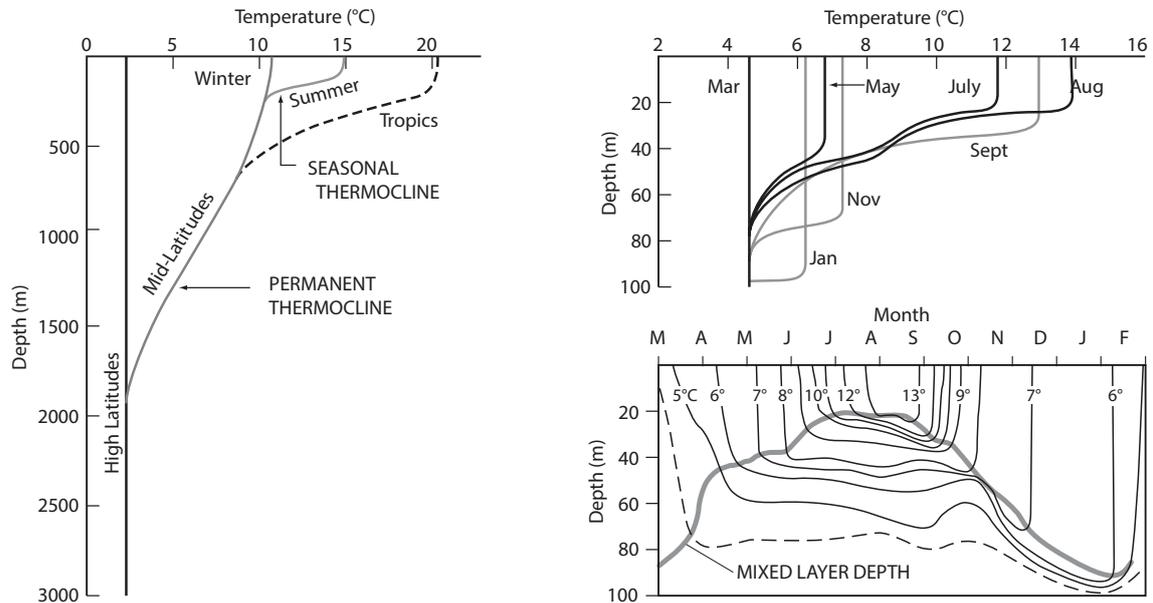


Figure 8.7: Schematic illustration of the vertical variation of temperature through the seasonal cycle. Figure from *Sarmiento and Gruber* [2006].

Although this elegant calculation is very appealing, we must not conclude that simple eddy diffusion and large-scale upwelling are the sole mechanisms that determine ocean mixing. At this point, these are merely reasonable parameterisations of several distinct physical processes. The formation and transport of water masses is the result of complex mechanisms which involve processes such as surface wind stress, atmosphere-ocean heat and freshwater fluxes, interior wave motion, instabilities, tidal action, flow over topography and more. The considerations above also ignored processes in the upper 1000 meters of the water column. There, the seasonal cycle is important and imprints a large, time-dependent signal on temperature, salinity and all tracers. Figure 8.7 shows the change in the temperature profile in the course of one year in the northern hemisphere. The uppermost layers which are exposed to the action of surface winds show temperature profiles with very little vertical variation. This is the so-called mixed layer. Mixed-layer depth undergoes a pronounced seasonal cycle (Fig. 8.7 right). Additionally, the structure of the upper km strongly depends on the geographical location. In high latitudes, stratification is weak due to much stronger winds and strong cooling which extracts sufficient buoyancy to promote vigorous vertical mixing.

Although many local processes determine the temperature and salinity of ocean water, there are conspicuous large-scale features which emerge from the observations. Figure 8.8 shows the vertical distribution of potential temperature (i.e., the temperature a water parcel would achieve under adiabatic expansion to the surface) along a track from the North Atlantic via the Southern Ocean to the North Pacific. The cold waters with little vertical temperature variations around Antarctica stand out. Also evident are the generally warmer waters in the North Atlantic Ocean compared to those in the Pacific. The general structure is dominated by the rather uniform decrease of surface temperature from the equator to the high latitudes.

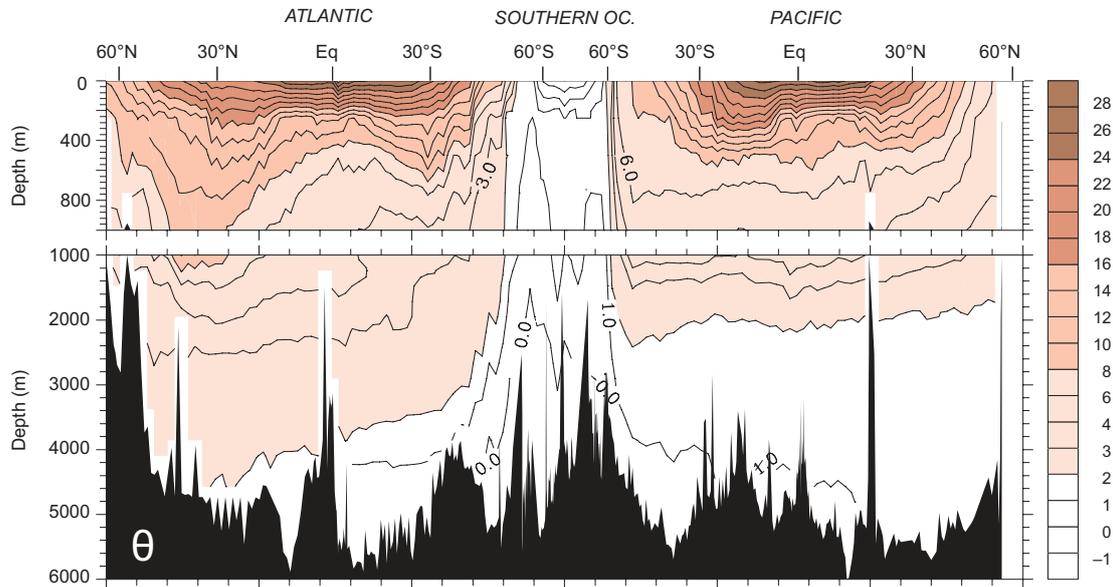


Figure 8.8: Vertical section of potential temperature along a track from the North Atlantic Ocean (left) via the Southern Ocean (middle) to the North Pacific (right) based on data from WOCE. Figure from *Sarmiento and Gruber* [2006].

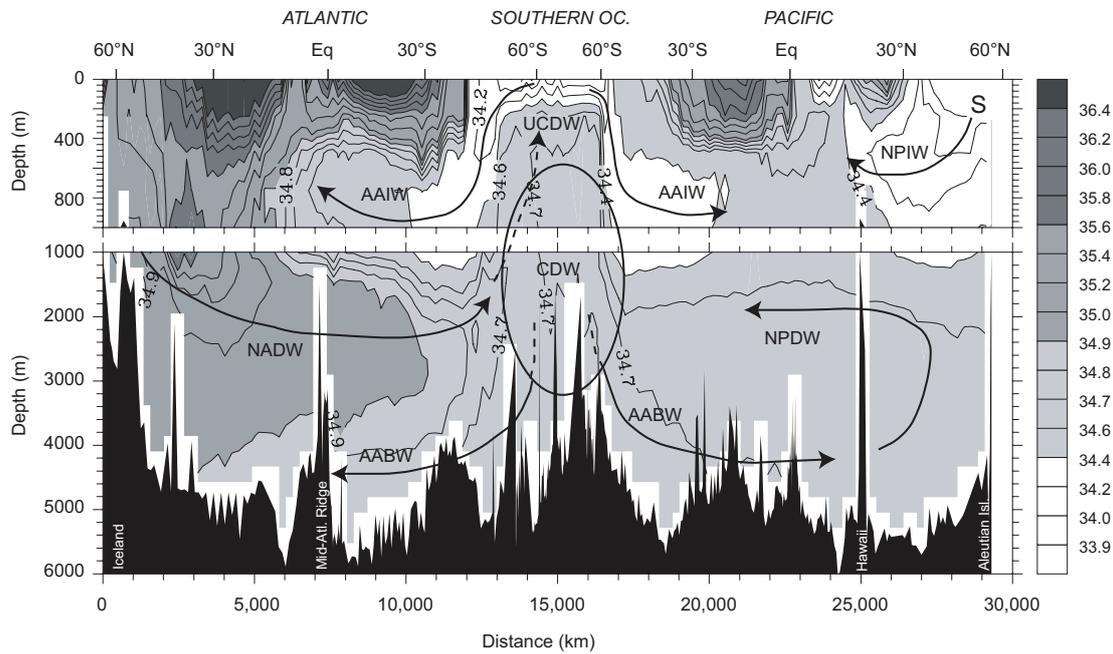


Figure 8.9: Vertical section of salinity along a track from the North Atlantic Ocean (left) via the Southern Ocean (middle) to the North Pacific (right) based on data from WOCE. Superimposed are the locations of the major water masses and schematic flow paths. The major water masses are NADW (North Atlantic Deep Water), UCDW (Upper Circumpolar Deep Water), AAIW (Antarctic Intermediate Water), NPDW (North Pacific Deep Water), AABW (Antarctic Bottom Water), NPIW (North Pacific Intermediate Water), and CDW (Circumpolar Deep Water). Figure from *Sarmiento and Gruber* [2006].

The distribution of salinity along the same track exhibits much more structure and permits the classification of waters into different water masses (Fig. 8.9). Water masses are distinguished essentially by the state variables T , S , and p (i.e., depth), but there is no rigorous definition.

Particularly evident is the "sandwich" structure of the water masses in the Atlantic Ocean (Fig. 8.9). The deep ocean is filled with Antarctic Bottom Water (AABW), the densest and coldest water mass which originates from a few deep water formation areas around the Antarctic continent. Further water mass abbreviations are given in the caption of Fig. 8.9. Another deep water formation area is in the Greenland-Norwegian Sea and the Labrador Sea where after mixing in the interior NADW is formed. The salinity structure in Fig. 8.9 suggests a competition between southward spreading NADW on the one hand, and northward flowing AAIW and AABW on the other hand.

8.4 Sea level rise

Sea level changes are caused by four processes. The density of sea water depends on temperature (Eq. 8.1). Warming or cooling, therefore, will cause an expansion or contraction of the water column which results in a change in sea level. This process is quantified by the thermal expansion coefficient of sea water (Table 8.1). This portion of sea level change is called *thermosteric*.

An additional important influence on sea level is due to changes in land ice mass. This is called *eustatic*. In addition, changes in the load on the Earth's crust (e.g., when large ice sheets form or disappear) also cause local sea level to change; these changes are referred to as *isostatic*. For sea level rise associated with a change from an ice age to an interglacial, the eustatic and isostatic contributions are dominant. A fourth process relevant particularly for local sea level change today is due to anthropogenic water storage changes, e.g. through groundwater pumping.

In order to estimate the thermosteric contribution to sea level, we consider a finite mass of water enclosed in a volume $A \cdot h$. Conservation of mass requires

$$\rho \cdot A \cdot h = (\rho + \Delta\rho) \cdot A \cdot (h + \Delta h), \quad (8.10)$$

assuming constant surface area A . From eq. 8.10 and neglecting higher-order terms, we obtain $\Delta h/h = -\Delta\rho/\rho$. If we now go to infinitesimal vertical extents of these volumes, i.e., $h \rightarrow dz$ and note that $\Delta h = -(\Delta\rho/\rho)h \rightarrow -(\Delta\rho/\rho)dz$, and integrate from an ocean depth $z = -H$

p [dbar]	$T = 2^\circ\text{C}$	$T = 10^\circ\text{C}$	$T = 25^\circ\text{C}$
0	$0.791 \cdot 10^{-4} \text{ K}^{-1}$	$1.672 \cdot 10^{-4} \text{ K}^{-1}$	$2.976 \cdot 10^{-4} \text{ K}^{-1}$
1000	$1.038 \cdot 10^{-4} \text{ K}^{-1}$	$1.859 \cdot 10^{-4} \text{ K}^{-1}$	
6000	$2.180 \cdot 10^{-4} \text{ K}^{-1}$		

Table 8.1: Thermal expansion coefficient α of sea water at salinity 35 for selected pressures and temperatures.

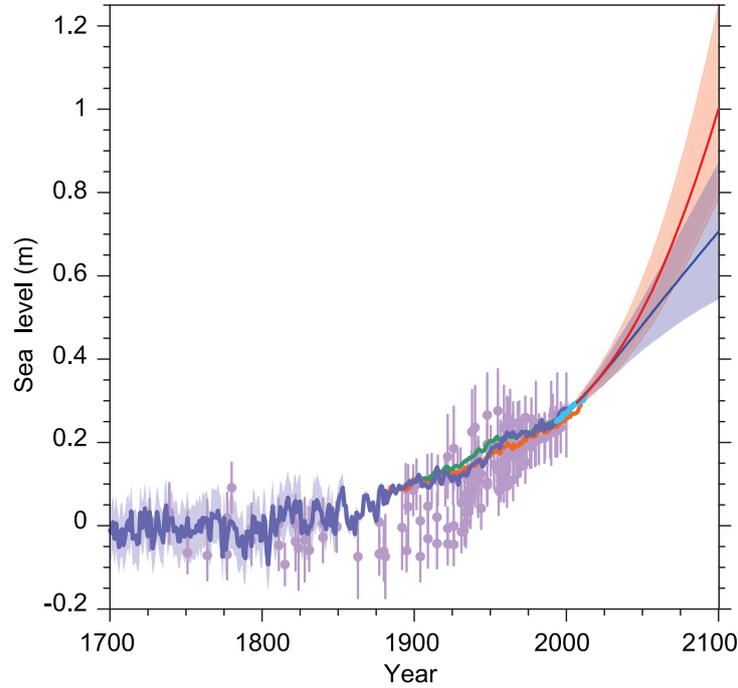


Figure 8.10: Compilation of sea level data from paleoclimate records, tide gauge data, and altimeter data, combined with model projections for two illustrative emissions scenarios (red: business as usual, blue: global mean temperature increase likely less than 2°C). Figure from *IPCC* [2013].

to the surface, we obtain

$$h_{\text{steric}} = - \int_{-H}^0 \frac{\Delta\rho}{\rho} dz = - \int_{-H}^0 \frac{1}{\rho} \frac{\partial\rho}{\partial T} \Delta T dz = \int_{-H}^0 \alpha \Delta T dz \quad . \quad (8.11)$$

The thermosteric sea level rise h_{steric} therefore depends on three quantities: (i) the change in temperature (we have neglected changes in salinity here), (ii) the depth H to which the change in temperature has penetrated, and the expansion coefficient α of sea water.

Note that α is a strong function of S , T and p as evident in Table 8.1. Therefore, it is very critical over which height of the water column a certain amount of heat is distributed. Vertical mixing, which controls the efficiency of heat transport into the interior of the ocean, is thus crucial to understand and simulate credible estimates of thermosteric sea level rise.

Sea level rise is one of the most dramatic and long-lasting impacts of anthropogenic climate change. It is caused by four major contributions: 1. the thermal expansion of warming water; 2. the melting of land glaciers; 3. the melting of the Greenland and Antarctic ice sheets; and 4. change in the storage of land water. To the mean rate of sea level rise in the period 2006 to 2015 is about 3.6 mm/yr and results from contributions from thermal expansion 1.4 mm/yr, glaciers 0.61 mm/yr, Greenland 0.77 mm/yr, Antarctica 0.43 mm/yr, and land storage -0.21 mm/yr. The remaining difference is within the estimated error bounds [*IPCC*, 2013; *IPCC*, 2019].

The historical context of sea level rise is shown in Fig. 8.10. Since 1900, global mean sea level

has risen by 19 cm. Due to the emissions of greenhouse gases since the Industrial Revolution we are committed to another 40 cm rise, even in a scenario in which the warming will likely not exceed 2°C by the end of the 21st century. Due to the contribution of the ice sheets and the slow warming of the world ocean, and the consequent thermal expansion, sea level will continue to rise for centuries. This is evident from the slope of the model projections at year 2100 in Fig. 8.10. The long-term (i.e., several centuries to millennia) amount of sea level rise caused by the emissions of greenhouse gases crucially depends on potential instabilities of large ice masses in West Antarctica and the amount of melting in Greenland [*Clark et al.*, 2016].

Chapter 9

Large-scale ocean circulation

9.1 Simplified coordinate system on the Earth and Coriolis acceleration

The most important influence of the large-scale circulations in the atmosphere and ocean is the rotation of the Earth. This implies that we must include the effect of inertial accelerations such as the centrifugal and Coriolis accelerations.

In this course, we will provide only a very basic and simplified introduction into the general circulation of the ocean. A more in-depth explanation of the different circulation types, including their mathematical description, is given in the lecture notes *Introduction to Climate Modelling*, available at <https://climatehomes.unibe.ch/~stocker/stocker22icm.pdf>.

Consider now the Earth as a rotating sphere which completes one rotation about its axis in 24 hours. Therefore, we can define an angular velocity vector pointing along the Earth's rotation axis $\boldsymbol{\Omega}$ with the magnitude $\Omega = 2\pi/(86,400 \text{ s})$. In such a system the two most important inertial accelerations to be considered are the centrifugal and the Coriolis accelerations. The former causes the deformation of the Earth to a rotation ellipsoid, and its effect on the circulation on this modified surface can be accounted for by slight spatial variations of the gravity acceleration g . This will not be further considered here. We also ignore, from now on, the deformation of the Earth and assume a sphere with radius $R = 6371 \text{ km}$.

The Coriolis acceleration depends on the angular velocity of the system and the relative velocity within the system, and its direction is perpendicular to both:

$$\mathbf{a}_c = -2(\boldsymbol{\Omega} \times \mathbf{v}) \quad (9.1)$$

where \mathbf{v} is the velocity in the rotating system, e.g., of ocean water or air in the atmosphere. The physical origin of the Coriolis acceleration is the conservation of angular momentum in a rotating system.

We now place a Cartesian coordinate system (x,y,z) on the Earth's surface and fix it at a geographical latitude φ (Fig. 9.1). This rectangular system is a rather poor but still useful approximation to the spherical surface on which the flow of the atmosphere and ocean will be studied. However, the most fundamental insights into the dynamics of these circulations can be obtained considering the simplified situation in the (x,y,z) -system.

In this system we can decompose $\boldsymbol{\Omega}$ into a component along the y -axis and one along the z -axis; the component along the x -axis is 0. In this coordinate system the velocity components are given by $\mathbf{v} = [u, v, 0]$, whereby it is assumed – again for simplicity – that the motion of the fluid is only horizontal, i.e. in the (x,y) -plane.

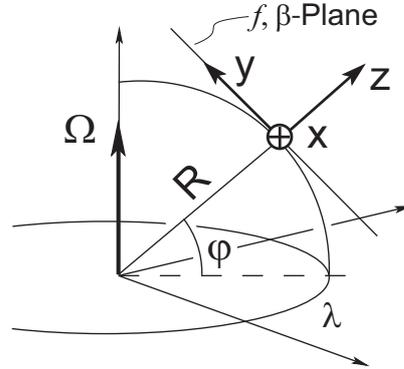


Figure 9.1: (x, y, z) -coordinate system attached to the rotating Earth at a fixed geographical latitude. The x -axis is pointing eastward, the y -axis northward and the z -axis upward.

Equation 9.1 can then be written as

$$\mathbf{a}_c = -2 \begin{bmatrix} 0 \\ \Omega \cos \varphi \\ \Omega \sin \varphi \end{bmatrix} \times \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} = \begin{bmatrix} 2v\Omega \sin \varphi \\ -2u\Omega \sin \varphi \\ 2u\Omega \cos \varphi \end{bmatrix} \quad (9.2)$$

The x and y directions of the Coriolis acceleration contain the common factor $f = 2\Omega \sin \varphi$, which we call *Coriolis parameter*.

Examining the relative directions of \mathbf{a}_c and $\mathbf{v} = [u, v, 0]$ we find that in the northern hemisphere the Coriolis acceleration is oriented 90° towards the *right* of the direction of the flow, while in the southern hemisphere, where $f < 0$, the Coriolis acceleration is 90° towards the *left* of the flow.

9.2 Hydrostatic balance

The Earth's gravity field causes the pressure in a fluid to increase with depth. This is shown in Fig. 9.2 (left). The resulting force in the z -direction on the fluid enclosed in the volume $A\delta z$ is given by

$$F_z = -Ap(z + \delta z) - mg + Ap(z). \quad (9.3)$$

The fluid is at rest if $F_z = 0$. Substituting $m = \rho A\delta z$ in Eq. 9.3, dividing by $\rho A\delta z$, and taking the limit $\delta z \rightarrow 0$ we obtain

$$a_z = \frac{1}{m} F_z = \frac{1}{\rho} \lim_{\delta z \rightarrow 0} \frac{-p(z + \delta z) + p(z)}{\delta z} - g = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g. \quad (9.4)$$

If the fluid is not accelerated vertically we have

$$a_z = 0 = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g. \quad (9.5)$$

This is the *hydrostatic balance*.

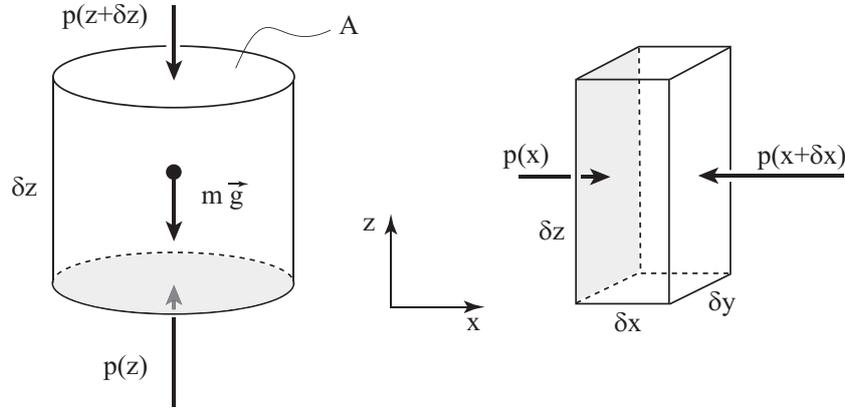


Figure 9.2: Left: Illustration of the hydrostatic pressure in a fluid of density ρ . Right: Pressure differences on a small volume of air or water cause a net force which acts on the mass enclosed in the volume.

9.3 Pressure gradient geostrophic balance

In order to calculate the motion of the fluid we now need to consider the true forces that act on a volume of fluid (either water or air). The most important force is caused by horizontal pressure differences in the fluid as illustrated in Fig. 9.2 (right).

The resulting component of the force along the x -axis, F_x , is given by

$$F_x = ma_x = p(x)\delta y\delta z - p(x + \delta x)\delta y\delta z \quad (9.6)$$

where m is the mass of the cube of water. Substituting $m = \rho\delta x\delta y\delta z$ in Eq. 9.6, dividing by m and then making the volume infinitesimally small, we obtain

$$a_x = \frac{1}{\rho} \lim_{\delta x \rightarrow 0} \frac{p(x) - p(x + \delta x)}{\delta x} = -\frac{1}{\rho} \frac{\partial p}{\partial x} . \quad (9.7)$$

By analogy, the acceleration due to pressure differences along the y -axis is

$$a_y = -\frac{1}{\rho} \frac{\partial p}{\partial y} . \quad (9.8)$$

We can now formulate the fundamental balance of forces (or acceleration) which holds between the pressure gradient and the Coriolis acceleration according to

$$0 = \mathbf{a}_c + \mathbf{a} = \begin{bmatrix} fv - \frac{1}{\rho} \frac{\partial p}{\partial x} \\ -fu - \frac{1}{\rho} \frac{\partial p}{\partial y} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} . \quad (9.9)$$

Note that here we have neglected all other forces such as e.g., friction. We rewrite Eq. 9.9 as

$$\begin{aligned} -fv &= -\frac{1}{\rho} \frac{\partial p}{\partial x} , \\ fu &= -\frac{1}{\rho} \frac{\partial p}{\partial y} . \end{aligned} \quad (9.10)$$

This is the *geostrophic balance*: pressure gradients in the x -direction cause velocities in the y -direction; pressure gradients in the y -direction cause velocities in the x -direction. This

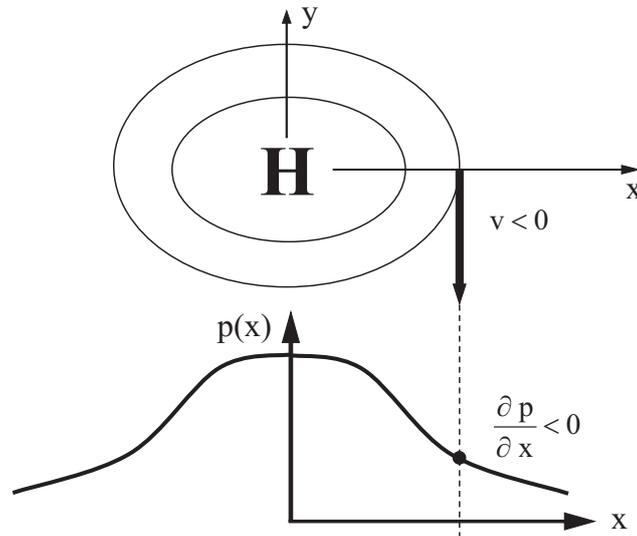


Figure 9.3: Geostrophic balance in the northern hemisphere where the Coriolis parameter f is positive. Top: view of the isobars with the high-pressure system. Bottom: pressure along the x -axis. The geostrophic circulation is tangential to the isobars and is clockwise around a high pressure system in the northern hemisphere.

is the most fundamental large-scale circulation type in the atmosphere and ocean on the rotating Earth and is readily observed in the atmosphere as the circulation around high and low pressure systems (Fig. 9.3).

9.4 Ekman circulation at the surface

At the ocean surface energy, water and momentum are exchanged between the atmosphere and ocean. The structure of the circulation close to the ocean surface (and also near the bottom of the ocean) is more complicated, and effects in addition to pressure gradients and Coriolis must be taken into account.

The circulation close to boundaries (e.g., at the air-ocean interface or at the bottom) is called *Ekman circulation*. It results from a balance of the wind stress, Coriolis acceleration, and friction within the fluid. Friction causes vertical transfer of momentum, i.e., the wind imparts momentum on the fluid which then starts to move, or momentum is removed from the fluid at the bottom of the ocean where the fluid is slowed down. Due to the Coriolis force, the velocity at the surface of the northern hemisphere is rotated 45° towards the right. The derivation of this Ekman velocity profile is beyond the scope of this introduction. Momentum is dissipated downwards which results in a rapid exponential decrease of the velocity with depth. The typical Ekman depth in mid latitudes is about 150 m and increases towards the equator (Ekman depth $\propto f^{-1/2}$). The continuous action of the Coriolis force causes further rotation of the velocity vector down the water column, as illustrated in Fig. 9.4. This is called the *Ekman spiral*. In the southern hemisphere, the Ekman spiral rotates anti-clockwise.

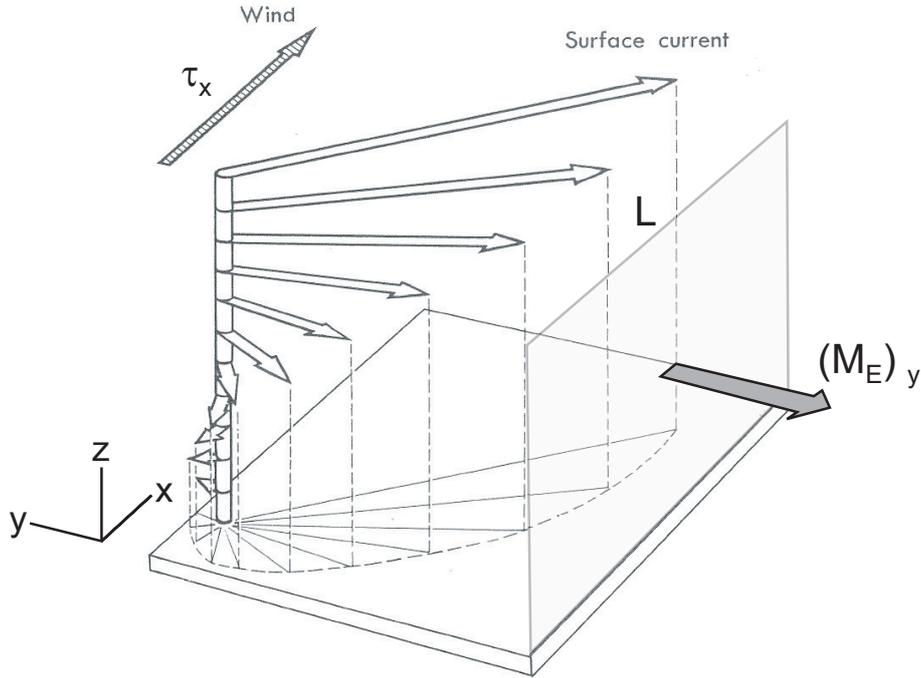


Figure 9.4: Direction of the water flow close to the ocean surface as forced by the wind stress. In the northern hemisphere the flow direction rotates clockwise. It decreases in strength as one moves downwards from the surface. The complicated Ekman flow results in an average Ekman transport across the shaded rectangle that is perpendicular to the direction of the wind. Figure modified from *von Arx* [1979].

Without derivation, which is beyond this introduction, we note that the integrated volume transport (in m^3/s) through an area of width L and depth over the entire Ekman layer (Fig. 9.4) is given by

$$\begin{aligned} (M_E)_x &= \frac{L}{\rho f} \cdot \tau_y , \\ (M_E)_y &= -\frac{L}{\rho f} \cdot \tau_x , \end{aligned} \quad (9.11)$$

where $\tau = [\tau_x, \tau_y]$ is the horizontal wind stress vector in units of N/m^2 , and L is the width of the area across which the fluid flows. The quantity M_E in Eq. 9.11 is called *Ekman transport*, and is perpendicular to the direction of the wind. In the northern hemisphere ($f > 0$), the Ekman transport is directed 90° towards the right of the wind direction, in the southern hemisphere ($f < 0$) it is 90° towards the left. Again, this is an effect due to the combination of friction and the Coriolis force.

Ekman transport is the most important process to set up the large-scale circulation in the world ocean. It is also the process that determines the circulation close to topographic boundaries such as coasts, sea mountains and the ocean floor.

A prominent feature in the climate system are the trade winds in the tropics. Their direction, both north and south of the equator, is from east to west. Ekman transport is particularly strong as one approaches the equator because $M_E \propto f^{-1}$ as in Eq. 9.11. North of the equator

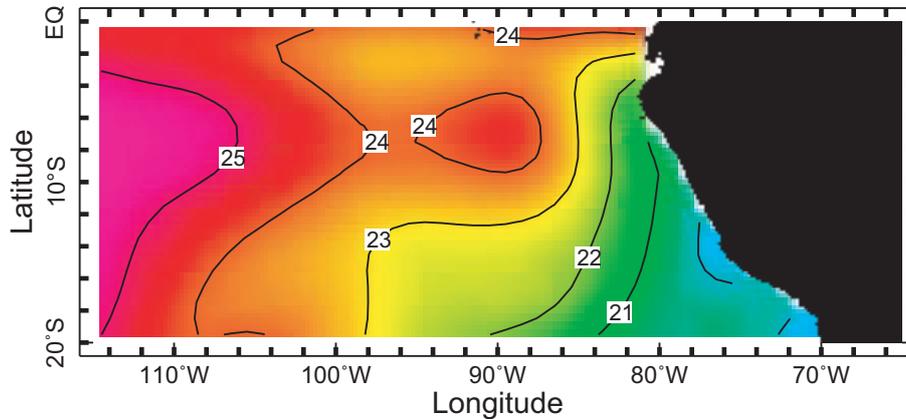


Figure 9.5: Annual mean sea surface temperature off the coast of Peru. The lowest temperatures are found close to the coast due to cold waters upwelling from several 100 m depth caused by Ekman upwelling. Figure prepared online using the IRI/LDEO Climate Data Library.

tor, Ekman transport is northward, south of the equator it is directed to the south. This results in a divergence of the flow, which implies that water must be replenished from below. Strong equatorial upwelling with effects on temperature and salinity along the equator can be observed. For example, the isotherms and isopycnals show a dome-like structure across the equator (Fig. 8.8).

Likewise, coastal winds produce strong Ekman transports with important regional consequences. Along the coast of Peru, for example, the mean wind has a significant northward component. Because the location is in the southern hemisphere, Ekman transport is thus directed towards the west, hence away from the coast, and water must upwell there. This is the origin of colder and nutrient-rich waters off the Peruvian coast that supports very productive fisheries in this region (Fig. 9.5). The fishery industry is very vulnerable to changes in the regional wind system. For example, during El Niño years, the northward winds weaken therefore the supply of nutrients due to Ekman upwelling ceases. In consequence, fish harvest collapses. There are many other coastal areas where Ekman transport dominates the local oceanographic conditions.

9.5 Sverdrup balance

We now turn to the large-scale circulation which is basin-wide and extends over many 1000 km. The conservation of angular momentum permits a basic physical understanding of the flow characteristics in the ocean.

Consider a mass m rotating around a fixed point O (Fig. 9.6). Angular momentum is given by

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \mathbf{r} \times (m\mathbf{v}) = m(\mathbf{r} \times (\boldsymbol{\Omega} \times \mathbf{r})) = m(\boldsymbol{\Omega} \cdot (\mathbf{r} \cdot \mathbf{r}) - \mathbf{r} \cdot (\mathbf{r} \cdot \boldsymbol{\Omega})) = mr^2 \cdot \boldsymbol{\Omega}, \quad (9.12)$$

and we have assumed that $\mathbf{r} \perp \boldsymbol{\Omega}$ as in Fig. 9.6. Angular momentum thus depends on the angular velocity and the distance from the rotation axis. Even if a mass is not moving relative

to the Earth's surface it has angular momentum, called *planetary angular momentum*, due to the rotation of the Earth. The magnitude of planetary angular momentum depends on the geographical latitude (Fig. 9.6). Increasing the latitude (moving poleward) reduces the distance from the rotation axis, and therefore also the planetary angular momentum decreases. The angular momentum of a mass m at latitude φ is given by

$$L(\varphi) = mR^2\Omega \cos^2 \varphi. \quad (9.13)$$

The angular momentum changes if torques act on the system. If the sum of all torques is zero, then angular momentum is conserved. This is a very powerful concept that helps understand the basics of the circulation of the ocean or the atmosphere in the rotating system of the Earth.

We now seek to qualitatively explain the circulation that results from the action of a typical wind pattern in mid latitudes of the northern hemisphere. At latitudes higher than about 30°N winds are blowing from the west to the east (Westerlies), at latitudes south of 30°N , the winds are blowing in the opposite direction (Easterlies, trade winds). This particular latitude dependence of the wind stress τ_x causes Ekman convergence and hence a downwelling of water. The downwelling waters move a bit closer to the Earth's rotation axis and thus lose angular momentum. To compensate for this loss, water must move towards the equator until the original distance of the water from the rotation axis is restored and hence the angular momentum is conserved (Fig. 9.7, left). This large-scale drift is called *Sverdrup transport*.

For the wind stress conditions in the mid-latitudes the Sverdrup transport is directed equatorward. It is a dynamical process which occurs in the absence of friction, other than the transfer of momentum from the atmosphere to the ocean, over the major part of the ocean basin (Fig. 9.8). *Sverdrup balance* is a consequence of the spherical shape of the Earth and therefore, in mathematical terms, of the dependence of the Coriolis parameter f on latitude. Quantitatively one can show that the meridional Sverdrup transport (in m^3/s) across a zonal

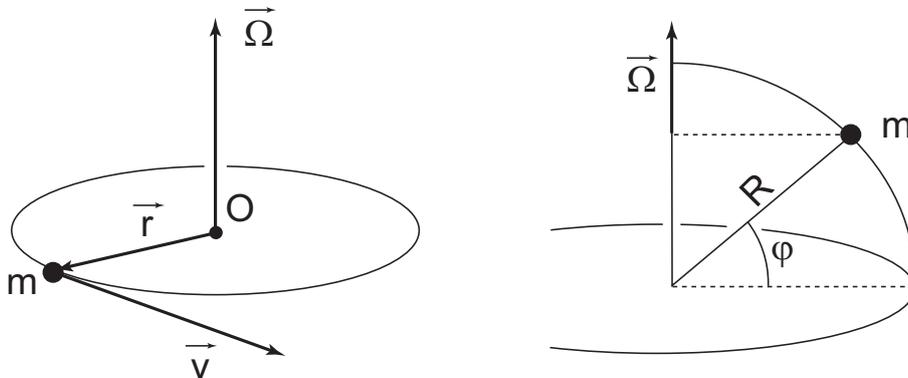


Figure 9.6: Left: rotation of a mass point around a fixed point O with angular velocity $\vec{\Omega}$, radius r and velocity \vec{v} . Right: mass point on the rotating Earth fixed at geographical latitude φ and at radius R .

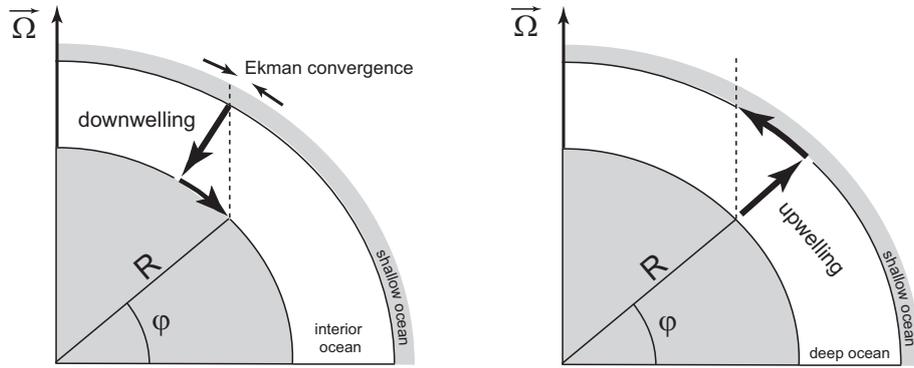


Figure 9.7: Schematic view of the large-scale motion (Sverdrup transport, left panel) caused by Ekman convergence and consequent downwelling in the subtropical gyres, and by upwelling of deep water (right panel). Total angular momentum is determined by the distance from the rotation axis. In order to conserve total angular momentum in the absence of any forces, water must move equatorwards while downwelling, and polewards while upwelling.

extent of width L and over the entire depth of the basin is given by

$$M_y = -\frac{L}{\rho\beta} \frac{\partial \tau_x}{\partial y}, \quad (9.14)$$

where β is the change of the Coriolis parameter f with latitude:

$$\beta = \frac{\partial f}{\partial y} = \frac{1}{R} \frac{\partial f}{\partial \varphi} = \frac{2\Omega}{R} \cos \varphi. \quad (9.15)$$

The Sverdrup transport is a consequence of the latitude-dependence of the wind stress and the spherical shape of the Earth, embodied in β . From Fig. 9.8 it becomes clear that the Sverdrup balance alone cannot provide an understanding of the complete circulation in a closed basin. If most of the basin exhibits a southward circulation, how is the fluid brought back northward?

An additional physical process must be considered to explain the full circulation.

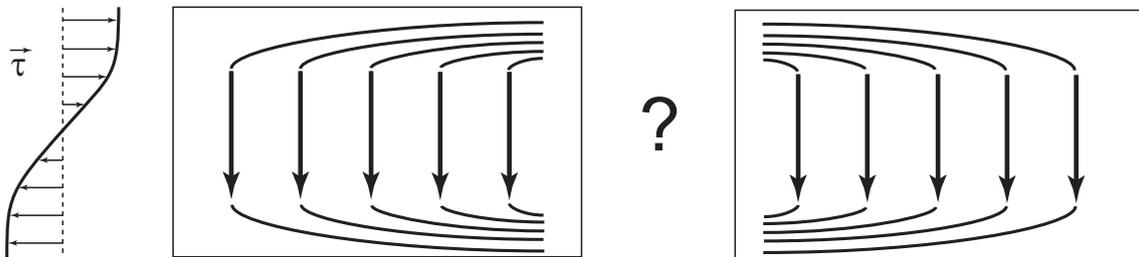


Figure 9.8: A typical wind stress profile in the northern hemisphere (westerlies in the mid-latitudes and easterlies in the subtropics) produces a basin-wide equatorward drift of the water. This is the Sverdrup transport which is due to the latitude dependence of the Coriolis parameter and wind stress. In a closed basin mass is conserved, and therefore the flow must return in a strong northward boundary current, either along the eastern or along the western boundary of the basin.

9.6 Western boundary currents

The Sverdrup balance is not able to provide us with an explanation of a consistent circulation in a closed basin. As illustrated in Fig. 9.8 there are in principle two possibilities to close the flow in the basin. Either there is a very strong and narrow northward current along the eastern boundary (left basin in Fig. 9.8) or along the western boundary (right basin in Fig. 9.8). There are two ways how we can physically argue which case will be realized in nature.

The first argument uses the Ekman flow that was explained in section 9.4. The eastward blowing wind in the midlatitudes (northern part of the basin) produces in the northern hemisphere a southward Ekman transport that is proportional to the strength of the wind stress. Likewise, a northward Ekman transport results from the westward blowing wind in the southern part of the basin. Hence, along the centre latitude of the basin there is an Ekman convergence which raises the ocean surface. The associated pressure gradients will balance the Coriolis force and this results in eastward flow in the northern part of the basin and westward flow in the southern part of the basin. Hence, the only possibility for a closure is a *western boundary current*.

The second explanation is more fundamental in that it directly explains the closing current using the balance of angular momentum. The wind stress imparts negative angular momentum on the flow. Also, the required northward movement brings the water mass closer to the rotation axis of the Earth and hence decreases planetary angular momentum further. These two processes lead to a strong reduction of angular momentum. For a steady state, therefore, angular momentum must be returned back to the fluid in the boundary current, or equivalently, the negative angular momentum must be transferred to the bottom of the ocean.

This is possible through friction which occurs when flow exhibits strong shear (shear = strong changes of velocity with location). Shear in the fluid represents local angular momentum of the flow as is readily seen if one examines the orientation of rotation of a small propeller that is held into the fluid. In the western boundary current, such propellers would rotate clockwise due to the decreasing velocity as one moves away from the western boundary. The flow therefore has negative local angular momentum.

Friction within this flow (lateral mixing) and at the ocean bottom both tend to weaken the shear. The magnitude of the strong negative local angular momentum would thus decrease and hence friction acts as a source of positive angular momentum to the fluid in the western boundary current.

In summary, the decrease of total angular momentum through wind stress and northward movement is compensated by an input of positive angular momentum by bottom friction and lateral mixing in the boundary current. Were the strong current to occur along the eastern boundary, friction would deliver negative angular momentum to the fluid, and a balanced flow could not be realized. Hence, the only possibility is a *western boundary current*.

The result of the full calculation for a closed basin was first presented by *Stommel* [1948] and is depicted in Fig. 9.9. It shows a clockwise circulation with slow speeds in most of the ocean basin and very large current velocities along the western boundary of the basin. This

”circular” flow is often referred to as a *gyre*. Note that outside the western boundary current, in the interior of the ocean basin, the velocity vector always has a southward component. This is fully consistent with the requirement of a Sverdrup balance in the interior.

The effect that is responsible for the formation of a western boundary current is the variation of the Coriolis parameter with latitude. This is equivalent to the notion that the planetary angular momentum depends on latitude. In Fig. 9.9 (right) the flow lines are shown for the case of a rotating but flat Earth, i.e. $f = \text{constant}$. In this case the flow remains symmetric, and no western boundary current is formed.

9.7 Deep circulation in the ocean

The consideration of angular momentum also helps us understand the deep circulation in an ocean basin. The structure of temperature and other tracers in the water column has suggested that there must exist wide-spread upwelling in the ocean, although this motion is extremely slow and cannot be measured directly.

Let us consider an idealised situation of deep water in an ocean basin of the northern hemisphere. The deep water is assumed to upwell in a uniform way, as discussed in section 8.3 (Eq. 8.9), and friction is neglected. As shown in Fig. 9.7 (right), the upwelling water moves away from the rotation axis of the Earth. In the absence of forces on the water (i.e., no friction and no wind stress because we are in the deep ocean), the angular momentum of the water must be conserved. By this upwelling movement the water gains planetary angular momentum. The only way to satisfy the balance of angular momentum is to move the water northward while upwelling so that the original distance to the rotation axis is restored. Upwelling of deep water in ocean basins thus leads to poleward motion. This is the qualitative explanation of the famous Stommel-Arons model (Fig. 9.10) which was sketched in 1958 by

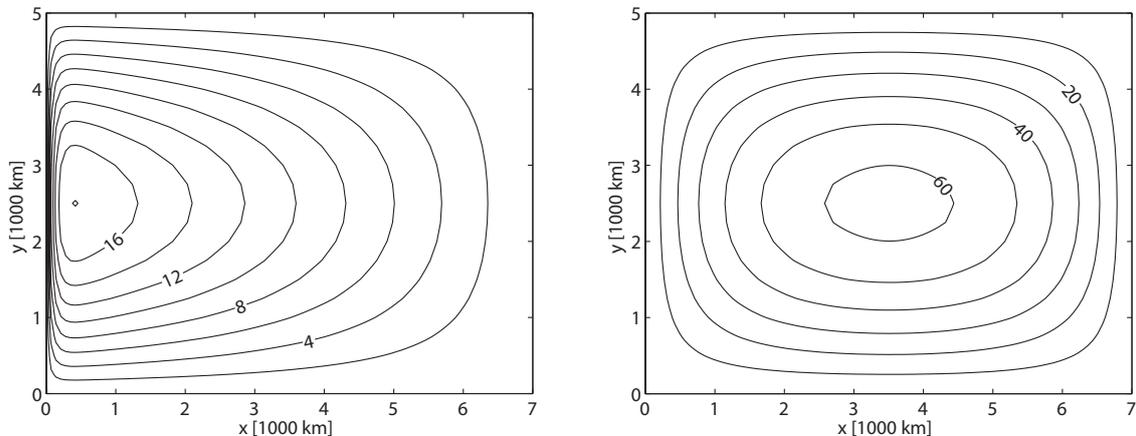


Figure 9.9: Volume streamfunction (in Sverdrup, $1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$) of the full calculation (Stommel model) of the wind-driven circulation in a rectangular, northern hemispheric basin for the case in which the Coriolis parameter f varies with latitude (left) and a case where f is constant (right). The current is along the streamlines and clockwise in both cases.

Henry Stommel and thoroughly formulated in 1960 [*Stommel and Arons, 1960*].

There is another interesting implication of this. The basin-wide, very slow northward and upward movement of deep water must be fed from somewhere. First, in the high latitudes of the Atlantic Ocean (Labrador Sea, Greenland-Norwegian-Iceland Seas) there is localised deep water formation, i.e., sinking of surface water to large depths. These are sources of newly formed deep water which compensate the large-scale upwelling. The northward movement, on the other hand, is compensated by a *deep western boundary current* which was postulated in 1958 when Henry Stommel first presented this concept in his landmark paper (Fig. 9.10).

One of the most spectacular breakthroughs of physical oceanography was the actual discovery of this deep western boundary current in the North Atlantic by current measurements at a depth of 2-3 km in the early 1960ies. This current closely hugs the western slope of the Atlantic basin and is strong enough to cross the equator. This way, North Atlantic Deep Water gets distributed to the other ocean basins of the world ocean. Similar western boundary currents are predicted to occur in the Pacific and Indian Oceans. These are essential elements of the global deep ocean circulation system.

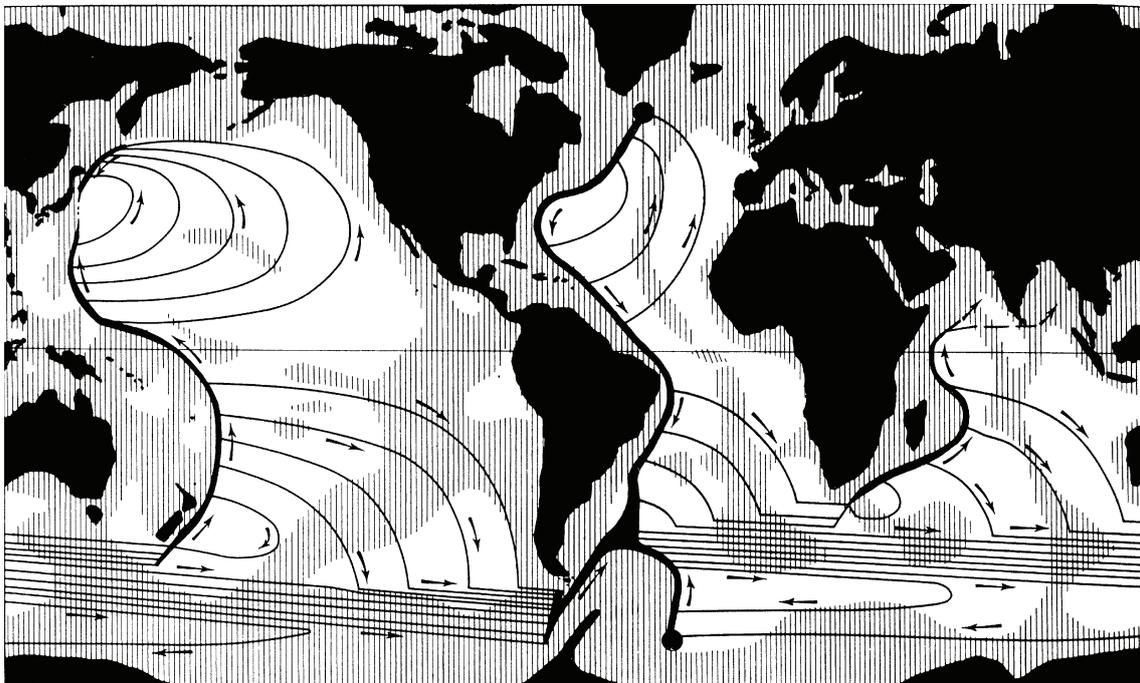


Figure 9.10: Illustration of the deep circulation as conceived by Henry Stommel in 1958 before the discovery of the intensive deep western circulation in the Atlantic Ocean. There is basin-scale poleward flow resulting from the conservation of angular momentum and postulated strong boundary currents indicated by bold lines. The western boundary currents feed the interior flow. Two isolated deep water formation areas are marked by black dots. Figure from *Stommel [1958]*.

9.8 Synthesis view of the basin-scale circulation

We can now attempt to synthesize the different circulation types and depict them schematically in Fig. 9.11. This figure shows a sectorial ocean basin in the northern hemisphere with an upper ocean of 1000 m depth over a deep ocean below. The zonal wind stress pattern drives the circulation at the surface. It causes meridional Ekman transport resulting in Ekman divergence and convergence which deform near-surface isopycnals (purple line labelled ρ) and eventually limits the gyre circulation to the upper ocean due to compensating pressure gradients.

The gyre circulation shows the intense western boundary currents and the interior where the Sverdrup balance holds. The wind stress pattern forces a counter-clockwise subpolar gyre and a clockwise subtropical gyre. The connection between the upper ocean and the deep ocean is through basin-wide upwelling (blue vertical arrows) and localised deep water formation in the high latitudes (red vertical arrow). The former drives the poleward drift in the interior deep ocean, while the latter feeds the deep western boundary current. If the deep water formation is stronger than the upwelling, the deep western boundary current will cross the equator and extend into the southern hemisphere. Note that the water flow can only cross the equator in

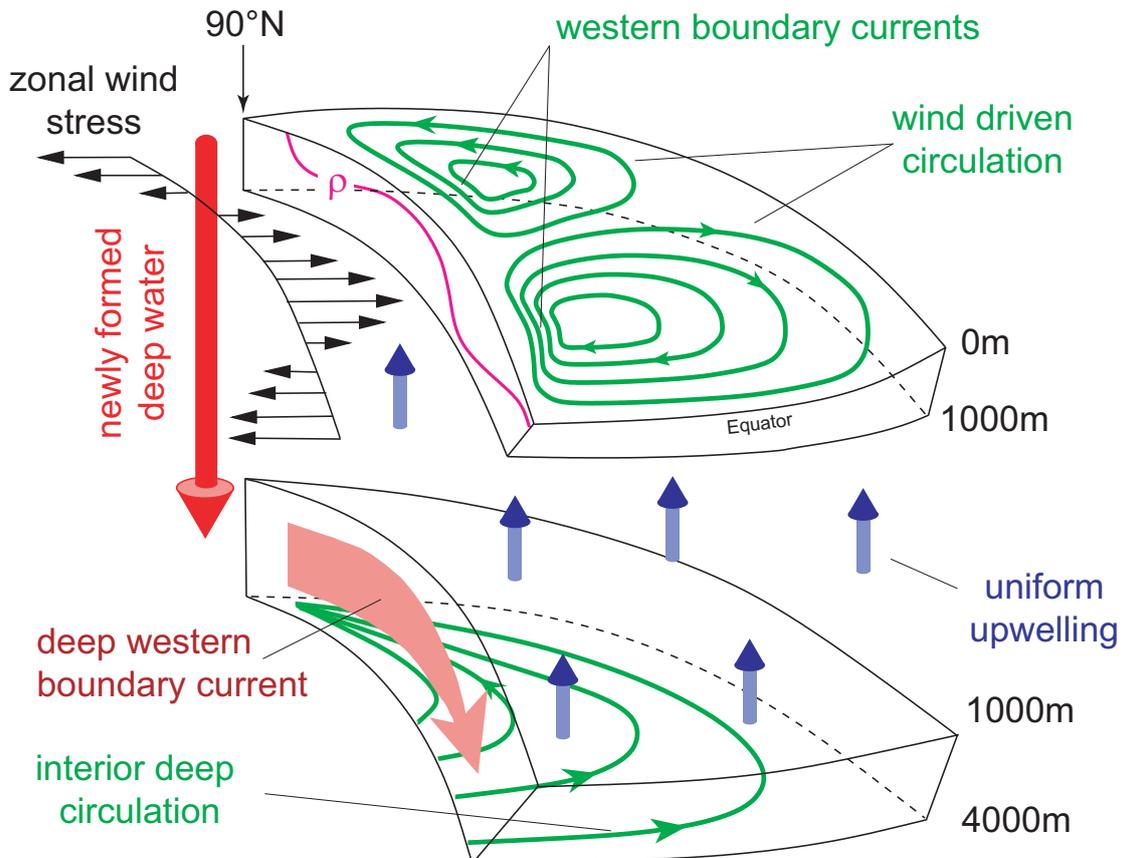


Figure 9.11: Schematic overview of the different types of steady-state circulations in a sectorial ocean basin extending from the equator to the North Pole.

currents close to basin boundaries, i.e., where friction or mixing play an important role, but not in the interior of the ocean.

This is the dynamical mechanism by which a global circulation is set up that is often referred to as the *global conveyor belt*, a notion coined by Wallace Broecker in 1987 and shown in Fig. 9.12. While very appealing and inspiring, note that the artist's impression of Fig. 9.12 ignores the various aspects of ocean dynamics which have been briefly introduced in this section. Nevertheless, some features of this iconic ocean circulation are in qualitative agreement with the dynamical understanding: (i) a large-scale meridional overturning extending over the entire Atlantic Ocean, (ii) upwelling in the Indian and Pacific Oceans, and (iii) exchange of water masses via the Southern Ocean in the Antarctic Circumpolar Current.

The Atlantic Ocean is the only location where warm surface water flows northward and cold deep water flows southward, across the equator. This exchange of water with different temperature operates like a giant heat pump in the Atlantic sector. It significantly influences the climate in the North Atlantic region and delivers large amounts of heat which makes the climate warmer there than at similar latitudes in the Pacific region.

The Atlantic Ocean is the only location where warm surface water flows northward and cold deep water flows southward, across the equator. This exchange of water with different temperature operates like a giant heat pump in the Atlantic sector. It significantly influences the climate in the North Atlantic region and delivers large amounts of heat which makes the climate warmer there than at similar latitudes in the Pacific region.

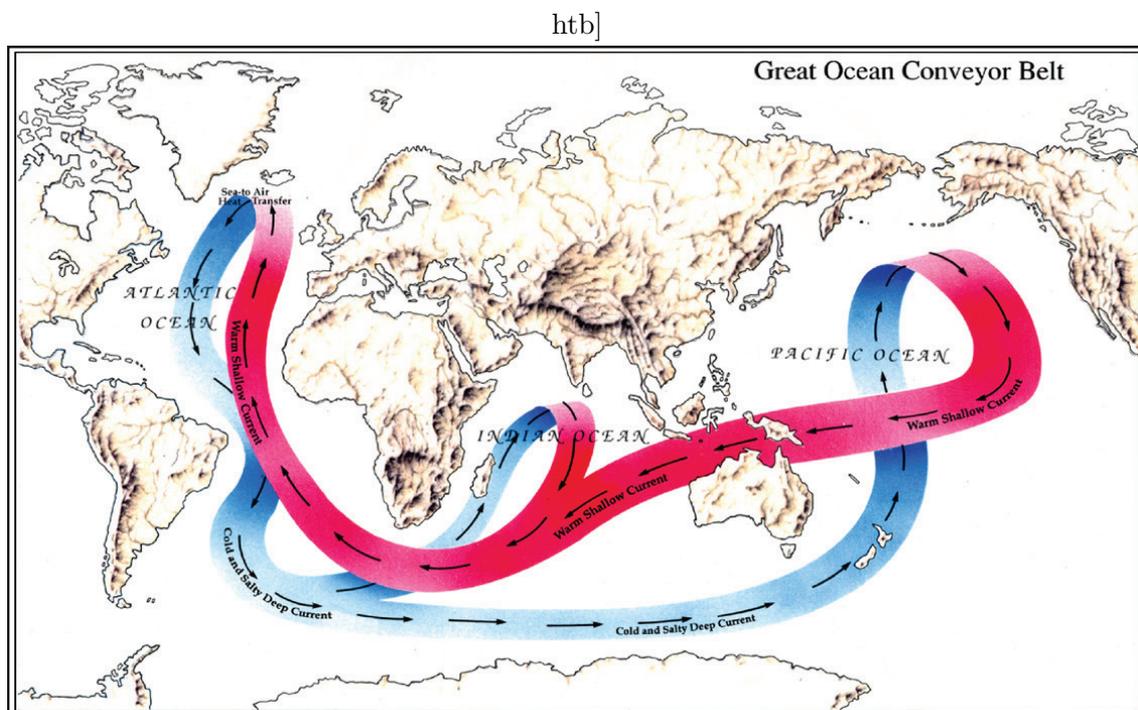


Figure 9.12: Reproduction of the original and by now iconic illustration of the "Great Ocean Conveyor Belt" as conceived by Wally Broecker. Figure from *Broecker* [1987].

Chapter 10

Multiple equilibria in the climate system

10.1 Abrupt climate transitions recorded by polar ice cores

The most detailed information about past climate states of the last 800,000 years can be retrieved from polar ice cores. One example is presented in Fig. 10.1. The Holocene, the present inter-glacial, has started some 10,000 years ago after the abrupt end of the last glacial period, 11,650 years ago. This end had been announced by certain signs already 20,000 years ago. For this period, an increase in the isotope concentrations in the Antarctic ice core could be detected. Stable isotopes of the water molecule are a measure for the local temperature (see Chapter 2). The temperature indicators prove that the climate changed in an abrupt way more than 20 times in Greenland over the course of the last glacial period. These abrupt warming events, numbered in Fig. 10.1, are referred to as Dansgaard-Oeschger events (D/O events) in remembrance of the research of the two pioneers Willy Dansgaard (1922-2011, University of Copenhagen) and Hans Oeschger (1927-1998, University of Bern).

These climate phases in Greenland all show an abrupt warming within about one decade and a subsequent continuous cooling over about 1000 to 3000 years. Interestingly, the isotope maxima and minima during the glacial periods are all at the same level. Already in 1984, Hans Oeschger surmised that the climate system operated in a way similar to a physical flip-flop, and that the ocean circulation in the Atlantic is likely to be responsible for these climate jumps. Flip-flop systems are characterized by several stable equilibria.

After *Bryan* [1986] demonstrated by means of a 3-dimensional ocean circulation model that several states of the deep circulation can be realized, Wally Broecker synthesized the results from different climate archives and argued that rapid oscillations of the "Atlantic heat pump" (the great conveyor belt circulation, Fig. 9.12) are responsible for the abrupt climate transitions found in the Greenland ice cores, in tree rings, in sea and lake sediments and in numerous other paleo-archives.

10.2 The Bipolar Climate Seesaw

Evidence from many climate archives supports the hypothesis that the ocean is primarily responsible for these abrupt transitions. A sudden shut-down of the North Atlantic deep water formation causes a reduction of the meridional heat flux and therefore an abrupt cooling in the North Atlantic region. This hypothesis makes some predictions that can be tested in various climate archives. An active meridional current draws heat from the South Atlantic. A shut-down of the heat pump will consequently cause a cooling of the North Atlantic and a

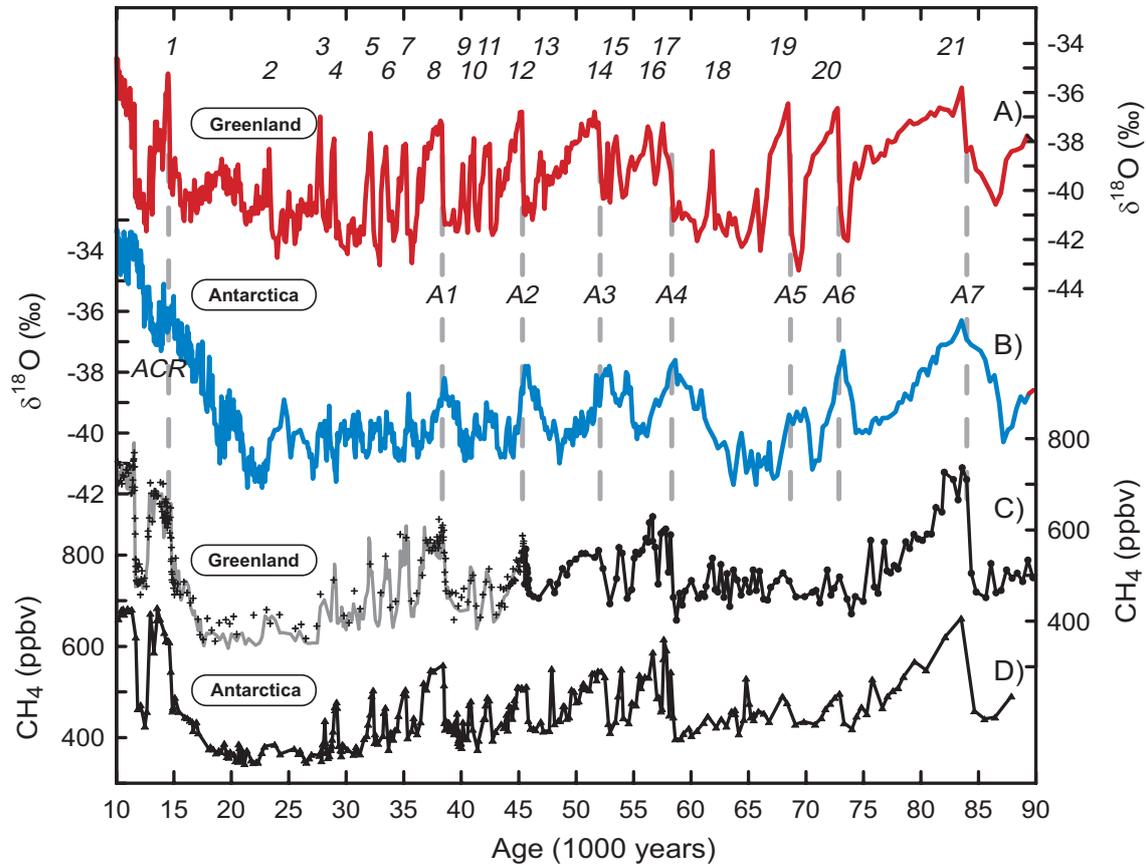


Figure 10.1: Climate history of the last 90,000 years recorded in ice cores of Greenland and of Antarctica. A) Oxygen-isotope concentration ^{18}O (in per mille deviation from a predefined standard) in the GRIP ice core from Greenland; B) ^{18}O in the Byrd core from Antarctica; C) Methane concentration in the GRIP core; D) methane concentration in the Byrd core. In the ice core, 21 Dansgaard/Oeschger events with a differing duration of the preceding cooling are recorded. The longest D/O events exhibit a corresponding warm event in the Antarctic core; labeled A1 to A7. All of the D/O events are marked by abrupt peaks in the methane, enabling a synchronization of the time scales of Greenland and Antarctic ice cores. Figure from *Blunier and Brook* [2001].

concomitant warming of the South Atlantic which should be noticeable in climate reconstructions from these regions. This has led to the formulation of the so-called "Bipolar Seesaw", a paradigm for the interaction of the northern and southern hemispheres during abrupt climate transitions [Broecker, 1998; Stocker, 1998]. The Bipolar Seesaw is illustrated in Fig. 10.2 and exemplifies that an abrupt warming in the north leads to an abrupt cooling of the South Atlantic and vice-versa. Coupling a large heat reservoir to this seesaw leads to the situation that an abrupt cooling in the South Atlantic induces a slow continuous cooling in the whole Southern Ocean. In this simplistic model, the different characteristics of temperature signals as they are reconstructed based on ice cores from Greenland and Antarctica (Fig. 10.1) can be explained [Stocker and Johnsen, 2003].

The simple bipolar thermal climate seesaw is formulated as an energy balance of the Southern

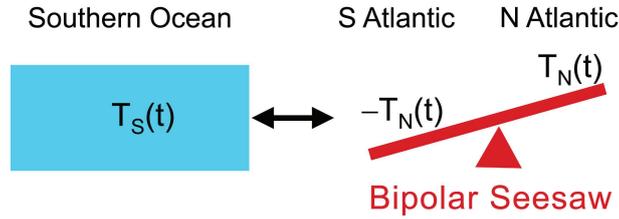


Figure 10.2: Bipolar Climate Seesaw coupled to a southern heat reservoir (Figure after *Stocker and Johnsen* [2003]).

Ocean:

$$\frac{dT_S(t)}{dt} = \frac{1}{\tau} (-T_N(t) - T_S(t)) \quad , \quad (10.1)$$

where T_S is the temperature anomaly of the Southern Ocean and T_N represents the temperature anomaly of Greenland. τ is a characteristic time scale for the heat equilibration in the Southern Ocean. If $T_N(t)$ is given, the temporal evolution of $T_S(t)$ can be determined by a Laplace transformation of Eq. 10.1:

$$T_S(t) = -\frac{1}{\tau} \int_0^t T_N(t-t') e^{-\frac{t-t'}{\tau}} dt' + T_S(0) e^{-\frac{t}{\tau}} \quad . \quad (10.2)$$

Hence, T_S is completely determined by the temporal evolution of T_N and represents the integrated northern temperature signal with a "fading memory". Let us consider this simple model in order to explain the different temporal evolution of the temperatures in Greenland and Antarctica. By tuning the only free parameter τ of this model we aim at producing the largest possible correlation between the modeled T_S based on Eq. 10.2 with the known temperature from the ice core T_N as input and the measured T_S derived from the Antarctic ice core. For $\tau \approx 1000$ years a maximum correlation of 0.77 is achieved. This allows us to predict the Antarctic temperature based on the temperature of Greenland in a surprisingly accurate way (Fig. 10.3).

10.3 Multiple equilibria in a simple ocean model

The deep circulation in the Atlantic is associated with a large heat transport that considerably affects climate in the North Atlantic region. This heat transport is responsible for a relatively mild climate there. Already at the beginning of the 20th century, geologists suggested that the change in the ocean circulation could have been responsible for climate oscillations. In 1961, Henry Stommel formulated an elegant conceptual model that is able to reproduce such changes, since it contains several equilibria [*Stommel, 1961*]. This model is presented in its simplified form according to *Marotzke* [2000]. The existence of multiple equilibria is owing to the fact that heat and water fluxes respond differently to anomalies in ocean surface temperature and salinity, respectively. Surface temperature anomalies cause changes in atmosphere-ocean heat fluxes that damp the anomalies efficiently. Conversely, anomalies in surface salinity have no effect on the atmospheric water cycle, i.e. evaporation and precipitation are not responding to salinity changes. In this conceptual model, the ocean

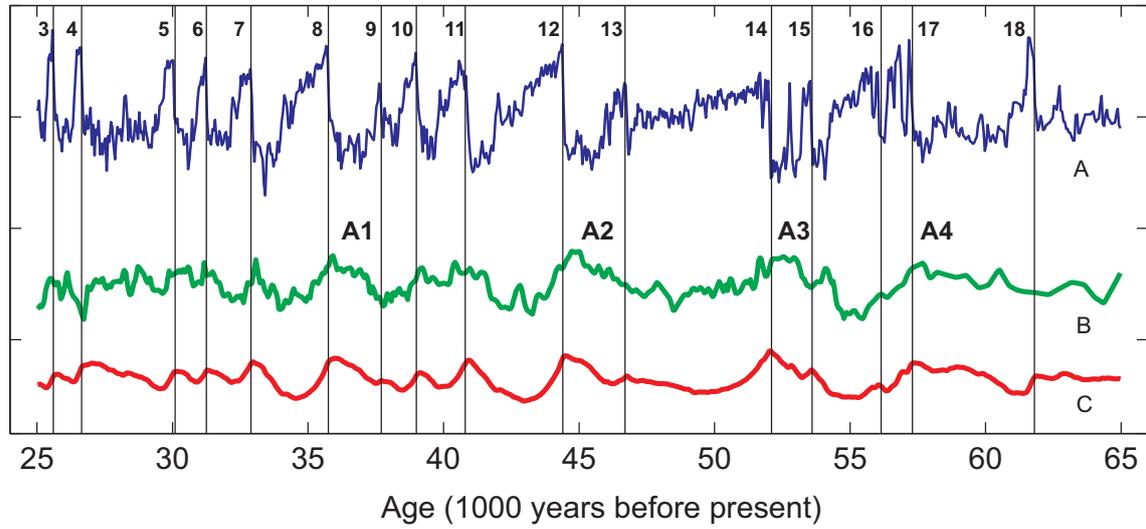


Figure 10.3: High-pass filtered time series of the temperatures in Greenland (A) and Antarctica (B) derived from ice cores. Curve C is the simulated temperature according to Eq. 10.2 with input (A). The abrupt Dansgaard/Oeschger events of the north hence become manifest in the series of warming events in Antarctica (A1, A2, ...) (Figure from *Stocker and Johnsen* [2003]).

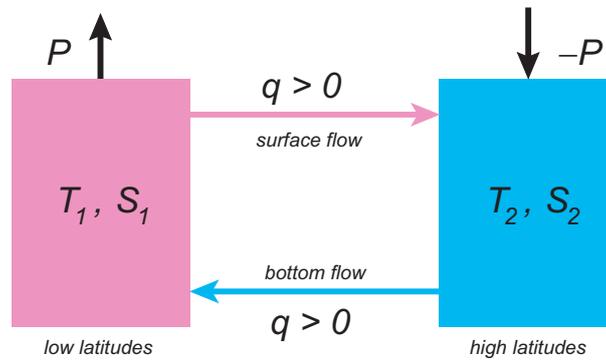


Figure 10.4: 2-Box model of the thermohaline circulation (after *Stommel* [1961] and *Marotzke* [2000]).

is strongly simplified, consisting of two boxes: one for latitudes where evaporation dominates and hence with a positive water flux P , and one for the high latitudes where precipitation dominates (Fig. 10.4). T_i and S_i represent the temperatures and salinities of the two boxes. A fixed temperature difference ΔT between the boxes is assumed which is maintained by heat fluxes between the atmosphere and the ocean. Between high and low latitudes a water transport q results, driven by the density gradient according to

$$q = k(\rho_2 - \rho_1) = k\rho_0 (\alpha(T_2 - T_1) + \beta(S_2 - S_1)) \quad , \quad (10.3)$$

which is the linear form of eq. 8.1. The balance of the salinity in the two boxes is

$$\frac{dS_1}{dt} = |q|(S_2 - S_1) + P \quad , \quad \frac{dS_2}{dt} = |q|(S_1 - S_2) - P \quad , \quad (10.4)$$

where $P > 0$ denotes the net water flux. In Eq. 10.4 the absolute value of q appears, because for the exchange the direction of the current is irrelevant. Stationary solutions for Eq. 10.4 can only be found if

$$\Delta S = S_2 - S_1 = \begin{cases} -\frac{\alpha\Delta T}{2\beta} \pm \sqrt{\left(\frac{\alpha\Delta T}{2\beta}\right)^2 - \frac{P}{k\rho_0\beta}} & q > 0 \quad , \\ -\frac{\alpha\Delta T}{2\beta} + \sqrt{\left(\frac{\alpha\Delta T}{2\beta}\right)^2 - \frac{P}{k\rho_0\beta}} & q < 0 \quad , \end{cases} \quad (10.5)$$

where $\Delta T = T_2 - T_1 < 0$. For the direct circulation, $q > 0$ and hence $\rho_2 > \rho_1$ two solutions are possible, one with a smaller contrast in salinity and one with a large negative ΔS . We now put

$$\delta = -\frac{\beta\Delta S}{\alpha\Delta T}, \quad E = \frac{\beta P}{k\rho_0(\alpha\Delta T)^2} \quad , \quad (10.6)$$

and obtain from (10.5)

$$\delta = \begin{cases} \frac{1}{2} \pm \sqrt{\frac{1}{4} - E} & q > 0 \quad , \\ \frac{1}{2} + \sqrt{\frac{1}{4} - E} & q < 0 \quad . \end{cases} \quad (10.7)$$

The transport q is given by

$$q = k\rho_0\alpha\Delta T(1 - \delta) \quad . \quad (10.8)$$

For $\delta > 1$ the circulation is indirect, meaning that water sinks, where it is warmer. In order to attain a sufficiently high density that enables a sinking, the salinity must be accordingly high. For $\delta < 1$ two solutions result, of which one is unstable (Fig. 10.5). For the direct circulation (water sinks where it is colder) $q > 0$. In case P increases, E and δ increase as well. But this leads to a decrease of q . An amplified hydrological cycle slows down the thermohaline circulation. For $0.5 < \delta < 1$ and hence $0 < q < \frac{1}{2}k\rho_0\alpha\Delta T$ the circulation is unstable. The model shows a threshold for q below which the thermohaline circulation does not exist. It must be noted, that in this simple model the meridional temperature contrast directly determines this threshold.

The existence of multiple equilibria of the thermohaline circulation can be made plausible by considering the heat and water transport as schematically illustrated in Fig. 10.6. In Fig. 10.6 (left), the overturning circulation is direct, i.e. water sinks where it is cold and upwells where it is warm. Under the typical profiles of T and S (with respectively high values at the surface) the ocean transports heat and salt northwards. The cycle of the fluxes is closed by an excess of heat in the equatorial region and a cooling in the north, and by the atmospheric water transport. However, the same water transport can also result from an opposite circulation as shown in Fig. 10.6 (right) in case the vertical gradient of S changes sign.

Hence, a significant rearrangement of the salinity distribution is necessary in order to provoke changes in the oceanic circulation. Due to the different ocean feedback processes associated with temperature and salinity perturbations, the surface salinity may change in a substantial way, which is a precondition for attaining state (b) in Fig. 10.6, with a very weak and shallow overturning circulation.

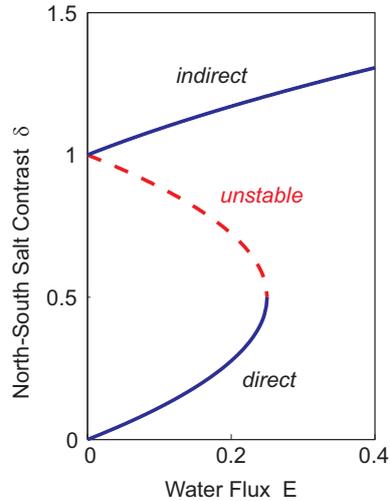


Figure 10.5: Multiple equilibria of the thermohaline circulation for different values of the water flux in the Stommel box model.

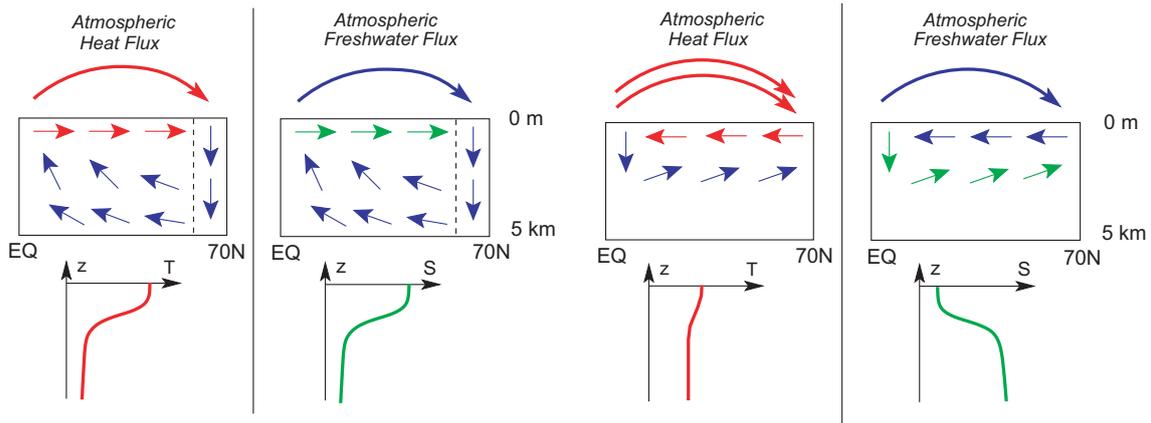


Figure 10.6: Schematic depiction of the thermohaline circulation and the meridional heat and water fluxes. (left) direct circulation: water sinks where it is cold; (right) indirect circulation: water sinks in warm areas. Figure from *Stocker and Marchal* [2000].

10.4 Multiple equilibria in coupled climate models

Manabe and Stouffer [1988] reported for the first time results from a coupled climate model in which, for present climate parameters, two different states were found. They primarily differed in their thermohaline circulation in the Atlantic. One of the states had an active deep water formation in the North Atlantic, the other state showed an overturning similar to the one in the Pacific. Transitions can be triggered by short-term perturbations in the water balance in the North Atlantic. Similar results were also found with other coupled models.

Therefore, it is probable that the deep water circulation in the Atlantic responds sensitively to changes in the water balance. This is a plausible mechanism to explain the abrupt transition found in climate time series from the last ice age (e.g., Fig. 10.1). One hypothesis claims that

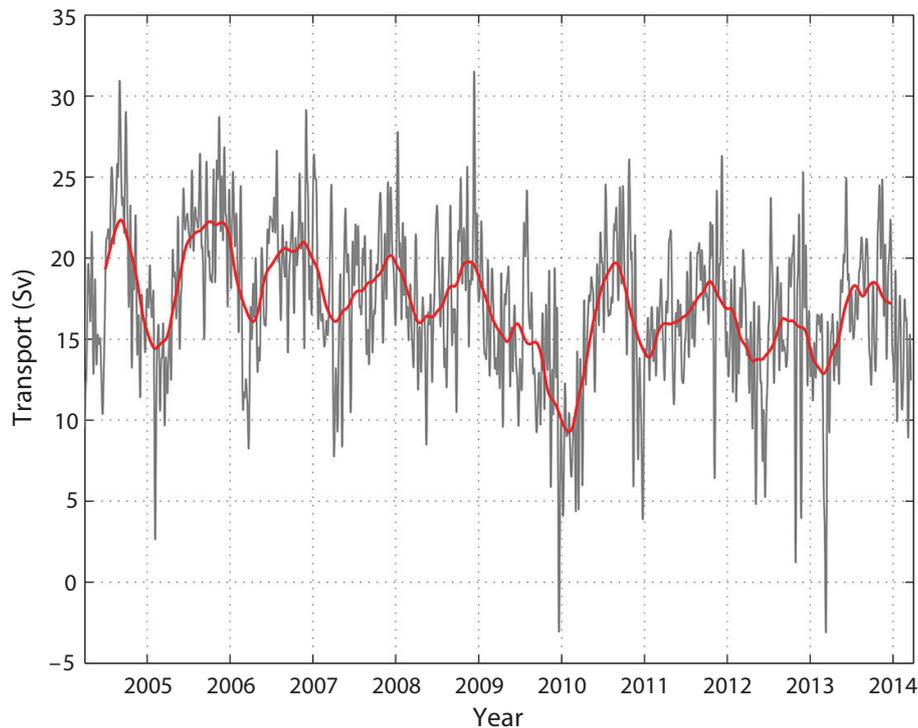


Figure 10.7: 180-day running mean (red) over 10 years of AMOC data observed at 26.5°N in the North Atlantic Ocean. These data indicate a slight downward trend. Whether this is the beginning of a slow AMOC shutdown is not confirmed yet because of the highly variable nature of the signal and the relatively short time series. Figure from *Srokosz and Bryden* [2015].

during glacial periods the ice sheets located around the North Atlantic cause large freshwater discharges owing to the advancement of ice streams. This situation was reinforced towards the end of the last glacial period, when the melting of the northern hemispheric ice sheets led to a sea level rise of about 120 m. During that time, the last abrupt climate transitions were observed.

However, the water balance is altered not only by the input of melting water, but depends also on the strength of the meridional water flux and hence on the hydrological cycle. In a warmer climate, the water cycle is generally intensified due to the increased amount of water in the form of vapour in the atmosphere and to the higher evaporation rates induced by higher temperatures. The question remains, whether this additional water input reaches the North Atlantic where it would increase the buoyancy and so reduce deep water formation. Some coupled models indicate that this is the case. This, in combination with the fact that a warming of the surface waters also reduces density and hence hinders the sinking of water masses, points to the conclusion that the Atlantic meridional overturning circulation (AMOC) is likely to weaken in the future. Whether a threshold will be exceeded and a complete shutdown of this – for Northern Europe crucial – current follows, is the object of current research. Since 2004, the international oceanographic program RAPID observes the currents and the water mass characteristics in the North Atlantic with a focus on the AMOC. Figure 10.7

shows the summary of these measurements.

Simple models show that a threshold of the circulation exists below which a complete shut-down of the current results without additional external inputs, meaning that there exists a transition to a second stable equilibrium. This behaviour is also simulated in more complete models such as 3-dimensional coupled atmosphere-ocean models as described in *Stocker* [2013]. Simulations with a simplified coupled model (Bern 2.5d model), consisting of a zonally averaged 3-basin ocean model and an energy balance model for the atmosphere, show that the crossing of a threshold depends on several important quantities in the climate system. Figure 10.8 gives a summary of the results. For the respective simulations, simplified CO₂-scenarios were chosen: after an exponential growth at different rates, the CO₂ concentration is held constant. The evolution of the thermohaline circulation may be split into two classes. The first class exhibits a linear behaviour in which a temporarily strong reduction of the circulation is simulated, followed by a recovery on the long-term. The reduction of the current is a response to the maximum value of the CO₂ increase and hence to the rate of warming. In the second class, the circulation shuts down completely and does not recover. An irreversible transition to the second stable equilibrium has been realized. It is interesting to notice that a reduction of the maximum concentration of CO₂ (from experiment 750 to 650) as well as a reduction of the rate of CO₂ increase (from experiment 750 to 750S) avoids the crossing of the critical threshold. Hence, the rate of future warming in the climate system plays a significant role. Depending on the rate and amount of warming, irreversible changes in the AMOC may result.

Similar behaviour is produced by a more complex model (Fig. 10.9). In this simulation with a state-of-the-art coupled climate model, the AMOC slows down over the 21st and 22nd centuries in an extreme business-as-usual scenario in which the global mean warming will exceed 8 °C by year 2500. In the scenarios with reductions of greenhouse gas emissions and hence lesser warming, the AMOC reduction is also limited and it slowly recovers after 2100. A bifurcation can be identified in these simulations which occurs between a global warming of 4 and 8 °C. It is characterised by near-linear reductions of the AMOC in response to the

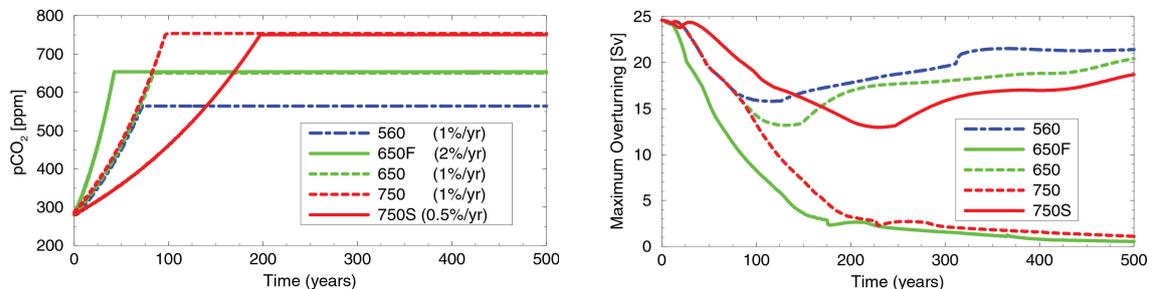


Figure 10.8: Simulations with the Bern 2.5d model for the evolution of the Atlantic meridional overturning circulation (AMOC) in the Atlantic considering a warming scenario. The different simplified CO₂-scenarios (left) consist of an exponential increase at different rates, levelling off at a given maximum value. The AMOC reveals a bifurcation in its behaviour: for small maximum values or slow rates of increase, the threshold for a complete shut-down is not exceeded. Figure from *Stocker and Schmittner* [1997].

warming and slow recovery for moderate warming, and, on the other hand, a near-complete shut-down of the AMOC for high levels of warming. This represents a "tipping point" in the AMOC which is consistent with the simulations based on a much simpler model (Fig. 10.8).

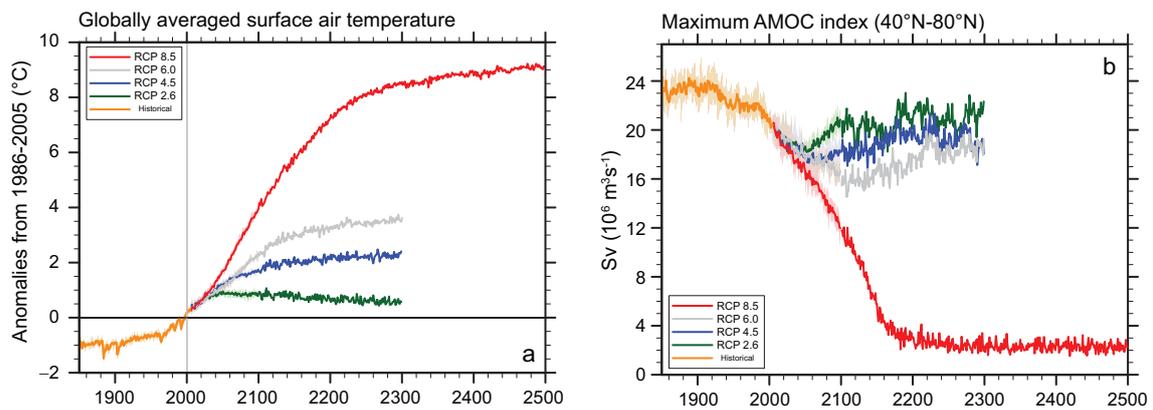


Figure 10.9: Long-term simulations with a comprehensive coupled climate model. a) Global mean temperature changes in four emissions scenarios. b) Response of the Atlantic meridional overturning circulation. Figure from *Stocker* [2013].

Table of constants

Name	Symbol	Value
Planck constant	h	$6.626 \cdot 10^{-34} \text{ J s}$
Speed of light	c	$2.998 \cdot 10^8 \text{ m s}^{-1}$
Boltzmann constant	$k = R/N_A$	$1.38 \cdot 10^{-23} \text{ J K}^{-1}$
Ideal gas constant	R	$8.314 \text{ J mol}^{-1} \text{ K}^{-1}$
Avogadro number	N_A	$6.022 \cdot 10^{23} \text{ mol}^{-1}$
Stefan Boltzmann number	σ	$5.67 \cdot 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$
Solar constant	S_0	1368 W m^{-2}
Distance Earth-Sun	d_E	$150 \cdot 10^9 \text{ m}$
Radius of Earth	R, R_E	$6.371 \cdot 10^6 \text{ m}$
Angular velocity of Earth	Ω	$7.27 \cdot 10^{-5} \text{ s}^{-1}$
Standard pressure	p_0	$1.013 \cdot 10^5 \text{ Pa}$
Standard temperature	T_0	273.15 K
Molar mass of air	M_{mol}	$0.029 \text{ kg mol}^{-1}$
Mass of proton	m_p	$1.67 \cdot 10^{-27} \text{ kg}$
Mass of electron	m_e	$9.11 \cdot 10^{-31} \text{ kg}$
Mass of neutron	m_n	$1.67 \cdot 10^{-27} \text{ kg}$
Elementary charge	e	$1.602 \cdot 10^{-19} \text{ C}$

Table 10.1: Table of constants

References

- Adkins, J. F., K. McIntyre, and S. D. P., The salinity, temperature, and $\delta^{18}\text{O}$ of the glacial deep ocean, *Science*, *298*, 1769–1773, 2002.
- Aeschbach-Hertig, W., F. Peeters, U. Beyerle, and R. Kipfer, Palaeotemperature reconstruction from noble gases in ground water taking into account equilibration with entrapped air, *Nature*, *405*, 1040–1044, 2000.
- Andrews, J. N., S. N. Davis, J. Fabrykamartin, J. C. Fontes, B. E. Lehmann, H. H. Loosli, J. L. Michelot, H. Moser, B. Smith, and M. Wolf, The insitu production of radioisotopes in rock matrices with particular reference to the stripa granite, *Geochimica et Cosmochimica Acta*, *53*, 1803–1815, 1989.
- Archer, D., A. Winguth, D. Lea, and N. Mahowald, What caused the glacial/interglacial atmospheric pCO_2 cycles?, *Reviews of Geophysics*, *12*, 159–189, 2000.
- Aston, F. W., Isotopes and atomic weights, *Nature*, *105*, 617–619, 1920.
- BAG, 20 Jahre seit dem Reaktorunfall von Tschernobyl: Die Auswirkungen auf die Schweiz, *Bundesamt für Gesundheit, Tech. Rep.*, 4pp, <http://www.bag.admin.ch>, 2006.
- Battisti, D. S., Dynamics and thermodynamics of a warming event in a coupled tropical atmosphere-ocean model, *Journal of Atmospheric Sciences*, *45*, 2889–2919, 1988.
- Battisti, D. S., and A. C. Hirst, Interannual variability in a tropical atmosphere-ocean model: Influence of the basic state, ocean geometry and nonlinearity, *Journal of Atmospheric Sciences*, *46*, 1687–1712, 1989.
- Battisti, D. S., A. C. Hirst, and S. E. S., Instability and predictability in coupled atmosphere-ocean models, *Phil. Trans. R. Soc. London*, *329*, 237–247, 1989.
- Beer, J., S. Baumgartner, B. Dittrich-Hannen, J. Hauenstein, P. W. Kubuk, C. Lukaszczyk, W. Mende, R. Stellmacher, and M. Suter, 1994, *Solar variability traced by cosmogenic isotopes*, pp. 291–300. Cambridge University Press.
- Bigeleis, J., Chemistry of isotopes, *Science*, *147*(3657), 463+, 1965.
- Bjerknes, J., Atmospheric teleconnections from the equatorial Pacific, *Monthly Weather Review*, *97*, 163–172, 1969.
- Blunier, T., and E. J. Brook, Timing of millennial-scale climate change in Antarctica and Greenland during the last glacial period, *Science*, *291*, 109–112, 2001.
- Bollhöfer, A., C. Schlosser, S. Schmid, M. Konrad, R. Purtschert, and R. Kraus, Half a century of krypton-85 activity concentration measured in air over central Europe: Trends and relevance for dating young groundwater, *Journal of environmental radioactivity*, *205*, 7–16, 2019.

- Bopp, L., K. E. Kohfeld, C. Le Quere, and O. Aumont, Dust impact on marine biota and atmospheric CO₂ in glacial periods, *Geochimica et Cosmochimica Acta*, 66, A91–A91, 2002.
- Broecker, W. S., The biggest chill, *Natural History*, 96, 74–82, 1987.
- Broecker, W. S., Paleocean circulation during the last deglaciation: A bipolar seesaw?, *Paleoceanography*, 13, 119–121, 1998.
- Bryan, F., High-latitude salinity effects and interhemispheric thermohaline circulations, *Nature*, 323, 301–304, 1986.
- Bütikofer, R., E. O. Flückiger, L. Desorgher, and M. R. Moser, The extreme solar cosmic ray particle event on 20 January 2005 and its influence on the radiation dose rate at aircraft altitude, *Science of the Total Environment*, 391, 177–183, 2008.
- Chahine, M., The hydrological cycle and its influence on climate, *Nature*, 359, 373–380, 1992.
- Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, C. A., R. DeFries, J. Galloway, M. Heimann, C. Jones, C. Le Qur, R. Myneni, S. Piao, and P. Thornton, 2013, *Carbon and Other Biogeochemical Cycles*, pp. 465–570, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- Clark, P. U., J. D. Shakun, S. Marcott, A. C. Mix, M. Eby, S. Kulp, A. Levermann, G. A. Milne, P. L. Pfister, B. D. Santer, D. P. Schrag, S. Solomon, T. F. Stocker, B. H. Strauss, A. J. Weaver, R. Winkelmann, D. Archer, E. Bard, A. Goldner, K. Lambeck, R. T. Pierrehumbert, and G.-K. Plattner, Consequences of twenty-first-century policy for multi-millennial climate and sea-level change, *Nature Climate Change*, 6, 360–369, 2016.
- Craig, H., Standard for reporting concentrations of deuterium and oxygen-18 in natural waters, *Science*, 133(346), 1833–1834, 1961.
- Dansgaard, W., Stable isotopes in precipitation, *Tellus*, 16, 436–468, 1964.
- Denman, K. L., G. Brasseur, A. Chidthaisong, P. Ciais, P. Cox, R. E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S. Ramachandra, P. L. da Silva Dias, S. Wofsy, and X. Zhang, 2007, *Couplings between changes in the climate system and biogeochemistry*, pp. 498–587, Cambridge United Kingdom and New York, NY, USA: Cambridge University Press.
- Dommenget, D., and M. Latif, A cautionary note on the interpretation of EOFs, *Journal of Climate*, 15, 216–225, 2002.
- EPICA community members, Eight glacial cycles from an Antarctic ice core, *Nature*, 429, 623–628, 2004.
- Fischer, H., M.-L. Siggaard-Andersen, U. Ruth, R. Rthlisberger, and E. Wolff, Glacial/interglacial changes in mineral dust and sea salt records in polar ice cores: sources,

transport, deposition, *Reviews of Geophysics*, 45, doi:10.1029/2005RG000192, RG1002, 2007.

Geiss, J., H. Oeschger, and U. Schwarz, The history of cosmic radiation as revealed by isotopic changes in the meteorites and on the earth, *Space Science Reviews*, 1, 197–223, 1962.

Genthon, C., Simulations of desert dust and sea-salt aerosols in Antarctica with a general circulation model of the atmosphere, *Tellus*, 44B, 371–389, 1992.

Gonfiantini, R., W. Stichler, and R. K., Standards and intercomparison materials, *International Atomic Energy Agency for stable isotope measurements*, Tech. Rep., No. IAEA-TECDOC-825, 1995.

Heisinger, B., and E. Nolte, Cosmogenic in situ production of radionuclides: Exposure ages and erosion rates, *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, 172, 790–795, 2000.

Houghton, R. A., Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850-2000, *Tellus Series B-Chemical and Physical Meteorology*, 55, 378–390, 2003.

IPCC, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 2007.

IPCC, *Climate Change 2013: The Physical Science Basis Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 2013.

IPCC, *IPCC 2019: Special Report on the Ocean and Cryosphere in a Changing Climate*, Cambridge University Press, Cambridge, 2019.

Joos, F., R. Meyer, M. Bruno, and M. Leuenberger, The variability in the carbon sinks as reconstructed for the last 1000 years, *Geophysical Research Letters*, 26(10), 1437–1440, 1999.

Jouzel, J., V. Masson-Delmotte, O. Cattani, G. Dreyfus, S. Falourd, G. Hoffmann, B. Minster, J. Nouet, J.-M. Barnola, J. Chappellaz, H. Fischer, J. Gallet, S. Johnsen, M. Leuenberger, L. Loulergue, D. Luethi, H. Oerter, F. Parrenin, G. Raisbeck, D. Raynaud, J. Schwander, R. Spahni, R. Souchez, E. Selmo, A. Schilt, J. Steffensen, B. Stenni, B. Stauffer, T. Stocker, J.-L. Tison, M. Werner, and E. Wolff, Orbital and millennial antarctic climate variability over the last 800,000 years, *Science*, 317, 793–796, 2007.

Keeling, R. F., S. C. Piper, and M. Heimann, Global and hemispheric CO₂-sinks deduced from changes in atmospheric O₂ concentrations, *Nature*, 381, 218–221, 1996.

- Key, R., A. Kozyr, C. Sabine, K. Lee, R. Wanninkhof, J. Bullister, R. Feely, F. Millero, C. Mordy, and T.-H. Peng, A global ocean carbon climatology: Results from global data analysis project (GLODAP), *Global Biogeochemical Cycles*, *18*, doi:10.1029/2004GB002247, GB4031, 2004.
- Kohfeld, K., C. LeQur, S. Harrison, and R. Anderson, Role of marine biology in glacial-interglacial CO₂ cycles, *Science*, *308*, 74–78, 2005.
- Köhler, P., H. Fischer, G. Munhoven, and R. E. Zeebe, Quantitative interpretation of atmospheric carbon records over the last glacial termination, *Global Biogeochem. Cycles*, *19*, DOI: 10.1029/2004GB002345, 2005.
- Lal, D., and B. Peters, 1967, *Cosmic-ray produced radioactivity on the Earth.*, Volume 46/2, pp. 551–612, Berlin: Springer Verlag,.
- Lehmann, B. E., M. Lehmann, A. Neff, and S. V. Tarakanov, Radon-222 monitoring of soil diffusivity, *Geophysical Research Letters*, *27*, 3917–3920, 2000.
- Levitus, S., and T. P. Boyer, *NOAA Atlas NESDIS 4, World Ocean Atlas 1994, Volume 4: Temperature*, NOAA, U.S. Dept. Commerce, 117 pp., 1994.
- Lindner, G., W. Pfeiffer, J. ROBBINS, and E. Recknagel, 1989, *Long-lived Chernobyl radionuclides in Lake Constance: Speciation, sedimentation, and biological transfer*, pp. 295–300. Visby, Sweden, Verlag, TUV Rheinland GMBH, Köln.
- Loulergue, L., A. Schilt, R. Spahni, V. Masson-Delmotte, T. Blunier, B. Lemieux, J.-M. Barnola, D. Raynaud, T. Stocker, and J. Chappellaz, Orbital and millennial-scale features of atmospheric CH₄ over the past 800,000 years, *Nature*, *453*, 383–386, 2008.
- Lüthi, D., M. Le Floch, B. Bereiter, T. Blunier, J. M. Barnola, U. Siegenthaler, D. Raynaud, J. Jouzel, H. Fischer, K. Kawamura, and T. F. Stocker, High-resolution carbon dioxide concentration record 650,000–800,000 years before present, *Nature*, *453*, 379–382, 2008.
- Lyons, T. W., C. T. Reinhard, and N. J. Planavsky, The rise of oxygen in earth's early ocean and atmosphere, *Nature*, *506*, 307–315, 2014.
- Manabe, S., and R. J. Stouffer, Two stable equilibria of a coupled ocean atmosphere model, *Journal of Climate*, *1*, 841–866, 1988.
- Marotzke, J., Abrupt climate change and thermohaline circulation: mechanisms and predictability, *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 1347–1350, 2000.
- Martcorena, B., and G. Bergametti, Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme, *Journal of Geophysical Research*, *100*, 16415–16430, 1995.
- Munk, W., Abyssal recipes, *Deep-Sea Research*, *13*, 707–730, 1966.

-
- Muscheler, R., F. Joos, J. Beer, S. A. Muller, M. Vonmoos, and I. Snowball, Solar activity during the last 1000 yr inferred from radionuclide records, *Quaternary Science Reviews*, *26*, 82–97, 2007.
- Naegler, T., and I. Levin, Closing the global radiocarbon budget 1945-2005, *Journal of Geophysical Research*, *111*, 14 –14, D12311, 2006.
- Neelin, J. D., , D. S. Battisti, A. C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. E. Zebiak, ENSO theory, *Journal of Geophysical Research*, *103*, 14261–14290, 1998.
- Neftel, A., E. Moor, H. Oeschger, and B. Stauffer, Evidence from polar ice cores for the increase in atmospheric CO₂ in the past 2 centuries, *Nature*, *315*, 45–47, 1985.
- Oeschger, H., The carbon-dioxide problem - an interdisciplinary survey - conclusions, *Experientia*, *36*, 808–812, 1980.
- Oeschger, H., U. Siegenthaler, U. Schotterer, and G. A., A box diffusion model to study the carbon dioxide exchange in nature, *Tellus*, *27*, 170–192, 1975.
- Petit, J. R., J. Jouzel, D. Raynaud, N. I. Barkov, J.-M. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davis, G. Delaygue, M. Delmotte, V. M. Kotlyakov, M. Legrand, V. Y. Lipenkov, C. Lorius, L. Pepin, C. Ritz, E. Saltzman, and M. Stievenard, Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature*, *399*, 429–436, 1999.
- Prospero, J., P. Ginoux, O. Torres, S. Nicholson, and T. Gill, Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 total ozone mapping spectrometer (TOMS) absorbing aerosol product, *Reviews of Geophysics*, *40*, doi:10.1029/2000RG000095, 2002.
- Reimer, P. J., E. Bard, A. Bayliss, J. W. Beck, P. G. Blackwell, C. B. Ramsey, C. E. Buck, H. Cheng, R. L. Edwards, M. Friedrich, P. M. Grootes, T. P. Guilderson, H. Haffidason, I. Hajdas, C. Hatte, T. J. Heaton, D. Hoffmann, A. G. Hogg, K. A. Hughen, K. F. Kaiser, B. Kromer, S. W. Manning, M. Niu, R. W. Reimer, D. A. Richards, E. M. Scott, J. R. Southon, R. A. Staff, C. S. M. Turner, and J. van der Plicht, IntCal13 and Marine13 radiocarbon age calibration curves 0 – 50,000 years cal BP, *Radiocarbon*, *55*, 1869–1887, 2013.
- Reiter, E., 1978, *Part 4: Radioactive Tracers*, Springfield VA: US Department of Energy.
- Revelle, R., and H. Suess, Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO₂ during the past decades, *Tellus*, *9*, 18–27, 1957.
- Roedel, W., *Physik unserer Umwelt - Die Atmosphre*, Springer, Heidelberg, 1994.
- Sabine, C. L., M. Heimann, P. Artaxo, D. C. E. Bakker, C. T. A. Chen, C. B. Field, N. Gruber, C. Le Qur, R. Prinn, J. E. Richey, P. R. Lankao, J. A. Sathaye, and R. Valentini,

2004, *Current Status and Past Trends of the Global Carbon Cycle*, pp. 165–186. Island Press, Washington DC, USA.

Santschi, P. H., Tschernobyl Radionuklide in der Umwelt: Tracer für die enge Kopplung zwischen atmosphärischen, terrestrischen und aquatischen Systemen, *EAWAG Mitteilungsblatt*, 23, 2–9, 1987.

Sarachik, E. S., and M. A. Cane, *The El Niño-Southern Oscillation Phenomenon*, Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK, 2010.

Sarmiento, J. L., and N. Gruber, *Ocean Biogeochemical Dynamics*, Princeton University Press, Market Place, Woodstock, Oxfordshire OX20 1SY, 2006.

Schilt, A., M. Baumgartner, T. Blunier, J. Schwander, R. Spahni, H. Fischer, and T. F. Stocker, Glacial-interglacial and millennial-scale variations in the atmospheric nitrous oxide concentration during the last 800 000 years, *Quaternary Science Reviews*, 29, 182–192, 2010.

Schopf, P. S., and M. J. Suarez, Vacillations in a coupled ocean-atmosphere model, *Journal of Atmospheric Sciences*, 45, 549–566, 1988.

Seinfeld, J., and S. Pandis, *Atmospheric chemistry and physics*, John Wiley & Sons, Hoboken, NJ, 2006.

Siegenthaler, U., and H. Oeschger, Predicting future atmospheric carbon dioxide levels, *Science*, 199, 388–395, 1978.

Siegenthaler, U., and H. Oeschger, Correlation of ^{18}O in precipitation with temperature and altitude, *Nature*, 285, 314–317, 1980.

Siegenthaler, U., T. Stocker, E. Monnin, D. Lthi, J. Schwander, B. Stauffer, D. Raynaud, J.-M. Barnola, H. Fischer, V. Masson-Delmotte, and J. Jouzel, Stable carbon cycle-climate relationship during the late pleistocene, *Science*, 310, 1313–1317, 2005.

Sigman, D. M., and E. Boyle, Glacial/interglacial variations in atmospheric carbon dioxide, *Nature*, 407, 859–869, 2000.

Soddy, F., The structure of the atom, *Nature*, 92, 452–452, 1914.

Spahni, R., J. Chappellaz, T. F. Stocker, L. Loulergue, G. Hausamann, K. Kawamura, J. Flckiger, J. Schwander, D. Raynaud, V. Masson-Delmotte, and J. Jouzel, Atmospheric methane and nitrous oxide of the late pleistocene from antarctic ice cores, *Science*, 310, 1317–1321, 2005.

Speer, K., and G. Forget, 2013, *Global distribution and formation of mode waters*, pp. 211–226. Academic Press.

Srokosz, M. A., and H. L. Bryden, Observing the Atlantic Meridional Overturning Circulation yields a decade of inevitable surprises, *Science*, 348, doi: 10.1126/science.1255575, 2015.

-
- Stephenson, D. B., V. Pavan, and R. Bojariu, Is the North Atlantic Oscillation a random walk?, *International Journal of Climatology*, *20*, 1–18, 2000.
- Stocker, T., 2013, *The ocean as a component of the climate system*, pp. 3–30. Academic Press.
- Stocker, T., and S. J. Johnsen, A minimum thermodynamic model for the bipolar seesaw, *Paleoceanography*, *18*, 1087, 2003.
- Stocker, T. F., The seesaw effect, *Science*, *282*, 61–62, 1998.
- Stocker, T. F., and O. Marchal, Abrupt climate change in the computer: Is it real?, *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 1362–1365, 2000.
- Stocker, T. F., and A. Schmittner, Influence of CO₂ emission rates on the stability of the thermohaline circulation, *Nature*, *388*, 862–865, 1997.
- Stommel, H., The westward intensification of wind-driven ocean currents, *Transactions of the American Geophysical Union*, *29*, 202–206, 1948.
- Stommel, H., The abyssal circulation, *Deep-Sea Research*, *5*, 80–82, 1958.
- Stommel, H., Thermohaline convection with two stable regimes of flow, *Tellus*, *13*, 224–230, 1961.
- Stommel, H., and A. B. Arons, On the abyssal circulation of the world ocean - I. Stationary planetary flow patterns on a sphere, *Deep-Sea Research*, *6*, 140–154, 1960.
- Tarnocai, C., J. G. Canadell, E. A. G. Schuur, P. Kuhry, G. Mazhitova, and S. Zimov, Soil organic carbon pools in the northern circumpolar permafrost region, *Global Biogeochemical Cycles*, *23*, 1–11, GB2023, 2009.
- Thomson, J. J., The appearance of helium and neon in vacuum tubes, *Nature*, *90*, 645–647, 1913.
- Turekian, K. K., and W. C. Graustein, 2003, *Natural Radionuclides in the Atmosphere*, pp. 261–279. Elsevier.
- von Arx, W. S., *Introduction to Physical Oceanography*, Addison-Wesley, 1979.
- von Storch, H., and F. W. Zwiers, *Statistical Analysis in Climate Research*, Cambridge University Press, The Edinburgh Building, Cambridge CB2 2R, 1999.
- Wagenbach, D., F. Ducroz, R. Mulvaney, L. Keck, A. Minikin, M. Legrand, J. S. Hall, and E. W. Wolff, Sea-salt aerosol in coastal antarctic regions, *Journal of Geophysical Research*, *103*(D9), 10961–10974, 1998.
- Walker, G. T., Correlations in seasonal variations of weather, *IX. Mem. Ind. Meteorol. Dept.*, *24*, 53–84, 1924.

Warneck, P., *Chemistry of the natural atmosphere*, Academic Press, San Diego, CA, 1987.

Webb, R. S., D. H. Rind, S. J. Lehman, R. J. Healy, and D. Sigman, Influence of ocean heat transport on the climate of the last glacial maximum, *Nature*, *385*, 695–699, 1997.

Wild, M., D. Folini, C. Schr, N. Loeb, E. Dutton, and G. Knig-Langlo, The global energy balance from a surface perspective, *Climate Dynamics*, *40*, 3107–3134, 2013.

Worthington, L. V., 1981, *The water masses of the world ocean: Some results of a fine-scale census*, pp. 42–69. MIT Press.

Yang, X., J. Pyle, and R. Cox, Sea salt aerosol production and bromine release: Role of snow on sea ice, *Geophysical Research Letters*, *35*, L16815, 2008.