OCEAN BIOGEOCHEMICAL DYNAMICS

CHAPTER 2
TRACER CONSERVATION AND OCEAN TRANSPORT

JORGE L. SARMIENTO
Atmospheric and Oceanic Sciences Program
Princeton University
Sayre Hall, Forrestal Campus
P.O. Box CN710
Princeton, NJ 08544-0710
U.S.A.

e-mail: jls@princeton.edu

NICOLAS GRUBER
Department of Atmospheric Sciences and IGPP
5853 Slichter Hall
University of California, Los Angeles
Los Angeles, CA 90095-1567
U.S.A.

e-mail: gruber@atmos.ucla.edu

September 11, 2004

©J. L. SARMIENTO AND N. GRUBER
September 14, 2004
Chapter 2

TRACER CONSERVATION AND OCEAN TRANSPORT

The tracer conservation equation establishes the relationship between the time rate of change of tracer concentration at a given point and the processes that can change that concentration. These processes include water transport by advection and mixing, and sources and sinks due to biological and chemical transformations. This book is dedicated primarily to a study of the sources and sinks due to biological and chemical transformations. However, we begin our study in this chapter with a derivation of the conservation equation, including the transport terms, accompanied by an overview of ocean circulation. The ocean circulation overview begins with a discussion of advection driven by wind forcing. The next two subsections focus on the wind driven circulation in the upper kilometer of the ocean in the presence of stratification, and also the deep ocean circulation below that. A final section discusses non-steady state behavior of the oceanic circulation, with a focus on meso-scale eddies and interannual to decadal variability of the upper ocean.

Our discussion of ocean circulation focuses on only a few major features that are critical for understanding ocean biogeochemistry. Textbooks such as Pond and Pickard [1983] provide a more detailed discussion of the basic theory of ocean circulation (cf. also Pedlosky [1996]), and Siedler et al. [2001] provides an up-to-date overview of observations and modeling focused primarily on results from the World Ocean Circulation Experiment (WOCE) carried out in the 1990’s (cf. also Warren and Wunsch [1981]).

2.1 Tracer Conservation Equation

The tracer conservation equation for a volume of water at a fixed location has the form:

\[
\frac{\partial C}{\partial t} = \left. \frac{\partial C}{\partial t} \right|_{\text{advection}} + \left. \frac{\partial C}{\partial t} \right|_{\text{diffusion}} + SMS(C) \tag{2.1.1}
\]

where the term \(SMS(C)\) (mmol m\(^{-3}\) s\(^{-1}\)) represents internal sources minus sinks. We use \(C\) to represent tracer concentration in mmol m\(^3\), as in Chapter 1. Geochemists generally prefer to report concentration in units of moles per unit mass of water (micromoles per kilogram, i.e., \(\mu\text{mol kg}^{-1}\)) rather than per unit volume. This is because water is compressible and the volume is thus a function of the pressure and temperature, whereas the mass is not. As a consequence, mass is conserved with changes in state, whereas
volume is not. However, the per unit mass unit is unwieldy to use in the conservation equation because it requires including density and all of its derivatives. Furthermore, most present models of the ocean are based on approximations that imply conservation of volume. We will therefore generally use units of mmol m$^{-3}$ for concentration. These can be converted to $\mu$mol kg$^{-1}$ by dividing by the density in kg m$^{-3}$ x 10$^{-3}$.

In what follows we first derive the advection and diffusion components, and then consider the application of the conservation equation to box models such as those described in Chapter 1.

**ADVECTION AND DIFFUSION COMPONENTS**

We derive the advection and diffusion components of the tracer conservation equation by considering an infinitesimally small cube of water with fixed volume, such as that shown in Figure 2.1.1. The total amount of tracer in the cube is equal to the concentration $C$ times the volume of the cube $V$ (m$^3$). Panel 2.1.1 shows how we derive the advective contribution to the amount of tracer contained in the volume. The final equation for the advective component of the conservation equation is

$$\frac{\partial C}{\partial t}_{\text{advection}} = -\frac{\partial (C \cdot u)}{\partial x} - \frac{\partial (C \cdot v)}{\partial y} - \frac{\partial (C \cdot w)}{\partial z} \quad (2.1.2)$$

The velocities, $u$, $v$, and $w$, are the vector components of the advection in the $x$, $y$, and $z$ directions, respectively. By convention, $x$ is taken as positive towards the east, $y$ is positive towards the north, and $z$ is positive upwards. The velocities have units of m s$^{-1}$.

We next add the effect of diffusion across the walls of the cube. Diffusion is the flux of tracer that occurs due in part to the random motion of tracer molecules referred to as *molecular diffusion*. It also generally includes an *eddy diffusion* component that represents the net effect of small-scale motions of water parcels that do not result in net advection. Thus, if one face of the cube has equal amounts of water flowing in and out, but if the water flowing in generally has a higher tracer concentration on average than the water flowing out, there will be a net transport of tracer into the cube even though there is no net advection. One can readily see that the magnitude of this term will depend on the size of the volume under consideration, with larger volumes having larger eddy diffusion contributions. It will also depend on the amount of water flowing in and out of the face of the cube, with larger exchanges resulting in greater eddy diffusion.

The tracer flux resulting from molecular diffusion is directly proportional to the gradient in tracer concentration, with the constant of proportionality defined as the diffusivity. Fick's First Law gives this relationship as:

© J. L. SARMIENTO AND N. GRUBER

September 14, 2004
\[ \Phi_x = -\varepsilon \frac{\partial C}{\partial x}, \quad \Phi_y = -\varepsilon \frac{\partial C}{\partial y}, \quad \Phi_z = -\varepsilon \frac{\partial C}{\partial z} \quad (2.1.3) \]

The flux \( \Phi \) is in units of mol m\(^{-2}\) s\(^{-1}\), and the molecular diffusivity \( \varepsilon \) is in m\(^2\) s\(^{-1}\). Fick’s Second Law relates the change in concentration to the gradient in the diffusive flux.

Consider a tracer flux \( \Phi_x \cdot \Delta y \cdot \Delta z \) entering the left face of the cube in Figure 2.1.1 and a flux \( \Phi_x + \frac{\partial \Phi_x}{\partial x} \cdot \Delta x \cdot \Delta y \cdot \Delta z \) leaving the right face of the cube. The effect of the diffusive flux on the total amount of tracer in the box \( V \cdot \frac{\partial C}{\partial t} \) is obtained by subtracting the second expression from the first, i.e., \( V \cdot \left( \frac{\partial C}{\partial t} \right) = -V \cdot \left( \frac{\partial \Phi_x}{\partial x} \right) \). The derivation of similar expressions for the \( y \) and \( z \) directions gives:

\[ \frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left( \varepsilon \frac{\partial C}{\partial x} \right) + \frac{\partial}{\partial y} \left( \varepsilon \frac{\partial C}{\partial y} \right) + \frac{\partial}{\partial z} \left( \varepsilon \frac{\partial C}{\partial z} \right) \quad (2.1.4) \]

Substituting (2.1.3) into (2.1.4) gives:

\[ \frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left[ (D_x + \varepsilon) \frac{\partial C}{\partial x} \right] + \frac{\partial}{\partial y} \left[ (D_y + \varepsilon) \frac{\partial C}{\partial y} \right] + \frac{\partial}{\partial z} \left[ (D_z + \varepsilon) \frac{\partial C}{\partial z} \right] \quad (2.1.6) \]

The transport of tracer by eddy diffusion is generally far larger than transport by molecular diffusion. Panel 2.1.2 shows a derivation of the eddy diffusion processes that treats the water motions as consisting of a time mean component and random variations around that mean. This treatment shows that transport of tracer by eddies results from the degree of correlation between the direction of the random motions of the water parcels and the tracer concentration. For example, if water parcels moving in the positive \( x \) direction have higher concentrations on the average than water parcels moving in the negative \( x \) direction, the tracer will be transported in the positive \( x \) direction.

Unfortunately, we generally do not have adequate information on the flow field and concentrations to calculate the necessary correlations. The usual practice is to approximate eddy diffusion by assuming it behaves like Fickian diffusion, with one important difference: eddy diffusion in the ocean is much larger in the horizontal than in the vertical. We therefore allow for the possibility of different magnitudes of mixing in each direction. The eddy and molecular contributions to the diffusivity are assumed to be additive, giving:

\[ \frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left[ (D_x + \varepsilon) \frac{\partial C}{\partial x} \right] + \frac{\partial}{\partial y} \left[ (D_y + \varepsilon) \frac{\partial C}{\partial y} \right] + \frac{\partial}{\partial z} \left[ (D_z + \varepsilon) \frac{\partial C}{\partial z} \right] \quad (2.1.6) \]
\[
\frac{\partial C}{\partial t} = -\frac{\partial (C \cdot u)}{\partial x} - \frac{\partial (C \cdot v)}{\partial y} - \frac{\partial (C \cdot w)}{\partial z} \\
+ \frac{\partial}{\partial x}\left[(D_x + \varepsilon) \cdot \frac{\partial C}{\partial x}\right] + \frac{\partial}{\partial y}\left[(D_y + \varepsilon) \cdot \frac{\partial C}{\partial y}\right] + \frac{\partial}{\partial z}\left[(D_z + \varepsilon) \cdot \frac{\partial C}{\partial z}\right] (2.1.7)
\]

We conclude the derivation by simplifying the tracer conservation equation. We expand the advection component (2.1.2) using the product rule:

\[
-\frac{\partial (C \cdot u)}{\partial x} - \frac{\partial (C \cdot v)}{\partial y} - \frac{\partial (C \cdot w)}{\partial z} = -C \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) \\
- \left( u \cdot \frac{\partial C}{\partial x} + v \cdot \frac{\partial C}{\partial y} + w \cdot \frac{\partial C}{\partial z} \right) (2.1.8)
\]

The mass conservation equation for incompressible flow derived in Panel 2.1.3 enables us to cancel the first term in parentheses on the right-hand side of (2.1.8). This gives:

\[
\frac{\partial C}{\partial t} = -u \cdot \frac{\partial C}{\partial x} - v \cdot \frac{\partial C}{\partial y} - w \cdot \frac{\partial C}{\partial z} \\
+ \frac{\partial}{\partial x}\left[(D_x + \varepsilon) \cdot \frac{\partial C}{\partial x}\right] + \frac{\partial}{\partial y}\left[(D_y + \varepsilon) \cdot \frac{\partial C}{\partial y}\right] + \frac{\partial}{\partial z}\left[(D_z + \varepsilon) \cdot \frac{\partial C}{\partial z}\right] (2.1.9)
\]

Finally, we make use of the definition of the divergence operator:

\[
\nabla = i \frac{\partial (\ )}{\partial x} + j \frac{\partial (\ )}{\partial y} + k \frac{\partial (\ )}{\partial z} (2.1.10)
\]

to obtain:

\[
\frac{\partial C}{\partial t} = -\mathbf{U} \cdot \nabla C + \nabla \cdot (\mathbf{D} \cdot \nabla C) + SMS(C) (2.1.11)
\]

where \(i, j, \) and \(k\) are unit vectors in the \(x, y, \) and \(z\) directions, respectively, \(\mathbf{U}\) is the velocity vector \(i u + j v + k w\), and \(\mathbf{D}\) is a diffusivity tensor with diagonal terms equal to \(D_x + \varepsilon, D_y + \varepsilon, \) and \(D_z + \varepsilon\).
Box models such as those described in the first chapter are the quantitative tool most frequently used by geochemists to investigate the behavior of chemicals. We define a box that encloses some region of particular interest, for example the surface ocean. The goal of developing a box model is to determine the processes that control the mean tracer concentration within that box. The usual way of proceeding is to assemble an equation that relates the change in mean concentration within the box to the fluxes across the boundaries of the box and sources and sinks internal to the box. In our example of a surface box, inputs across the boundaries of the box might be river input, air-sea fluxes, and upward flow of deep waters into the box. Losses across the boundaries of the box might include downward flow of surface waters. A sink internal to the box might be biological uptake, and an internal source might be remineralization of organic matter.

We demonstrate next how to construct a box model using the conservation equation (2.1.11). We wish to have an equation in terms of the mean concentration of the box. We accomplish this by integrating (2.1.11) over the volume of the box. The time rate of change term now represents the total change within the box. Dividing by the volume of the box will give the time rate of change of the mean concentration within the box. Note that our model does not require that the concentration be uniform within the box. A common misconception of geochemists is that box models require a uniform concentration within the box. This is useful as an assumption and is commonly employed, but it is not a necessary condition.

We next discuss the processes that can change the mean concentration within the box of our model. These include fluxes across the faces of the box. For example, a box at the surface of the ocean would have gas exchange with the atmosphere as well as advective and diffusive exchange with the layer immediately below. A box representing the deep ocean would have exchange with upper ocean boxes as well as exchange fluxes with the sediments. When we perform the volume integral of (2.1.11) as described in the previous paragraph, we also integrate the advection and diffusion terms over the entire volume. However, it is only at the boundaries of the box that these terms can provide a net flux into or out of the box that will affect the mean concentration. Advection and diffusion within the box move tracer around but do not change the mean concentration. It is therefore useful to transform the volume integral of the advection and diffusion terms into a surface integral using the Divergence Theorem:

\[
\int\![-\nabla \cdot \mathbf{U} \cdot \nabla C + \nabla \cdot (\mathbf{D} \cdot \nabla C)]dV = -\int\!\mathbf{n} \cdot (\mathbf{U} \cdot \nabla C) dS
\]

(2.1.12)

where \( \mathbf{n} \) is a unit vector normal to the boundaries of the box and \( S \) is surface area in m\(^2\). Note that we have made use of the incompressibility assumption to convert \( \mathbf{U} \cdot \nabla C \) in
(2.1.11) to $\nabla \cdot UC$ in (2.1.12). This surface integral also includes exchange at the air-sea and sediment-water interfaces, which usually occur through diffusive processes.

Finally, we must integrate the $SMS(C)$ term over the entire volume, since any change in this term at any location within the box will have an impact on the concentration. The final integral form of the tracer conservation equation is therefore:

$$\int \left( \frac{\partial C}{\partial t} - SMS(C) \right) dV = - \oint n \cdot (U \cdot C - D \cdot \nabla C) dS$$  

(2.1.13)

The problems that we encountered when we tried to simulate oxygen with the two box model of the ocean in Chapter 1 could have been foreseen by a more careful analysis of the surface transport integral in (2.1.13) and a better understanding of the ocean circulation. Consider, for example, the advection term in (2.1.13). This is the integral of the product of concentration times velocity normal to the surface of the box. One can think of this as a weighted mean of the surface concentration, where the weight is the component of the velocity that is perpendicular to the interface between the surface and deep boxes. Thus, in the two box model of Chapter 1, the surface concentration that is representative of waters sinking into the deep ocean is not the mean of the entire surface ocean, but rather a concentration that is strongly weighted towards the high latitude regions where surface water sinks into the deep ocean. We were able to improve our box model solution in Chapter 1 by separating the high latitude surface ocean from the low latitude surface ocean, but we could also have done this by using the velocity weighted integral of the surface concentration as given by (2.1.13). This simple example shows how important it is to understand the major features of oceanic transport, to which the remainder of this chapter is dedicated. We begin with a discussion of the circulation that is driven by the winds, which are the main source of energy for ocean circulation.

### 2.2 Wind Driven Circulation

Figure 2.2.1 summarizes the mean wind velocity at the surface of the ocean for January and July. The basic pattern of the winds, schematically illustrated in Figure 2.2.2 consists of the Trade Winds in the tropics, the Westerlies in mid-latitudes, and the Polar Easterlies in high latitudes. One might expect that surface water would be forced to move in the same direction as the wind. However, the water that is set directly into motion by the wind feels the effect of the Earth's rotation, and the resulting *Ekman transport*, which is confined to the top 10 to 100 m of the water column, is 90° to the right of the wind in the northern hemisphere and 90° to the left of it in the southern hemisphere (see Figure 2.2.3 for the ocean flow in the top 50 m).
One can gain some insight into how the Earth’s rotation affects motions in the atmosphere as well as the ocean by considering, for example, what happens to a projectile fired directly towards the north from somewhere at 30°N. The projectile carries with it not only the northward velocity imparted to it at the beginning of its trajectory, but also the eastward velocity of the Earth at 30°N. The eastward velocity of the earth is smaller at higher latitudes because of the decreasing diameter of the circles formed by slicing the earth at a given parallel of latitude. For example, the Equator has an eastward velocity of 455 m s$^{-1}$, and the 30°N and 45°N latitudes have eastward velocities of 402 and 326 m s$^{-1}$, respectively. Thus, with respect to a fixed position on the earth, the projectile will appear to accelerate towards the east. This simple thought experiment also shows how a projectile fired in the southern hemisphere would result in acceleration to the left of the direction in which the projectile is fired.

The Ekman transport causes water to converge in some regions of the ocean and diverge in others. This has two major impacts on ocean circulation. It causes downwelling in regions of convergent flow (Ekman pumping) and upwelling in regions of divergent flow (Ekman suction). Upwelling is particularly important for biology, because it is one of the primary mechanisms by which nutrients from the deep ocean are brought to the surface.

The other major impact of Ekman transport is that it piles up water in some regions and removes it from others, thus creating large horizontal gradients in sea surface height (Figure 2.2.4). The pressure gradients that result from these sea surface height differences might be expected to lead to a down slope flow, but the earth’s rotation intervenes, and the flow that results goes right angle to the slope rather than down it. Figure 2.2.5 shows that the resulting upper ocean flows for depths between 0 and 500 m have a gyre like structure that is particularly marked in the central part of the ocean basins in both hemispheres of the Atlantic and Pacific Oceans, as well as the southern hemisphere of the Indian Ocean. One particularly interesting and unexpected aspect of these gyres is that the western boundary flows are extremely narrow and intense. This westward intensification of the western boundary currents is also due to the earth’s rotation. Note also the large sea surface height slope away from Antarctica and resulting large circumpolar current.

In what follows, we briefly introduce the equations of motion and derive a few important analytical solutions that provide powerful insights into the major circulation patterns we have described. The emphasis is on providing a set of simple tools that are particularly useful for determining the nature of the ocean circulation that one might expect under a particular set of circumstances; and on those aspects of the circulation, such as upwelling driven by Ekman transport, that have the largest impact on biogeochemical processes.
EQUATIONS OF MOTION

The equations of motion are based on Newton’s Second Law $F = ma$, where $F$ is force (in Newtons), $m$ is mass in kg, and $a$ is acceleration in m s$^{-2}$. We use for $m$ the volume $V$ times the density $\rho$. For a full derivation of the equations of motion, the reader is referred to a standard physical oceanography textbook, such as Pond and Pickard [1983]. All of the solutions we consider here assume that the flow is hydrostatic (i.e., that the vertical pressure gradient is balanced by gravity) and incompressible (see Panel 2.1.3). We further assume that the forces driving the ocean circulation are in balance such that there is no net acceleration $a$, i.e., $F/ (\rho \cdot V) = 0$. However, it should be noted that the inclusion of the net acceleration term in the equations allows for a rich time dependent array of oceanic phenomena including waves with scales ranging from capillary to planetary, turbulent motions, formation of eddies, and convective overturning. Observations of some of these processes are discussed in Section 2.5.

The forces we consider here, all expressed as accelerations $F/ (\rho \cdot V)$, include the pressure gradient acceleration, the acceleration due to the earth’s rotation which is referred to as the Coriolis acceleration, the effect of the wind stress acting on the surface of the ocean, and the gravitational acceleration $g$:

\[
a_x = 0 = - \frac{1}{\rho} \frac{\partial p}{\partial x} + \frac{f}{\rho} v + \frac{1}{\rho} \frac{\partial \tau_x}{\partial z} \tag{2.2.1a}
\]

\[
a_y = 0 = - \frac{1}{\rho} \frac{\partial p}{\partial y} - \frac{f}{\rho} u + \frac{1}{\rho} \frac{\partial \tau_y}{\partial z} \tag{2.2.1b}
\]

\[
a_z = 0 = - \frac{1}{\rho} \frac{\partial p}{\partial z} - g \tag{2.2.1c}
\]

The first of these equations is the balance of acceleration terms in the $x$-direction. The second and third equations are the balances of the accelerations in the $y$ and $z$-directions, respectively. The vertical velocity $w$ does not enter into these simplified versions of the equations. We can solve for it by using the mass conservation equation for incompressible flow (equation (12) of Panel 2.1.3). And now we describe each of the terms:

1. The first term on the right hand side of all three equations is the acceleration due to the pressure gradient force. Pressure has units of Newtons m$^{-2}$. If pressure increases as $x$, $y$, or $z$ increases, the pressure gradient force will induce a flow in the opposite direction.
(2) The second term on the right hand side of (2.2.1a) and (2.2.1b) is the acceleration due to the Coriolis force. Note that the Coriolis acceleration in the x-direction (equation (2.2.1a)) depends on \( v \), the velocity in the y-direction. Conversely, the Coriolis acceleration in the y-direction (equation (2.2.1b)) depends on the x-direction velocity \( u \). In other words, the acceleration is always at right angles to the direction of the flow. In effect, this means that the only way the Coriolis force can balance a pressure gradient or any other force is if the velocity is at right angles to that other force. The Coriolis acceleration is derived by considering the motion in a rotating frame of reference relative to one fixed with respect to the stars. After neglecting minor terms, the Coriolis acceleration turns out to have the simple form given in the equations with a magnitude equal to the velocity times the Coriolis parameter

\[
f = 2 \cdot \Omega \sin \theta
\]

where \( f \) has units of \( s^{-1} \). Here \( \theta \) is the latitude (-\( \pi/2 \) radians at the South Pole, 0 radians at the Equator, and \( \pi/2 \) radians at the North Pole), and \( \Omega = 7.3 \times 10^{-5} \text{ s}^{-1} \) is the angular velocity of the earth in radians per second. The sign of \( f \) determines the direction of the Coriolis acceleration. In the northern hemisphere, where \( f \) is positive, the Coriolis acceleration is to the right of the flow. In the southern hemisphere, \( f \) is negative and the Coriolis acceleration is to the left of the flow. The Coriolis parameter, and thus the Coriolis acceleration, is greatest at the poles and decreases to 0 at the Equator.

(3) The third term on the right hand side of (2.2.1a) and (2.2.1b) is the acceleration due to the direct influence of the winds in the top 10 to 100 m of the water column. The force of the wind on the ocean is generally represented as a horizontal stress \( \tau_0 \) (Newtons m\(^{-2}\)) acting on a horizontal surface at the top of the ocean. One way of estimating the wind stress is as a function of the wind speed \( U \) as measured at a nominal height of 10 m on board a ship. In units of kg m\(^{-1}\) s\(^{-2}\), a typical empirical relationship would be \( \tau_0 = C_D \cdot \rho \cdot U^2 \), where the dimensionless drag coefficient \( C_D \) would vary from \( 10^{-3} \) to \( 4 \times 10^{-3} \) as the wind speed increased from 2 m s\(^{-1}\) to 50 m s\(^{-1}\). Satellite measurements of microwave backscatter from wavelets are also being used to map wind stress. The surface wind stress propagates vertically into the interior of the ocean through frictional coupling of deeper layers of water by the shallower layers set directly in motion by the wind. The force resulting from this stress is equal to the vertical gradient of the stress times the volume \( V \). Dividing this force by the mass \( \rho \cdot V \) gives the acceleration expression shown in (2.2.1a) and (2.2.1b).

©J. L. SARMIENTO AND N. GRUBER
September 14, 2004
EKMAN TRANSPORT

In the surface waters that are under the direct frictional influence of the wind, the pressure gradient forces are small relative to the Coriolis and wind stress accelerations, and can thus be neglected. The Ekman transport solution for horizontal motion in this surface layer thus results from the balance of the Coriolis acceleration with the acceleration due to the wind stress. From (2.2.1a) and (2.2.1.b) we have:

\[ 0 = f v + \frac{1}{\rho} \frac{\partial \tau^x}{\partial z} \]
\[ 0 = -f u + \frac{1}{\rho} \frac{\partial \tau^y}{\partial z} \]  

(2.2.3)

The direct frictional influence of the wind forcing decreases very rapidly with depth so that it is already negligible by about 10 to 100 m depth. The details of what happens within the zone of frictional influence are not important for the problems we will consider here, so we integrate (2.2.3) vertically from a depth \(-h\), where the stress due to the wind is 0, to the surface, where the stress due to the wind is \(\tau_0\). We first multiply (2.2.3) by the density. We then separately integrate the wind stress acceleration term:

\[ \int_{-h}^{0} \frac{\partial \tau^x}{\partial z} dz = \tau_0^x \] 
\[ \int_{-h}^{0} \frac{\partial \tau^y}{\partial z} dz = \tau_0^y \]  

(2.2.4) 

where \(h\) is typically of order 10 to 100 m, and define an Ekman mass transport \(M_{Ek}\) as the vertical integral of the density times the velocity to the depth of frictional influence

\[ M_{Ek}^x = \int_{-h}^{0}(\rho u)dz \]
\[ M_{Ek}^y = \int_{-h}^{0}(\rho v)dz \]  

(2.2.5)

Note that this mass transport has units of kg m\(^{-1}\) s\(^{-1}\). It is the mass transport of water that occurs every second across a 1 m wide section of the water column. The final expression we obtain for the Ekman transport is

\[ M_{Ek}^x = \frac{\tau_0^y}{f} \]
\[ M_{Ek}^y = -\frac{\tau_0^x}{f} \]  

(2.2.6)
From these equations we see that the vertically integrated Ekman transport is always 90° to the right of the wind stress in the northern hemisphere where the Coriolis parameter $f$ is positive, and 90° to the left of the wind stress in the southern hemisphere where $f$ is negative.

Mass conservation requires that any water that is displaced by the Ekman transport must be replaced either by lateral transport from elsewhere or by water from below in regions of divergence (Ekman suction). In regions of convergent Ekman transport, water must be pumped downwards (Ekman pumping). There are three basic mechanisms by which the Ekman transport can lead to upwelling and downwelling. The first is by winds that have a component parallel to the coastline. If the wind blows in such a way as to drive Ekman transport away from the coast, there will be upwelling of deep waters (left panel of Figure 2.2.6). Conversely, if water is driven towards the coast by Ekman transport, there will be downwelling (right panel of Figure 2.2.6).

Examination of the wind patterns in Figure 2.2.1 show that the western margins of continents typically have conditions that favor upwelling (e.g., off California and Central America, off the coast of Perú, and off northwest and southeast Africa). This is confirmed by the circulation patterns shown in Figure 2.2.3 and by the sea surface temperature pattern of Figure 2.2.7, which shows cold waters offshore off all the upwelling regions. These are all areas of extremely high biological productivity, as we shall see in Chapter 4.

The second mechanism by which Ekman transport can give rise to upwelling is the blowing of wind along the Equator. Here, the Trade Winds have a strong eastward component both north and south of the equator (Figure 2.2.1). Because of the change in sign of the Coriolis acceleration, these winds lead to northward Ekman transport north of the Equator and southward Ekman transport south of the Equator (Figures 2.2.3 and 2.2.8a). Mass balance requires upwelling along the Equator. Both the Atlantic and Pacific Oceans have a large amount of upwelling due to this phenomenon. The Indian Ocean situation is more complex because of the large seasonal changes in the winds associated with the monsoonal circulation of the Indian subcontinent.

The third mechanism by which Ekman transport can cause upwelling or downwelling is as a consequence of the large-scale pattern of the winds in the open ocean. Consider the schematic wind stress pattern of Figure 2.2.2. Between about 20°N and 50°N, in the region referred to as the subtropics, the winds shift from westward moving to eastward moving. In the southern part of this zone, the westward moving Trade Winds will drive an Ekman transport to the north, whereas the eastward moving Westerlies in the northern part of the subtropics will drive an Ekman transport to the south. The magnitude of the flow is largest at the edges of the subtropics where the wind stress is at a maximum. In the middle of the gyre, the horizontal flow goes to 0. There will therefore be a convergent flow and downwelling in the central part of the subtropics.
(Figures 2.2.3 and 2.2.8b). Applying the same considerations to other regions of the schematic wind fields of Figure 2.2.2 gives upwelling in the subpolar region that lies to the north of the subtropics.

One can obtain a solution for the vertical velocity at the base of the Ekman layer at a depth \(-h\) by vertically integrating the mass continuity equation (equation (12) of Panel 2.1.3) to the depth \(-h\), first assuming a steady state. This gives

\[
0 = \frac{\partial M_x}{\partial x} + \frac{\partial M_y}{\partial y} + \rho \cdot (w_{z=0} - w_{z=-h}) \quad (2.2.7)
\]

We assume that the vertical velocity is 0 at the surface of the ocean and that the mass transport terms in (2.2.7) are equal to the Ekman transport defined by (2.2.6). This gives the following solution for the velocity at the base of the Ekman Layer

\[
w_{z=-h} = \frac{1}{\rho} \left[ \frac{\partial}{\partial x} \left( \frac{\tau_0^y}{f} \right) - \frac{\partial}{\partial y} \left( \frac{\tau_0^x}{f} \right) \right] = \frac{1}{\rho} \text{curl} \left( \frac{\tau_0}{f} \right) \quad (2.2.8)
\]

where \text{curl} is defined as \(\partial( )/\partial x - \partial( )/\partial y\). Consider again just the east-west component of the wind stress for the 20°N to 50°N subtropical region depicted in Figure 2.2.2. Equation (2.2.8) gives us for the downwelling velocity

\[
w_{z=-h} = \frac{1}{\rho} \left[ \frac{\partial}{\partial x} \left( \frac{\tau_0^y}{f} \right) - \frac{\partial}{\partial y} \left( \frac{\tau_0^x}{f} \right) \right] = \frac{1}{\rho} \text{curl} \left( \frac{\tau_0}{f} \right)
\]

Applying (2.2.8) to other regions of the schematic wind fields of Figure 2.2.2 gives upwelling in the subpolar region as well as the tropics. The large-scale patterns of Ekman pumping and Ekman suction have a dramatic impact on many biogeochemical processes. The lower latitudes of the subtropics, for example, are generally almost devoid of surface nutrients because there is no upwelling (and little mixing) to supply them from below. Consequently, they have substantially less biological activity than upwelling regions. As already noted, the equatorial region and many coastal upwelling zones are rich in biological activity driven by the high supply of nutrients from below. The subpolar gyres are also rich in nutrients and biological activity, but other processes such as light supply and deep winter mixing play a role in these regions that will be discussed in Chapter 4.
GYRE CIRCULATION

As we have noted, the Ekman transport accumulates water in the sub-tropics and removes it from the sub-polar and equatorial regions (Figures 2.2.3 and 2.2.4). In addition to causing downwelling, the accumulation of water into an elevated mound within the sub-tropics causes a pressure gradient force directed down the slope of the mound. If the earth did not rotate, water would accelerate in the direction of this pressure gradient. However, because of the earth’s rotation and resulting Coriolis force, the water that is set into motion by this pressure gradient force gets deflected 90° to the right in the northern hemisphere. The only way that the Coriolis force can balance the pressure gradient force is if the water flow is at right angles to the slope of the mound. The Coriolis force leads to a clockwise circulation around the center of the high pressure mound in the subtropical gyre regions of the northern hemisphere, and a counterclockwise rotation around the corresponding subtropical mound in the southern hemisphere (Figure 2.2.5). Such a circulation around a high-pressure region is referred to as anticyclonic.

In the sub-polar regions, the juxtaposition of the Westerlies and the Polar Easterlies causes a divergent Ekman flow and the generation of a low pressure region in between. The resulting pressure gradient force would lead to acceleration toward the interior of this low pressure field, but the Coriolis force again deflects this flow 90° to the right in the northern hemisphere and to the left in the southern hemisphere. The net effect is a cyclonic circulation (opposite to anticyclonic) around the low pressure depression of the surface.

Over most of the gyres that we have discussed, the Coriolis force roughly balances the pressure gradient force induced by the wind. The equations formed by the balance of these two terms are known as the geostrophic equations. We shall return to a full discussion of the geostrophic equations per se in Section 2.3. If we add to the geostrophic equations the wind stress forcing at the top of the ocean in (2.2.2a) and (2.2.2b), and then vertically integrate the resulting equations from the ocean floor to the surface, we obtain a solution that is referred to as the Sverdrup transport (cf., Pond and Pickard [1983]). The Sverdrup transport solution is able to predict the broad meridional (i.e., north-south) currents on the eastern sides of the gyres but not the narrow intense currents on the western side of the gyres shown in Figure 2.2.5. The Gulf Stream off the coast of North America and the Kuroshio off the coast of Japan are two well-known examples of such energetic western boundary currents. The reason for the existence of these intense western boundary currents was not understood for a long time, until Stommel provided a very simple and elegant solution in 1948 (Stommel [1948]).

Stommel’s crucial insight was to understand the central role of angular momentum or “vorticity” conservation in determining the overall pattern of the gyre circulation. The pressure gradient, Coriolis, and wind stress forces together are able to conserve vorticity

©J. L. SARMIENTO AND N. GRUBER
September 14, 2004
with an equatorward flow of water in the subtropical gyre, but poleward movement is not possible without violating this constraint. The only way that poleward movement is possible is by removing vorticity from the flow. Stommel was able to demonstrate with a simple analytical solution of the equations of motion that frictional forces acting on narrow currents with high velocity shear provide a way of removing vorticity. A central feature of his analysis is the fundamental role played by the fact that the Coriolis parameter varies with latitude. If the Coriolis parameter is set equal to zero or to a constant, in other words, if it is not allowed to vary with latitude, Stommel’s analytical model gives a symmetric flow regime, with the poleward flow being as broad as the equatorward flow. Stommel gives an excellent discussion of why this is so in his book on the Gulf Stream (Stommel [1965]). It is important to note that there are alternative theories for the existence of narrow and intense western boundary currents, but all have in common the importance of meridional variations in the Coriolis force. The interested reader is referred to more advanced texts.

Figure 2.2.9 shows an idealized summary of the wind-driven circulation based on a model similar to Stommel’s but with a different expression for the friction term (lateral rather than bottom friction; Munk and Carrier [1950]). In the northern hemisphere, subpolar gyres are set up between the Westerlies and the Polar Easterlies. At mid-latitudes, anti-cyclonic subtropical gyres are dominant with broad interior circulation and strong western boundary currents. The equatorial circulation set up by the winds has a westward flowing Equatorial Current in both hemispheres and an Equatorial Counter Current at the Equator. A comparison of this idealized circulation pattern with the actual circulation patterns illustrated in Figure 2.2.5 shows that this theory provides valuable insights into the basic structure of the gyre circulation of the ocean.

2.3 Wind Driven Circulation in the Stratified Ocean

The discussion in Section 2.2 ignored variations in the density of the ocean. The stratification of the water column, which is due to variations in the ocean’s temperature and salinity, has a dramatic impact on the wind driven circulation. As we shall see, there is a strong tendency for stratification to trap the entire wind driven gyre transport in the surface ocean. There are two basic concepts that will help us understand why this is so and how the ocean is able to escape this constraint: (1) currents in the interior of the ocean are strongly constrained to occur along surfaces of constant density, and (2) the geostrophic equations need to be satisfied by the flow along the layers defined by the surfaces of constant density. In the following subsection, we discuss both of these in turn before applying them in the subsequent subsection to understand the gyre transport.
BASIC CONCEPTS

Ocean stratification

Density is a function of temperature $T$, salinity $S$ (defined in Table 1.2.1), and pressure obeying the Equation of State $\rho = f(T, S, p)$. Oceanographers always report density as $\sigma = (\rho - 1000)$, so that, for example, a density of $\rho = 1027$ kg m$^{-3}$ becomes $\sigma = 27$. Figure 2.3.1 shows how density at the surface of the ocean varies as a function of temperature and pressure. The full equation of state, given by Fofonoff [1985] is quite complex. A simple rule of thumb is that $\sigma$ increases by ~1 when temperature increases by 5°C and when salinity increases by 1. Because water is compressible, the density also increases as the pressure increases, at a rate of about 1 sigma unit per 200 db (about 200 m depth).

The density of a water parcel at the depth where it is located is referred to as the in situ density. We are commonly interested in identifying water parcels that would have the same density if they were at the same pressure. We thus find it convenient to define a potential density as the density a water parcel would have if moved adiabatically and while conserving salinity to a reference pressure, commonly 1 atm at the surface of the ocean for which we use the symbol $\sigma_\theta$. Surfaces of constant potential density are referred to as isopycnal surfaces or isopycnals. In addition to considering the flow along isopycnal surfaces, we will often find it useful to consider the distribution of deep ocean properties along such surfaces.

While useful, the potential density concept suffers from the fact that two water types that have the same density at one reference pressure, e.g. at a pressure of 4000 db, but a different temperature and salinity, will most likely have a different density if referenced to another depth, e.g., the surface. This is because the compressibility of seawater varies with both temperature and salinity. It is thus generally not clear exactly how to define an isopycnal surface along which a water parcel will actually move if a significant change in depth is involved. Reid and Lynn [1971] addressed this problem by using a reference pressure as close as possible to the depth range of the water mass that is being examined. For example, in examining deep ocean observations, they used a reference pressure of 4000 db (about 4000 m). The potential density referenced to 4000 db is referred to as $\sigma_4$. The difficulty of spanning a broad depth range in the analysis of a particular water type has led to the introduction of the idea of a neutral surface (McDougall [1987]), which in effect continually readjusts the reference depth for calculation of the potential density of adjacent locations so as to make it as close as possible to the actual depth of the water type being considered at the two locations.

Most of the ocean stratification is due to temperature, and in what follows we will often use potential temperature distributions to depict the main features of the oceanic
stratification. By analogy with potential density, the potential temperature, for which we use the symbol $\theta$, is defined as the temperature a water parcel would have if it were brought adiabatically to the surface. Equations for calculating it are given by Fofonoff [1977] and Fofonoff [1978]. Surfaces of constant potential temperature are referred to as isothermal surfaces.

Figure 2.3.2a gives a schematic illustration of the vertical profile of potential temperature at several locations. Typically, one finds that the deep ocean has relatively uniform temperatures below a depth of 1 or 2 km down to the floor at a depth of 5 or 6 km. Above this, the vertical temperature profile depends on the region of the ocean. Cold high latitude regions have temperatures similar to the deep ocean. The mid-latitudes and tropics have a steep temperature gradient beginning at a depth of ~1 km and extending all the way to a mixed layer of uniform temperature at the surface that is typically of order 50 to 100 meters thick. The mixed layer depth is determined by the direct frictional forcing of the wind, which only penetrates to depths of approximately 10 to 100 m, as well as vertical overturning caused by densification of surface waters by surface cooling, an increase of salinity due to evaporation, or lateral inflow of dense waters at the surface. The mixed layer depth varies with season (Figure 2.3.2b) and can achieve depths of several hundred meters in the high latitudes where surface cooling destabilizes the water column in wintertime. As previously noted, the region of steep temperature gradients is known as the thermocline. The corresponding density gradient region is referred to as the pycnocline. Gradient regions in salinity, which do not always correspond to those in the temperature, are referred to as haloclines. The thermocline is further divided into the main or permanent thermocline, which is below the reach of seasonal fluctuations at the surface of the ocean, and the seasonal thermocline nearer the surface, which becomes part of the surface mixed layer during the winter.

Figure 2.3.3b shows a section of observed potential temperature that begins on the left-hand-side with the North Atlantic, then goes south to the Southern Ocean, the ocean that rings the Antarctic, before turning back north into the Pacific. The contours are lines of constant potential temperature. The data are from the WOCE tracks shown in Figure 2.3.3a. The sections show that the main thermocline within the subtropical gyre has a basin-like convex downward structure confined to each hemisphere. The basin-like structure is particularly striking in the North Atlantic. Cold waters bound the subtropical gyre basins to both the north and south. On the poleward side, the cold waters rise all the way to the surface and the high latitude thermocline is very weak or essentially non-existent. In the Equatorial regions, the surface waters have very high temperatures, but the thermocline is much thinner, a few hundred meters, than in the subtropical regions to the north and south where the thermocline has a thickness in excess of a thousand meters. The vertical temperature gradient is much greater in the equatorial region than elsewhere. The places where isothermal surfaces encounter the surface of the ocean are referred to as outcrops. The outcropping of the isothermal surfaces (and thus also of the isopycnal
surfaces) allows for penetration of water from the surface into the interior of the thermocline, referred to as ventilation. The ventilation of the thermocline from the outcrop regions will play a central role in the discussion of the gyre circulation that is to follow.

The ocean circulation in the interior of the ocean is strongly favored to occur along isopycnal surfaces. This is because any water parcel that is forced away from such a surface to deeper waters of greater density or shallower waters of lower density will have a strong buoyancy restoring that will force the water back onto a layer with waters of the same density. The only way to escape this constraint is to alter the density of the water parcel. One way to alter the density of water is at the surface by adding or removing heat through solar forcing and latent and sensible heat transfer, or by changing the salinity by exchanging water with the atmosphere. River input and ice formation and melting can also change the surface salinity. Another way to alter the density is by mixing of the water parcel with water of a different density. Such exchange can be quite large over rough topography, for example (e.g., Polzin et al. [1997; Ledwell et al. [2000; Mauritzen et al. [2002; Garabato et al. [2004]). However, the amount of mixing in the interior of the ocean away from the surface and the sides and floor of the ocean is generally very small (e.g., Ledwell et al. [1993] and Ledwell et al. [1998]) because the input of energy for mixing is low. We therefore assume in what follows that ocean transport within the interior of the ocean away from the surface, sides and bottom, occurs largely along isopycnal surfaces.

Geostrophic Equations

As noted in Section 2.2, the steady state motion in the interior of the ocean can be characterized by a balance between the pressure gradient and Coriolis forces referred to as the geostrophic equations:

\[
\frac{1}{\rho} \frac{\partial p}{\partial x} = f v \tag{2.3.1}
\]

\[
\frac{1}{\rho} \frac{\partial p}{\partial y} = -f u \tag{2.3.2}
\]

We derive here an expression for how these equations constrain the flow along layers of thickness \( H \) defined by two isopycnal surfaces. An approximation to the above equations that simplifies our derivation is to assume that the density in the pressure gradient term can be represented by its mean, \( \rho_0 \). This is one of a set of assumptions that are collectively referred to as the Boussinesq Approximation. Multiplying through by this density and then differentiating (2.3.1) with respect to \( y \) and subtracting the differential of (2.3.2) with respect to \( x \) to eliminate the pressure gradient term gives:
Here the density has been taken out of the differentials based on the assumption that its variations are negligible. In the next step, we note that $f$ is only a function of $y$, so that the differential of $f$ with respect to $x$ is 0; and we define $\partial f/\partial y$ as $\beta$. This gives:

$$f\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \beta v = 0$$

(2.3.4)

We next substitute the incompressible mass continuity equation

$$-\frac{\partial w}{\partial z} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$$

(2.3.5)

into (2.3.4) to obtain the \textit{geostrophic vorticity equation}:

$$\beta v - f \frac{\partial w}{\partial z} = 0$$

(2.3.6)

We now proceed through a further series of manipulations to arrive at an expression of the geostrophic vorticity conservation in terms of the isopycnal thickness. We use the total derivative definition given by equation (9) of Panel 2.1.3, to note that $df/dt = v \partial f/\partial y = \beta v$ ($f$ is constant with time and with respect to the $x$ and $z$ coordinates). Substituting this into (2.3.6) gives:

$$\frac{df}{dt} - f \frac{\partial w}{\partial z} = 0$$

(2.3.7)

Next we consider how to interpret the $\partial w/\partial z$ term. If we follow a parcel of water flowing in a layer of thickness $H$ defined by two isopycnals, this term can be thought of as representing essentially the vertical compression or expansion of the water parcel as the layer defined by the two isopycnals thins or thickens. The change in thickness of the water parcel will be equal to the difference between vertical velocity at the top of the water parcel $\partial z_{top}/\partial t$ and that at the bottom of the parcel $\partial z_{bottom}/\partial t$ divided by the thickness of the layer $H$. Since $z_{top} - z_{bottom} = H$, we have that $\partial w/\partial z$ is equal to $(1/H)(dH/dt)$ (see Pedlosky [1987] for various ways of integrating (2.3.7) that give similar solutions). Substituting this into (2.3.7) and combining the two terms thus obtained gives:

$$\frac{df}{dt} - f \left(\frac{1}{H}\right)\frac{dH}{dt} = 0$$

September 14, 2004
\[ H \frac{d}{dt} \left( \frac{f}{H} \right) = 0 \]  

which requires that

\[ \frac{f}{H} = \text{constant} \]  

We refer to \( f/H \) as the potential vorticity or \( PV \).

A simple example of the application of (2.3.9) is the behavior of a water column flowing over topography where \( H \) is the depth of the ocean. As the water column is forced to thin, the water, in order to conserve potential vorticity, must flow towards regions of smaller \( f \), i.e., towards the equator. On the downstream side of a topographic feature the effect is reversed. This helps to explain important excursions in, for example, the Circumpolar Current around the Antarctic, which is observed to deflect towards the north as it flows over the mid-ocean ridges.

GYRE CIRCULATION WITH STRATIFICATION

According to the arguments we have made thus far, the flow is strongly favored to occur on density surfaces and to follow contours of constant \( PV \). Consider the idealized case of flat layers such as those depicted in Figure 2.3.4. Because \( H \) is constant, a line of constant \( PV \) must be along a line of constant \( f \), i.e., it must be east west. This is depicted in the lower part of the diagram. The lines of \( PV \) intersect the eastern boundary (we do not draw \( PV \) lines in the western boundary region where the force balance includes contributions from other than the geostrophic terms). Because the flow must be along lines of constant \( PV \), and because there can be no flow out of the boundary, we are forced to conclude that flat layers such as those depicted in the figure can have no transport. In such a case, the gyre transport would be trapped in the surface layer of the ocean. What does the real ocean look like?

**Insights from the Potential Vorticity Distribution**

In this and subsequent subsections, our discussion focuses primarily on the North Atlantic subtropical gyre, though the basic approach we use can be applied to the other subtropical gyres of the world as well. Figure 2.3.5 shows a plot of depth and \( PV \) contours on two isopycnal surfaces. These are based on density observations in the main thermocline of the North Atlantic. The thickness \( H \) is calculated as described in the figure caption. The first thing to note from the depths and outcrops of these two density surfaces is that both isopycnals outcrop at the surface of the ocean within the subtropical
gyre (compare the outcrop positions with the extent of the subtropical gyre depicted in Figure 2.2.5). The possibility thus exists of the gyre circulation being pumped down from the surface onto these density surfaces. We note next that there are a large number of $PV$ contours in the tropics that do in fact intersect the eastern boundary as in our idealized case of Figure 2.3.4. As we might infer from our previous discussion, this region defines a shadow zone of low circulation and slow exchange with the surface of the ocean at the outcrop of the isopycnal (Luyten et al. [1983]). Figure 2.3.6 shows one year long flow trajectories on the 26.5 density surface as calculated by an ocean general circulation model. The dot marks the end of the trajectory. This model simulation confirms that the shadow zone is indeed a region of slow circulation and poor ventilation from the outcrop of the density surface.

Both isopycnals in Figure 2.3.5 also show some $PV$ contours that emanate from the outcrop at the surface and then flow around the subtropical gyre. Horizontal currents, such as those that form part of the subtropical gyre circulation (cf., Figure 2.2.5), can flow in from the surface along these constant $PV$ lines (a process referred to as horizontal induction), as can water that is pumped down from the surface by Ekman convergence. The particle trajectories in Figure 2.3.6 illustrate this circulation very nicely. The region where inward flow from the surface is permitted is referred to as the ventilated region (Luyten et al. [1983]). Both layers also have what are referred to as pool regions of homogenous vorticity on the western side of the subtropical gyre. Rhines and Young [1982] postulated the existence of such pool regions, arguing that they would form by frictional coupling of deep isopycnal surfaces with the gyre transport on shallower surfaces, even if the friction is weak. The closed $PV$ contours of the pool region permit gyre transport that is not required to flow in or out of the outcrop of the density surface, as illustrated by the particle trajectories in Figure 2.3.6.

We thus see that the gyre transport escapes being trapped at the surface by entering the main thermocline along lines of constant $PV$ that emanate from the surface outcrops of the density surfaces; and by weak frictional coupling of deeper surfaces with shallower surfaces, which generates regions of homogeneous potential vorticity around which the gyre transport can occur. One of the most remarkable consequences of $PV$ conservation is the shallowing of the thermocline towards the equator noted in Figure 2.3.3b and which can be seen in the isopycnal depth contours of Figure 2.3.6. This can readily be understood as a precondition for the equatorward transport of the subtropical gyre circulation. The equatorward flow towards regions of lower Coriolis parameter $f$ can only occur if the thickness of the layers thins.

We end this subsection by noting that the thermocline in the subpolar gyre and equatorial regions is fundamentally different from the ventilated thermocline of the subtropical gyres. The subpolar and equatorial regions are places where the Ekman divergence leads to upwelling, and the interior gyre circulation, which carries less dense
low latitude waters towards the pole, is in the wrong direction for it to penetrate into the
denser waters of the thermocline. If this were all that were going on in these regions, the
upwelling would thin the thermocline, as is indeed observed to be the case. The diffusive
theories of the thermocline assume that downward mixing of heat from above
counterbalances this thinning tendency of the upwelling. The diffusive theories can be
 contrasted with the so-called adiabatic theories we have been discussing, in which it is
 assumed that there is no exchange of heat or salinity between layers once water leaves the
surface. The thermocline layer that is predicted by the diffusive theory is referred to as a
boundary layer. It has been proposed that this diffusive boundary layer is a continuous
blanket lying near the surface in the subpolar gyre and flowing deep beneath the so-called
ventilated thermocline of the subtropical gyre that we discussed above (Samelson and
Vallis [1997]). The deep boundary layer in the subtropical gyre manifests itself in Figure
2.3.3b as a region of sharper vertical temperature gradient at a depth of several hundred
meters underneath the mode water layer, defined in the North Atlantic by a thick layer of
relatively uniform temperature at 18°C.

Insights from Tracers

Among the most powerful tracers for studying thermocline ventilation are tritium
produced by nuclear bomb tests and its decay product helium-3, as well as
chlorofluorocarbons. We will discuss tritium and helium-3 observations here, as these
have been more thoroughly investigated. Tritium, the radioactive isotope of hydrogen,
has a half-life of 12.4 years and decays to helium-3 by emitting a weak beta particle. It is
produced in the upper atmosphere by cosmic ray spallation and fast neutron interactions
with \(^{14}\)N to form one tritium atom and \(^{12}\)C. The global tritium inventory maintained by
the balance between these formation processes and decay is about 3.6 kg. Nuclear bomb
tests of the late 1950’s and early 1960’s introduced about 550 kg of tritium into the
atmosphere, totally overwhelming the natural signal (Figure 2.3.7). The bomb produced
tritium washed out onto the land and into the ocean as water. Its total inventory has been
decreasing with time as it decays away. In what follows, we first show the distribution of
tritium in the main thermocline and then follow with a discussion of age maps determined
from the combined tritium/helium-3 distribution. The discussion centers on the North
Atlantic although much work has been done also in the North Pacific and will be
mentioned briefly at the end of this section.

Figure 2.3.8 shows the tritium distribution in 1972 as mapped on two density
surfaces within the main thermocline of the North Atlantic by Sarmiento et al. [1982].
The depths of these surfaces are shown in Figure 2.3.5. The outstanding feature is the
strong front centered at about 15°N that reflects the boundary between the ventilated and
pool regions of the gyre to the north, and the shadow zone of poor ventilation to the
south. This provides strong confirmation of the thermocline theories discussed earlier,
including the circulation inferred from the potential vorticity pattern of Figure 2.3.5, and
the particle trajectories of Figure 2.3.6. Within the ventilated and pool regions, we see that the highest values are near the outcrop. On the $\sigma = 26.8$ surface, the tritium drops by quite a large amount as one traces the westward path of ventilation region. Lower concentrations are found in the recirculation region in the northwestern corner of the gyre, which is ventilated less efficiently than the so-called ventilated region, presumably by exchange of tracer across the flow lines.

Sarmiento [1983] developed a simple box model of thermocline ventilation in order to estimate the exchange rate between the surface ocean and the interior (Figure 2.3.9). The mechanisms of ventilation are presumed to be Ekman pumping, horizontal inflow (which is referred to as induction), and lateral mixing. These are represented by a simple exchange term $1/\tau$. The equation solved was:

$$\frac{\partial [{}^3\text{H}]_{\text{interior}}}{\partial t} = \frac{1}{\tau} \left( [{}^3\text{H}]_{\text{surface}} - [{}^3\text{H}]_{\text{interior}} \right) - \lambda \cdot [{}^3\text{H}]_{\text{interior}}$$  \hspace{1cm} (2.3.10)

where $\lambda$ is the decay constant of tritium (yr$^{-1}$), equal to $\ln 2$ over the half life. The surface tritium concentration was taken from Dreisigacker and Roether [1978]. The model was run with various values of $\tau$ until the 1972 value of the interior tritium concentration was correctly predicted. The results are shown in Figure 2.3.10. Also shown in the Figure are some revised calculations of ventilation time scales by Doney and Jenkins [1988]. The main result of these calculations is that the thermocline ventilation time scales are of the order of one or two decades or less and that the ventilation rate exceeded the Ekman pumping by a factor of 5. Horizontal induction is the largest contributor to the ventilation by far (e.g., Huang [1990]; Speer and Tziperman [1992]; and Marshall et al. [1993]).

A powerful application of combined parent tritium/daughter helium-3 observations is the determination of an age since the water was last at the surface. Helium-3 equilibrates rapidly with the atmosphere by gas exchange. It builds-up by decay of the parent once the water is isolated from the surface. The isolation time can be estimated from:

$$\tau_{\text{THE}} = \frac{\ln \left( \frac{[{}^3\text{H}] + [{}^3\text{He}]}{[{}^3\text{H}]} \right)}{\lambda}$$  \hspace{1cm} (2.3.11)

Figure 2.3.11 shows age contours on four isopycnal surfaces in the North Atlantic (Jenkins and Wallace [1992]). The 15°N front is only hinted at due to the sparseness of observations in this region. However, the general patterns of gyre ventilation, and the increasing time scale with depth, are all evident in these figures. Jenkins [1998] has used the two tracers in combination to estimate the actual magnitude of the advection and diffusivity in the western side of the North Atlantic subtropical gyre.
In concluding this section, we show in Figure 2.3.12 a map of potential vorticity (cf. McPhaden and Zhang [2002]) and CFC-11 age (John Bullister, personal communication) from the North Pacific Ocean (see also the tritium distributions analyzed by Fine et al. [1987]). The CFC-11 age is determined by assuming that the concentration of this gas is saturated at the surface of the ocean, which makes it possible to determine the age of any water parcel simply by comparing its concentration directly with the temporal history of the surface concentration shown in Figure 2.3.7. Thiele and Sarmiento [1990] discusses the effect of mixing on this age determination, as well as tritium-helium-3 age estimates. These two tracers show the same patterns of the ventilated zone, shadow zone, and recirculation region of the main thermocline that we have seen in the Atlantic Ocean.

**Insights from the Thermal Wind Relationship**

In regions where the geostrophic balance obtains, one should be able, in principle, to calculate velocities from (2.3.1) and (2.3.2). However, the horizontal pressure gradients cannot be measured with sufficient accuracy to do this. We get around the difficulty of determining the horizontal pressure gradient from observations by recasting the geostrophic equations in terms of the horizontal density gradient (the so-called thermal wind relationship). We do this using the hydrostatic equation (2.2.1c), which gives the vertical pressure gradient in terms of the density:

\[-\rho g = \frac{\partial p}{\partial z}\]  (2.3.12)

Note that density variations in this equation cannot be neglected, even though it is still safe to ignore them in the geostrophic equations. Without density variations, the vertical pressure gradient in (2.3.12) would have to be constant everywhere, which is a poor assumption. We eliminate the pressure gradient terms from (2.3.1) and (2.3.2) by differentiating them with respect to \(z\) and subtracting the differential of (2.3.12) with respect to \(x\) and \(y\), respectively:

\[
\frac{\partial (\rho_0 f v)}{\partial z} + \frac{\partial (\rho g)}{\partial x} = \frac{\partial}{\partial z} \left( \frac{\partial p}{\partial x} \right) - \frac{\partial}{\partial x} \left( \frac{\partial p}{\partial z} \right) = 0
\]

\[
\frac{\partial (\rho_0 f u)}{\partial z} + \frac{\partial (\rho g)}{\partial y} = \frac{\partial}{\partial z} \left( \frac{\partial p}{\partial y} \right) - \frac{\partial}{\partial y} \left( \frac{\partial p}{\partial z} \right) = 0
\]  (2.3.13)

We now have:
\[
\frac{\partial}{\partial z} (\rho_0 f v) = -\frac{\partial}{\partial x} (\rho g) \\
\frac{\partial}{\partial z} (\rho_0 f u) = \frac{\partial}{\partial y} (\rho g)
\]  
(2.3.14)

The vertical gradient of \( f \) and the horizontal gradient of \( g \) are negligible, and the vertical gradient of \( \rho_0 \) is 0. Taking these terms out of the differentials gives the Thermal Wind Relationship:

\[
\frac{\partial v}{\partial z} = -\frac{g}{\rho_0 f} \frac{\partial \rho}{\partial x} \\
\frac{\partial u}{\partial z} = \frac{g}{\rho_0 f} \frac{\partial \rho}{\partial y}
\]  
(2.3.15)

The thermal wind equations give us the vertical gradient of the horizontal velocity as a function of the horizontal density gradient. In order to find the absolute velocity, we must integrate this equation vertically from some depth at which we know the velocity. This requirement of determining the velocity at some depth level is known as the reference level problem. The historical practice has been to pick a depth between oppositely flowing water masses where it seems reasonable to assume that the velocity should be close to 0. In recent years there have been attempts to combine the geostrophic constraint with mass conservation to estimate absolute velocities and transports of heat, salt and tracers (e.g., Roemmich [1980]; Wunsch [1984]; Rintoul and Wunsch [1991]; Martel and Wunsch [1993]; Wunsch [1994]; and Ganachaud [2003]). Here we apply the Thermal Wind Relationship qualitatively in a couple of locations of particular interest.

Figure 2.3.13 gives a section across the narrow Gulf Stream off the North American coast. For simplicity, we treat the section as though it were east-west, impinging upon a north-south coastline, although in reality the orientation of the coastline is rotated towards the east-west direction. The density and temperature both show a very sharp front with colder and denser waters near the coast and warmer and less dense waters in the interior. The horizontal gradient of density in the \( x \) direction across the front is negative (density decreases with increasing \( x \)). From (2.3.15) we see that this implies that the north-south velocity \( v \) must increase with increasing \( z \) (upwards). If we integrate this equation upward from some reference depth (e.g., 1500 meters) we see that the velocity in the frontal region is towards the north. The intense current defined by the strong horizontal gradient region is quite narrow, of order 100 kilometers. This narrow boundary current, which is the Gulf Stream, closes the interior Sverdrup Transport in Stommel’s simple model of the gyre (cf., Figures 2.2.5 and 2.2.9). Thus, we see that the large scale flow features defined by the simple wind driven theories of Section 2.2 manifest themselves very clearly in the density structure of the ocean.
Another region of considerable interest is the tropics. We see from Figure 2.2.9 that the large scale trend of flow on the equatorward side of the subtropical gyres is towards the west (the North and South Equatorial Currents; cf. Also Figure 2.2.5). However, Figure 2.2.5 also shows an eastward flow just north of the equator. How do these features manifest themselves in the density structure and geostrophic velocity? Figure 2.3.14 shows north-south sections of various properties across the Pacific Ocean. We use the temperature to obtain a measure of the density. This section shows low temperature (high density) waters at shallow depths at the Equator, and at about 8°N. The meridional density gradient is negative throughout most of the northern hemisphere (density decreases with increasing y) and positive in the southern hemisphere (density increases with increasing y). From (2.3.15) we see that the choice of a 0 velocity reference level at some depth below the thermocline implies a flow in the negative x direction in both hemispheres (recall that f changes sign from positive in the northern hemisphere to negative in the southern). This is as predicted by the Sverdrup Transport solution. Between 5°N and 8°N, the flow is reversed to form the North Equatorial Counter Current, a feature which is also predicted by the gyre circulation theories in analyses with realistic wind stresses. The only feature of the transport shown in the bottom panel of Figure 2.3.14 that is not predicted by the theories we have developed so far is the Undercurrent right at the Equator where the Coriolis force is 0. A simple pressure gradient force drives this current from the west towards the east.

2.4 Deep Ocean Circulation

There do not exist any generally accepted theories of the deep ocean circulation. Indeed, even the term thermohaline circulation which is often used to describe this flow is controversial because of its incorrect implication that the main source of energy for the deep ocean mass flux is buoyancy (heat and water fluxes) rather than mechanical energy provided by the wind and tides (cf. Munk and Wunsch [1998] and Wunsch [2002], who recommend using the term meridional overturning circulation or MOC instead). Buoyancy actually does play a role in the meridional overturning circulation, but mostly in determining where surface water sinks to the abyss. In this section, we will confine our attention primarily to a description of the main features of the meridional overturning circulation as inferred from observations and simulated by ocean general circulation models. In a final subsection, we will propose a synthesis.

OBSERVATIONS

We use the salinity and radiocarbon distributions to characterize the main features of the meridional overturning circulation and to estimate the time scale of deep ocean ventilation by this circulation. Figure 2.4.1 shows salinity along the two sections illustrated in Figure 2.3.3a, Figure 2.4.2 shows zonal mean radiocarbon in the Atlantic.
and Pacific Oceans produced from gridded data as described in the figure caption, and Figure 2.4.3 shows the age of radiocarbon at 3500 m determined by inverting the decay equation

\[ C = C_0 \cdot \exp(-\lambda t) \]  

(Matsumoto and Key [2004]). The ages are calculated with respect to the pre-industrial atmosphere (0‰). Radiocarbon has a decay constant of \( \lambda = 1.21 \times 10^{-4} \text{ y}^{-1} \). Thus, for example, a radiocarbon concentration of −100‰ implies an age of 871 years. Before discussing these figures, we need to say a little about the interpretation of the radiocarbon distributions and ages.

**Natural radiocarbon** produced by cosmic rays in the atmosphere has long been the tracer of choice for estimating the time scales of the deep ocean meridional overturning circulation. Panel 1.2.1 explains how natural radiocarbon is formed, how it enters the ocean, and how it is measured. The natural radiocarbon concentration has fluctuated over time due to variability in the cosmic ray flux (cf. Stuiver and Quay [1981]). However, on the time scale of the ocean circulation, these fluctuations are small relative to the contamination by the input of radiocarbon free fossil fuel CO\(_2\) that began with the start of the industrial revolution in the 19\(^{th}\) Century (called the **Suess Effect** after Suess [1955]); and to the production of radiocarbon by nuclear bomb tests in the 1950’s and 1960’s, which nearly doubled the atmospheric concentration by 1964 (see Figure 2.3.7).

The surface ocean radiocarbon concentrations before the Suess effect and nuclear bomb tests were about −40‰ in the low latitudes, (cf. Toggweiler et al. [1989]), dropping to about −67‰ in the North Atlantic and −140‰ in the surface waters of the Weddell Sea in the Southern Ocean (Broecker et al. [1998]). The surface North Atlantic and Weddell Sea concentrations correspond to ages of 570 years and 1250 years, respectively, and even −40‰ gives an age of 340 years. One would expect the concentrations of these surface waters to be close to atmospheric. The reason that the surface ocean does not equilibrate with the atmosphere is because the air-sea equilibration time for isotopic ratios of carbon is about 5 years for a 40 m thick layer of the ocean (see Chapter 3), and water does not spend that much time at the surface. In the Weddell Sea, this exposure time at the surface is so short that essentially no radiocarbon is able to enter from the atmosphere and the surface concentration essentially reflects the deep ocean concentration in that region (Weiss et al. [1979]).

The observed radiocarbon concentrations in the main thermocline of the Atlantic and Pacific Oceans shown in Figure 2.4.2 are clearly well above the surface values, indicating that the nuclear bomb test contamination has penetrated throughout the main thermocline to a depth of about 1000 m. We can infer from the distribution of CFC-11
and tritium in the North Atlantic that there is also contamination of bomb radiocarbon at the bottom of the ocean west of the mid-ocean ridge from Iceland all the way down to the equator and possibly as far as 5°S to 10°S (Smethie et al. [2000]). This means that the age estimates for this region of the ocean shown in Figure 2.4.3 are too young. However, the rest of the deep ocean at 3500 m has little if any such contamination.

The salinity distribution (Figure 2.4.1) shows a layered structure with distinct minima and maxima at different depths. Since there are no significant internal sources or sinks of salinity, these minima and maxima can only originate at the surface of the ocean. Their pattern thus implies the existence of a large-scale meridional overturning circulation involving deep water formation (sinking from the surface into the deep ocean) in the North Atlantic and around the Antarctic, and penetration of these water types into the Atlantic, Indian, and Pacific Oceans. We use the salinity and radiocarbon distributions to identify the deep water masses that form by this sinking, beginning first in the North Atlantic and then progressing through the Southern Ocean into the Pacific.

The North Atlantic has a thick layer of relatively salty North Atlantic Deep Water (NADW; see arrows in Figure 2.4.1a). This water mass originates in the deep basins to the north of Iceland, and also includes shallower components that form in the Labrador and Mediterranean Seas, the latter having particularly high salinity (see filled in circles in Figure 2.4.3 for formation locations). The radiocarbon content of NADW is the highest in the deep ocean (Figure 2.4.2a), giving the youngest ages (Figure 2.4.3). This would be true even without the bomb contaminant. The age map in Figure 2.4.3 shows that the youngest ages of the NADW are along the western boundary of the basin, suggesting that it flows out of the North Atlantic predominantly as a western boundary current.

At the bottom of the Atlantic Ocean, mostly south of the Equator, is found the relatively fresh (i.e., less salty) Antarctic Bottom Water (AABW; see Figure 2.4.1). The predominant source of the AABW within the Atlantic is the Weddell Sea (see filled in circle in Figure 2.4.3) though there is some input from other formation regions around the Antarctic, as well as some entrainment of the so-called Circumpolar Deep Water that fills the deep Southern Ocean. As the AABW penetrates to the north, its influence gradually erodes away due to mixing with salty NADW above. The AABW has lower radiocarbon than the NADW (Figure 2.4.2a). The 3500 m age map in Figure 2.4.3, which lies at the top of the AABW, gives an age of 1400 years near the Antarctic continent. We would like to be able to use the radiocarbon time scales in Figure 2.4.3 to estimate the ventilation rate of the Atlantic Ocean by the NADW and AABW. However, to do so we need to isolate that part of the aging that is due to actual decay of radiocarbon from that which is due to mixing of old AABW with young NADW and contamination by bomb radiocarbon in the North Atlantic. The most recent attempt to do this, which we describe here, is that of Broecker et al. [1998].
Consider a water parcel somewhere in the deep Atlantic that consists of some fraction $f_n$ of NADW, and a corresponding fraction $(1-f_n)$ of AABW. The measured radiocarbon content of this water parcel will equal $f_n$ times the initial radiocarbon concentration of the NADW at the time it left the surface, plus $(1-f_n)$ times the initial radiocarbon concentration of AABW, minus radiocarbon that has been lost to decay (the part that interests us), plus the radiocarbon contamination effects. Ignoring the contamination effects for now, we see that we can determine the decay if we know $f_n$ and the initial radiocarbon concentrations of NADW and AABW. Broecker et al. [1998] estimate the initial radiocarbon concentrations of NADW and AABW ($-67\%\epsilon$ and $-140\%\epsilon$, respectively) by using observations as in Broecker et al. [1991]. They remove the bomb contamination from NADW by plotting radiocarbon versus tritium (also produced both by cosmic rays and by bomb tests) and extrapolating the relationship to the presumed pre-bomb tritium concentration of near 0. The AABW is assumed to be free of bomb contamination.

Determining $f_n$ requires a conservative tracer that is clearly different in the NADW than in AABW. Salinity is conservative, but the separation into NADW and AABW is ambiguous because the four water masses that combine to form the NADW have different salinities. Broecker et al. [1991] proposed a new tracer, $PO_4^*$, that has nearly uniform values within the AABW and within the four components of the NADW ($1.95 \pm 0.07$ and $0.73 \pm 0.07 \mu$mol kg$^{-1}$, respectively; Broecker et al. [1998]). $PO_4^*$ consists of the weighted sum of two tracers used by organisms during photosynthesis and respiration, namely phosphate and oxygen. As noted in Chapter 1, observations suggest that deep ocean organisms remineralize organic matter in a fixed stoichiometric ratio of about 1 phosphate molecule released for every 170 oxygen molecules taken up. In other words, for every phosphate molecule added to the water by remineralization, 170 oxygen molecules are removed. Thus, if we were to add the phosphate concentration to the oxygen concentration divided by 170, the resulting tracer would be conserved except at the ocean surface where gas exchange can affect the $O_2$ concentration. We discuss the underlying concepts in more detail in Chapter 5. The specific definition for $PO_4^*$ proposed by Broecker et al. [1991] is

$$[PO_4^*] = [PO_4^{3-}] + \frac{[O_2]}{175} - 1.95 \mu$mol kg$^{-1}$, \hspace{1cm} (2.4.2)$$

which has a slightly higher stoichiometric ratio of 175. The constant of 1.95 $\mu$mol kg$^{-1}$ in (2.4.2) is arbitrary. Given the aforementioned NADW and AABW concentrations, the fraction of northern component water in a water sample can then be calculated from

$$f_n = \frac{1.95 - [PO_4^*]}{1.95 - 0.73} \hspace{1cm} (2.4.3)$$
Figure 2.4.4 shows the $PO_4^*$ distribution at 3000 m in the world ocean. The separation between NADW and AABW in the Atlantic is clear. Interestingly, the remainder of the deep waters in the Indian and Pacific Oceans have a nearly uniform $PO_4^*$ with a mean of $\sim$1.4 $\mu$mol kg$^{-1}$. Plugging this into (2.4.3) gives that the Indian and Pacific Oceans are about 45% NADW and 55% AABW. Globally, including the Atlantic, this means that about half the deep water is NADW and half AABW. This implies that the rate of AABW formation must be about the same as the rate of formation of NADW (Broecker et al. [1998]). So what is the rate of formation of NADW?

Once we know the fraction of northern and southern component waters from (2.4.3), we can estimate the change in radiocarbon concentration due to decay from:

$$\Delta \left( \Delta ^{14} C \right) = \Delta ^{14} C_{\text{initial}} - \Delta ^{14} C_{\text{observed}}$$

$$\Delta ^{14} C_{\text{initial}} = f_n (0.933) + (1 - f_n)(0.860) \quad (2.4.4)$$

Here 0.933 is $R^*$ for the NADW component and 0.860 is $R^*$ for the AABW component ($R^*$ is defined by equation (2) of panel 1.2.1). Figure 2.4.5 shows the radiocarbon deficiency due to decay calculated from (2.4.3) and (2.4.4) overlaid with age contours. Recall that this radiocarbon deficiency and therefore the age are with respect to the initial radiocarbon of the NADW and AABW. They are thus a measure of the time scale of ventilation of the deep ocean with respect to the surface deep water formation regions. The analysis in Figure 2.4.5 uses the vertically averaged data over water depths greater than 2000 m. The map shows very young water along the western boundaries, consistent with Figure 2.4.3 and our earlier inference that the southward flow of NADW is confined primarily to western boundary currents. The observed distribution of chlorofluorocarbon and tritium provide strong confirmation of this western boundary undercurrent in the North Atlantic (Jenkins and Rhines [1980]; Weiss et al. [1985]; Olsen et al. [1986]; Pickart et al. [1989]; Rhein et al. [1998]; and Smethie et al. [2000]).

The average radiocarbon deficit obtained from the data shown in Figure 2.4.5 is $18\%,e$ corresponding to a residence time of 161 years. After correcting for variations in the initial conditions due to the cosmic ray flux variability, Broecker et al. [1998] estimate the residence time as 180 years. For the deep Atlantic below 2000 m, this gives a deep water input of 22 Sv (recall 1 Sv = $10^6$ m$^3$ s$^{-1}$). Other studies suggest about 4 to 7 Sv of this is from AABW (cf. Broecker et al. [1998]), giving an NADW input of 15 to 18 Sv, with a corresponding input of AABW based on the global $PO_4^*$ analysis of about 15 Sv. If we take the volume of $8.2 \times 10^{17}$ m$^3$ for the ocean basins below 2000 m, this gives a ventilation time scale of $\sim$870 years.
Returning to our analysis of the salinity distribution, we see that northward flowing Antarctic Intermediate Water (AAIW) lies above the NADW (Figure 2.4.1). AAIW is relatively fresh and has quite high radiocarbon (Figure 2.4.2a). The AAIW originates in open ocean areas of the Southern Ocean. The upper part of the AAIW is actually a separate water mass identified as the Subantarctic Mode Water or SAMW, also flowing towards the north (cf. Chapter 7, Figures 7.3.2 to 7.3.4). Above the SAMW lies the main thermocline.

The circumpolar region, or Southern Ocean, that rings the Antarctic is characterized by a relatively uniform deep water mass called the Circumpolar Deep Water CDW. This water mass is a crossroads of the deep ocean that blends together NADW, and corresponding deep water masses from the Indian and Pacific Oceans, as well as including deep water that forms around the Antarctic continent. It is remarkably uniform in the vertical, and the radiocarbon distributions shown in Figure 2.4.2 suggest that this water is drawn towards the surface by upwelling, with a downwards return flow towards the south that becomes AABW, and a northward flow that becomes AAIW as well as SAMW.

The Pacific Ocean has an AABW tongue along the bottom that erodes as the water penetrates to the north (Figure 2.4.1). This AABW is often referred to as CDW to distinguish it from the slightly different AABW that flows into the Atlantic. The Pacific AABW has a strong component of salty NADW that it picks up around the Antarctic, making it the saltiest water in the deep Pacific (cf., Reid and Lynn [1971]). Note from the salinity section in Figure 2.4.1 that there is no evidence of deep water formation in the North Pacific. The radiocarbon observations shown in Figure 2.4.2b suggest instead that the bottom water upwells and returns to the south at intermediate depths. This low radiocarbon water mass is known as the North Pacific Deep Water (NPDW; see Figure 2.4.1). Above the NPDW in the South Pacific can be seen the fresher signature of AAIW penetrating to the north (Figure 2.4.1). A corresponding intermediate water mass formed in the North Pacific is referred to as the North Pacific Intermediate Water (NPIW).

The Indian Ocean, which is not shown in Figure 2.4.1, looks similar to the Pacific except that a modest amount of salty deep water sinks down from the Red Sea and forms a water mass referred to as the Red Sea Deep Water.

We conclude our discussion of observations with a brief summary of an important box model analysis of the deep radiocarbon distribution by Stuiver et al. [1983]. Their box model is illustrated in Figure 2.4.6. Their estimate of NADW formation is specified a priori, so we will not discuss it any further. Their flow diagram for the Indian and Pacific Oceans has water entering from the circumpolar region and upwelling across 1500 m. Our analysis of observations is inconsistent with this depiction of the flow. It appears instead that most of the outflow is in the deep waters well below 1500 m.
However, the view represented by this diagram does correspond to the way that oceanographers used to view the deep ocean circulation, and it is interesting to examine its implications.

In order to determine the magnitude of the inflow from the Southern Ocean into the Indian and Pacific Oceans in the Stuiver et al. [1983] box model we need to know the radiocarbon content of the inflow and outflow water. The radiocarbon content of the inflowing circumpolar water is $-158\%e$ corresponding to an $R^*$ of 0.842 (see definition in Panel 1.2.1, equation (2)). The water at 1500 m has a radiocarbon concentration of $-181\%e$ in the Indian Ocean, and $-207\%e$ in the Pacific, corresponding to $R^*$s of 0.819 and 0.793, respectively. Decay is equal to the decay constant times the deep ocean volume and the mean deep ocean concentration of $-184\%e$ in the Indian Ocean, and $-217\%e$ in the Pacific, corresponding to $R^*$s of 0.816 and 0.783, respectively. A solution is obtained from the radiocarbon balance in combination with the mass balance constraint that inflows from the circumpolar region $I$ balance the vertical upwelling $W$:

\[
\begin{align*}
\text{Indian Ocean:} & \quad I_i = W_i \\
I_i (0.842) = W_i (0.819) + V_i \lambda (0.816) \\
\text{Pacific Ocean:} & \quad I_p = W_p \\
I_p (0.842) = W_p (0.793) + V_p \lambda (0.783)
\end{align*}
\]

(2.4.2)

The volume of the Indian Ocean north of 50°S and below 1000 m is $1.6 \times 10^{17}$ m$^3$ and that of the Pacific is $4.1 \times 10^{17}$ m$^3$. Solution of these equations gives 20 Sv for the Indian Ocean inflow and 25 Sv for the Pacific. The ventilation time scales and 1500 m upwelling rates corresponding to these inflows are given in Table 2.4.1. We discuss these calculations further in the following two subsections.

MODELS

Given an appropriate specification of boundary conditions at the surface and bottom of the ocean (i.e., wind forcing, heat and water fluxes, and friction) it is possible in principle to solve the three equations of motion, the equation of mass continuity, the Equation of State relating $\rho$ to $T$ and $S$, and the conservation equations for heat and salt. This gives a total of seven equations, which match the number of unknowns: $u$, $v$, $w$, $p$, $\rho$, $T$, and $S$. However, there is no analytical solution for the full equations under realistic conditions. The solution is therefore approximated by breaking the ocean down into finite size boxes, doing a volume integral around each box as we did in equation (2.1.13), and then expressing the resulting integrals in finite difference form and solving them numerically on a computer. The models that result from this are called Ocean General Circulation Models or OGCM’s.
Any ocean circulation processes that are smaller than the resolution of the boxes in an OGCM have to be parameterized, usually as eddy diffusivity (see Panel 2.1.2). We would thus prefer to use as large a number of boxes as is necessary to resolve major features of the ocean circulation that we believe to be important. However, as illustrated in Figure 2.4.7, present computational resources are not up to the task, particularly for global circulation models such as those we discuss in this section, which correspond to the area denoted “T” in Figure 2.4.7. In particular, these models are unable to resolve mesoscale eddies (shaded area 1 of the figure) such as are discussed in the following section. In the next section we show some results from a model that does have sufficient resolution to predict such eddies, but that is able to solve for the circulation in only a small portion of the Pacific Ocean off California (see Figure 2.5.1). Another major problem with models is that the boundary conditions of wind stress, heat, and water fluxes, are difficult to determine. Taking all these problems together, we must conclude that the numerical solutions of the flow that are obtained are all approximations to the real flow in the ocean. Nevertheless, this powerful approach is being used with increasing frequency to study ocean circulation, supplemented in studies such as the ECCO project (cf. Stammer et al. [2002]), by increasingly sophisticated assimilation of observations (see Figures 2.2.3 and 2.2.5).

The value of tracers such as radiocarbon and chlorofluorocarbon in helping to constrain the meridional overturning circulation predicted by OGCMs is illustrated by the diagrams shown in Figure 2.4.8 from Matsumoto et al. [2004], which compare 19 different models (indicated by small diamonds) with each other and with observations of these tracers (indicated by dots and crosses giving estimates of the uncertainty). The abscissa of all these diagrams is the radiocarbon content of the in Circumpolar Deep Water CDW. Some of the conclusions that can be drawn from these diagrams include (cf. Matsumoto et al. [2004]):

(1) The radiocarbon in the CDW has a wide range in the models, with most of them being well outside the $2\sigma$ uncertainty shown in the figure. Only four models are within the uncertainty. The data are thus clearly indicating that most models are unable to simulate the radiocarbon content of the CDW correctly.

(2) The CDW radiocarbon correlates linearly with an almost 1:1 slope with the NPDW radiocarbon content (see sloping lines in Figure 2.4.8a), but is independent of the NADW radiocarbon content. Matsumoto et al. [2004] conclude from this that most models have roughly the same circulation time scale between the Southern Ocean and North Pacific. The wide range of NPDW radiocarbon contents between the models is due primarily to the differing ability of the models to simulate the rate at which fresh radiocarbon is fed into the Southern Ocean either by gas exchange from above or laterally by input of young NADW.
(3) The CFC-11 inventories in the Southern Ocean have basically the same pattern as the CDW radiocarbon inventory on a model-by-model basis. Since CFC-11 has only been in the ocean for a few decades (see Figure 2.3.7) and thus mostly reflects upper ocean processes, the CFC-11 results of Figure 2.4.8c imply that the models have a wide range of surface to deep ventilation rates in the Southern Ocean. This is most likely the primary cause of the wide range in CDW radiocarbon simulations.

In what follows, we present a few results obtained by the model identified by 19 in Figure 2.4.8. The general picture that emerges from this particular model is broadly consistent with a range of other studies (e.g. England [1993], Maier-Reimer et al. [1993], Hirst et al. [1996], and Hirst and McDougall [1998]). This simulation was carried out at Princeton/GFDL based on the code originally developed by Bryan [1969] as modified by Pacanowski and Griffies [1999]. The model has a nominal resolution of 4° in the horizontal and 24 levels in the vertical ranging in thickness from 25 to 450 m, with highest resolution in the upper part of the water column. The particular version of the model we discuss here is identified as P2A in Gnanadesikan et al. [submitted], where it is described. The small scale processes in the conservation equation that are not resolved are represented by eddy diffusion terms with diffusivities of 0.15 cm$^2$ s$^{-1}$ in the vertical increasing to 1.3 cm$^2$ s$^{-1}$ at depth, and a lateral along isopycnal diffusivity of 100 m$^2$ s$^{-1}$.

The boundary conditions used to force the model at the ocean surface include the momentum flux, equal to wind stress determined from observations; and heat and water fluxes specified also according to observations. In addition, the surface temperature and salinity are forced towards observations (which is equivalent to an additional heat or water flux). No momentum, heat or salinity flux is permitted across the ocean bottom. The lateral walls have a no slip boundary condition (i.e., 0 velocity) and no heat or salinity flux.

Figure 2.4.9a shows the global meridional overturning in the P2A model and Figure 2.4.9b and c show a breakdown of this into the Atlantic and Indo-Pacific Oceans. The contour interval is 2.5 Sv, i.e., the amount of water flowing between two contour lines is equal to 2.5 Sv, and the total flow between any two contour lines is equal to the difference between them. For example, the flow between a contour labeled 10.0 and one labeled –2.5 is 12.5 Sv. The motion is clockwise around maxima in the stream function, and counter-clockwise around minima in the stream function. The primary features of the meridional overturning circulation outside the Southern Ocean are:

(1) A massive amount of upwelling in the upper 100 to 200 m at the equator. This upwelling is fed almost entirely from the upper few hundred meters of the subtropical gyre thermocline. The equatorial upwelling feature is consistent with what we inferred from our analysis of the Ekman transport.
(2) The circulation in the deep main thermocline outside the Southern Ocean is characterized by a northward flow (Figure 2.4.9a) that is confined primarily to the Atlantic Ocean (Figure 2.4.9b), where it contributes to the formation of NADW. We have identified this as AAIW on the diagrams, as its behavior in the model, including the fact that it feeds into NADW formation, corresponds to what we believe the AAIW does in the real ocean (e.g., Sloyan and Rintoul [2001]).

(3) The southward flow of deep water at intermediate depths of 1000 to about 3500 m occurs in all the ocean basins (Figure 2.4.9). It has a contribution of about 20 Sv from NADW (Figure 2.4.9b), with 5 to 10 Sv of additional flow due to NPDW and the corresponding Indian Ocean Deep Water (Figure 2.4.9c). The Indo-Pacific deep waters are formed primarily by upwelling of AABW in the northern part of these basins, whereas the Atlantic deep waters are formed primarily by sinking from above of waters from the main thermocline.

(4) A northward flow of AABW below about 3000 to 3500 m in all the ocean basins. The model thus succeeds in qualitatively reproducing the main features of the deep ocean circulation that we inferred from observations in the previous subsection. It also gives about the same NADW formation rate as estimated by Broecker et al. [1998]. However, the bottom water inflow of ~8 Sv in the Indo-Pacific is much smaller than the 45 Sv obtained by Stuiver et al. [1983] (Table 2.4.1), despite the fact that the OGCM radiocarbon simulation agrees quite well with the observations (see Figure 2.4.8a). The difference between the Stuiver et al. [1983] box model analysis and the OGCM is due in part to the fact that in the OGCM, most of the outflow of deep water from the Indian and Pacific Oceans occurs horizontally below 1500 m rather than vertically across 1500 m as assumed, most likely incorrectly, in the box model. Furthermore, inflows at one depth level are often balanced by outflows at another location but at the same depth level so that they do not show up in the zonally integrated meridional overturning. In addition, lateral mixing, which is not reflected in the stream function but is parameterized as advection in the box model, is an important contributor to the transport of radiocarbon out of the circumpolar region into the ocean basins (cf., Gnanadesikan et al. [submitted]).

We conclude our discussion of the meridional circulation in the model with a brief description of the large overturning cell centered at about 50°S in the Southern Ocean, which is referred to as the Deacon Cell (cf. Speer et al. [2000]), also commenting on the smaller and deeper cell centered at about 60°S to 55°S. An observationally based analysis of the circulation in this region is presented in Chapter 7 (cf., Figures 7.3.2 to 7.3.4) and by Speer et al. [2000]. Many aspects of the circulation shown in Figure 2.4.9a are consistent with observational analyses, such as the large northward Ekman transport at the surface, which the model gives as about 40 Sv centered at 50°S to 45°S, whereas at.

©J. L. Sarmiento and N. Gruber
September 14, 2004
least one observational analysis suggest it is more like 25 Sv (Speer et al. [2000]); and
the upwelling of deep water toward the surface between about 65°S and 50°S, which
contains some NADW (see Figure 2.4.1a) but consists mostly of shallower Upper
Circumpolar Deep Water (UCDW). Also consistent with our understanding of the deep
ocean circulation is that some of the northward Ekman transport subsequently sinks and
flows to the north as part of the intermediate water (10 Sv in the model). The downward
limb of the Deacon Cell of about 20 Sv that occurs at ~40°S is primarily an artifact of the
way that the stream function is calculated along surfaces of constant depth rather than
along surfaces of constant density (Döös and Webb [1994] and Hirst et al. [1996])

The deep overturning cell of 10 Sv between 60°S and 50°S in Figure 2.4.9a likely
corresponds to the feature inferred from observations of downward sinking of water
along the Antarctic continent to form AABW which subsequently flows to the north (cf.
Figure 2.4.1a). However, the meridional circulation in the model shows this cell as being
isolated rather than connected to the AABW.

The difficulties of model P2A and most models in simulating the Southern Ocean
circulation correctly is due to many different problems that are difficult to resolve,
including the various ways that eddy mixing is parameterized in models and the difficulty
in correctly representing buoyancy forcing at the surface of the ocean (i.e., heat and water
fluxes).

SUMMARY OF DEEP OCEAN CIRCULATION

One of the most influential attempts to synthesize the deep ocean meridional
overturning circulation views it as a “great ocean conveyor” starting with sinking of
NADW, which then travels to the circumpolar region, where it is briefly brought back to
the surface and re-cooled, before continuing on its way into the deep Indian and Pacific
Oceans, there to upwell all the way to the thermocline and surface ocean. The great
ocean conveyor is closed by the return of Pacific Ocean surface water through the
Indonesian Straits where it joins Indian Ocean surface water to flow around South Africa
(cf. Broecker [1991]). This view grows in part out of the remarkable “tour-de-force”
model of Stommel [1958], which, was concerned primarily with deep water formation
and the flow paths that deep water takes after it leaves the deep water formation regions,
but had also to deal with the problem of how to get rid of old deep water. Stommel
[1958] imagined that the removal of old deep water occurred by upwelling which was
due to heating of deep water by vertical mixing, and that this deep ocean upwelling
joined up with the thermocline upwelling previously discussed by Robinson and Stommel
[1959]). More elaborate views of the meridional overturning circulation, exemplified by
Schmitz [1995], include the various deep and intermediate water types we have identified
above and confine the deep upwelling primarily to the North Pacific. These depictions
tend to focus one’s attention on the formation of NADW as the starting point for the
meridional overturning circulation, and the importance of vertical upwelling to close the thermohaline circulation.

An alternative view of the meridional overturning circulation that has begun to emerge in the last decade focuses primarily on the closure of the circulation (that is to say, the removal of old deep water from the deep ocean) as the starting point for understanding the deep circulation. The main organizing principle is the removal of deep water through wind-driven upwelling that occurs in the Southern Ocean (see Figures 2.4.10 and 2.4.11 based on Toggweiler and Samuels [1993b] and Gnanadesikan and Hallberg [2002], respectively). Southern Ocean upwelling is driven by the strong northward Ekman transport at the surface of the ocean due to the Westerlies (cf. Speer et al. [2000]). A significant fraction of the upwelling gets drawn from great depths for reasons that have to do with the dynamics of the Antarctic Circumpolar Current (ACC) in this unique area of the ocean where there is a completely open zonal path around the world (cf., Toggweiler and Samuels [1993a], Toggweiler and Samuels [1993b], Döös and Webb [1994], Gille [1997], Gnanadesikan [1999], Hallberg and Gnanadesikan [2001], and Lee and Coward [2003]). As a consequence of this deep upwelling, the main feature of the deep ocean circulation becomes a mid-depth flow to the south, and very little vertical mixing is required since the winds in the Southern Ocean do most of the work required to bring deep water to the surface. Buoyancy forces help determine where deep-water formation occurs, but do not drive the deep circulation as they do in the old theory.

The specific consequences that follow from the fact that upwelling in the Southern Ocean appears to be the primary return path for deep waters to the surface of the ocean include the following (cf. Figure 2.4.10):

1. The observational and model analyses suggest that there is a massive deep water flow to the south in all three ocean basins. Presumably, this is due to the drawing up of deep water by the Ekman upwelling in the Southern Ocean.

2. The deep water inflow to the Southern Ocean is fed by two circulation loops. The first is a shallow loop that exists mainly in the Atlantic and involves northward flow through the main thermocline (with input also from the Indian and Pacific Oceans) and formation of NADW in the North Atlantic. The second is a deep loop that involves AABW formation in the Southern Ocean, and northward flow of AABW in all three basins which then upwells and returns as deep water (NADW, NPDW, and RSDW in the Indian Ocean; cf. Schmitz [1995]).

3. Because essentially all of the upwelling of deep water to the surface of the ocean appears to occur in the Ekman divergence region of the Southern Ocean, the meridional overturning circulation loops associated with it are asymmetric. The deep overturning loop that results from southward flow at the surface and
formation of AABW occurs very close to where the upwelling is. Thus this water spends very little time at the surface before sinking again. This explains why AABW does not have time to take-up any radiocarbon from the atmosphere, and very likely plays a role in explaining why the surface nutrients are so high in this region, a feature we will discuss in Chapter 4.

(4) By contrast, the shallow overturning loop results from northward flow of AAIW and SAMW in the main thermocline and surface waters, and then sinking in the North Atlantic to form NADW. The upper branch of this loop spends a great deal of time in the main thermocline and surface waters. Thus, the NADW that forms from this upper ocean water has relatively high radiocarbon and also tends to have lost most of its nutrients before it is cooled and sinks to the abyss.

We show in Figure 2.4.11 a revision by Gnanadesikan and Hallberg [2002] of the Great Conveyor Belt Circulation diagram of Broecker [1991] which attempts to capture the main features of the above thermohaline circulation in a map view. In this diagram, the AABW and Deep Waters (NADW, NPDW, and RSDW) flows are depicted as horizontal dense water circulation loops within each basin in which there is no conversion of “dense” to “intermediate” or “light” waters. We note that the original Great Conveyor Belt Circulation diagram of Broecker [1991] had the bottom water upwelling all the way to the thermocline within the Indian and Pacific Oceans in association with vertical mixing as proposed by Robinson and Stommel [1959]. This is not supported by the observations and model we have discussed above (cf. Wunsch et al. [1983]). The conversion of dense (NADW, NPDW, and RSDW) to intermediate (AAIW and SAMW) water masses occurs instead within the Southern Ocean, where the strong upwelling associated with a large northward Ekman transport brings deep waters to the surface (cf. Toggweiler and Samuels [1995]; and Gnanadesikan [1999]). It is these intermediate waters which flow to the north and are converted to light thermocline waters that eventually return back to the North Atlantic to form NADW (cf. Sloyan and Rintoul [2001]), thereby closing the upper meridional overturning circulation loop. The Pacific light waters flow through the Indonesian Straits into the Indian Ocean, where they are joined by light waters formed in the Indian Ocean, and thence around South Africa. Note that NADW also includes an input of intermediate water that comes from the circumpolar current flowing around South America.

The remarkable “tour-de-force” model of the deep ocean circulation by Stommel [1958] is worth mention, as it is still highly influential in our thinking about deep ocean circulation. This model is based on the geostrophic vorticity equation (2.3.6) together with the assumption that interior mixing drives heat down into the ocean and leads to upwelling (as in the diffusive theories of the thermocline). This upwelling requires a compensatory input of fresh dense plumes of water sinking into the deep ocean from the high latitude regions where they form. Because the interior upwelling increases as $z$
increases ($z$ is positive upwards), the vertical velocity gradient $\partial w / \partial z$ is positive. From (2.3.6), we see that this requires a flow towards the pole in both hemispheres, directly toward the deep water formation regions! Stommel [1958] postulated that the flow away from the source regions must therefore be in narrow intense deep western boundary currents analogous to those at the surface. Indeed, one of the great successes of this model was the first observation of such currents after Stommel [1958] predicted their existence (cf. North Atlantic in Figures 2.4.3 and 2.4.5).

The geostrophic vorticity constraint provides powerful insights on the flow of the deep ocean in the ocean interior away from the boundaries. However, in its original conception, the “tour-de-force” model had water upwelling across density surfaces all the way from the deep ocean to the thermocline before returning from there to the deep water formation regions. The observations we have described are inconsistent with this, suggesting instead that most of the flow in the deep ocean is nearly horizontal, involving only modest interior upwelling.

We conclude our discussion of the meridional overturning circulation with a brief discussion of the influence of geothermal heating on the ocean circulation. Examinations of the impact of a single plume or a series of plumes of hydrothermal waters along a mid-ocean ridge show that the effect is mostly local (e.g., Stommel [1982] and Speer [1989]) and can probably be ignored in global ocean circulation models. (The release of excess mantle helium-3 by the hydrothermal plumes in the Pacific Ocean provides a remarkable illustration of the deep ocean circulation in this basin; Lupton [1998]). However, a more recent study of a spatially uniform geothermal heat flux shows that it can change the meridional overturning in the ocean by several Sverdrups (Scott et al. [2001]), which is quite significant compared to the deep ocean thermohaline circulation rates of order a few tens of Sverdrups obtained from the radiocarbon measurements.

### 2.5 Time Varying Flows

We began our discussion of ocean circulation by examining systematically the balances between the various forces in the equations of motion (2.2.1) assuming that the ocean circulation is in steady state (i.e., by setting the acceleration terms to zero). Inclusion of the acceleration terms gives rise to a rich set of additional behavior including waves, turbulent motions, eddies, convection, large-scale ocean-atmosphere interactions and many other such processes leading to variability on all time and space-scales (see Figure 2.4.7). This variability is a fundamental property of the ocean and the climate system, and therefore the ocean circulation can never be regarded as being in a true steady state. In this section, we discuss variability on two spatial and temporal scales. We first talk briefly about the role of meso-scale variability in the ocean, as these phenomena are among the most important contributors to oceanic variability in phytoplankton productivity (see e.g. Doney et al. [2003]). We then proceed to a
discussion of variability on interannual and decadal time scales. We will focus primarily
on an interannual phenomenon in the tropics, the El Niño-Southern Oscillation (ENSO),
which has received much attention in the public during recent years. We will then
proceed to a discussion of extratropical variability patterns that have emerged in various
recent analyses. Enfield [1989], Philander [1990], and Cane [1992] provide more
detailed treatments on ENSO and its influence on world climate, while Stocker [1996]
gives an overview of climatic variability on decadal and longer timescales.

MESO-SCALE VARIABILITY

The ocean is turbulent in nature. Any instantaneous image of sea-surface temperature
or surface chlorophyll reveals a rich spectrum of spatial variability with a particular
concentration of energy at the meso-scale, i.e. on spatial scales of tens to a few hundred
kilometers and at time-scales of a few days to weeks (Figure 2.5.1). The turbulent nature
of the ocean is a result of instabilities that occur because the full momentum equations
characterizing the flow have acceleration terms that depend on the velocity of the flow.
As the flow increases, the influence of these non-linear terms grows destabilizing the
large-scale flow. Starting from a laminar flow regime, i.e. where all paths along which a
given particle follows are parallel, the destabilization is often characterized first by the
appearance of meanders. After some critical threshold is reached, the flow becomes
progressively more unstable, first shedding vortices, i.e. circular flow features, and
eventually becoming fully turbulent.

There are many processes that create oceanic turbulence. Many energetic flow
features in the ocean, such as the western boundary currents and the Antarctic
Circumpolar Current (ACC; also known as the West Wind Drift; cf. Figures 2.2.3, 2.2.5,
and 2.2.9) are so intense that they are well beyond the critical threshold of instability,
explaining their intense meandering and eddy shedding. In this case the instability is
driven by the large-scale flow, with the energy mostly coming from the kinetic energy of
this flow (barotropic instability). Instabilities can also draw their energy from the
potential energy contained in tilted isopycnals associated with the flow (baroclinic
instability). Other sources of instability include flow past obstacles, such as islands, and
other interactions with complex topography, as well as external forces such as wind and
buoyancy fluxes.

Meso-scale eddies are a particularly prominent form of turbulence in the ocean
because (a) theirs is the scale that tends to be preferentially formed in unstable flows, and
(b) they tend to have a longer lifetime than most other turbulent phenomena. The scale of
meso-scale eddies is a few hundred kilometers in the tropics, of order a hundred
kilometers in the subtropics, and a few tens of kilometers in the high-latitudes (Chelton et
al., 1998). Meso-scale eddies tend to be geostrophically balanced, i.e. the pressure
gradient force balances Coriolis force (see equations (2.3.1) and (2.3.2)), explaining in

©J. L. SARMIENTO AND N. GRUBER
September 14, 2004
part their relatively long lifetime. The meso-scale eddies in the ocean are equivalent to the high and low pressure systems that characterize most of the atmospheric state at any given moment.

The most important effect of eddies on oceanic properties is to act as a mixing agent, thereby homogenizing gradients. However, it is now also well established that in certain regions, such as the Antarctic Circumpolar Current or in the equatorial regions, eddies can induce a mean transport, and that this transport can be of leading order importance in determining the mean-state of the system (e.g. Stammer [1998] and McWilliams and Danabasoglu [2002]). Evidence that eddies and other meso- and sub-mesoscale phenomena can also fundamentally alter the distribution and dynamics of marine ecosystems started to accumulate in the late 1970s (e.g. Denman [1976]). But it is only in the last 20 years, with the advent of satellite and autonomous in situ observation capabilities as well as high-resolution modeling, that such variations have come to be recognized as important components of marine biology and biogeochemical cycles. While earlier work focused on the role of mesoscale dynamics on phytoplankton patchiness (see review by Denman and Dower [2001]), much research in the past decade has studied the role of mesoscale processes in enhancing surface ocean productivity (e.g. Falkowski et al. [1991]; McGillicuddy et al. [1998]; Oschlies and Garcon [1998]; Abraham et al. [2000]).

The main reason for the enhancing effect of meso-scale eddies is that cyclonic eddies, i.e. those rotating around a depression of sea level, have upward tilting isopycnals in the interior that may bring deeper isopycnal surfaces with elevated nutrients into the euphotic zone (Figure 2.5.2). In nutrient limited ecosystems, this may lead to an enhancement of biological productivity. Anti-cyclonic eddies, i.e. those rotating around an elevation of sea level, have downward tilting isopycnal surfaces in the interior, leading to little expected change in productivity (Figure 2.5.2). Given this differential response, it has been suggested that the net effect of eddies in such nutrient limiting systems is to increase biological productivity. The magnitude of the influence of eddies on biological productivity is an issue of current debate (McGillicuddy et al. [1998]; Oschlies [2001]), but there is little doubt that meso-scale phenomena are among the dominant features in the spectrum of chlorophyll variability (Doney et al. [2003]).
Tropical Variability

The single most prominent signal in year-to-year variability in the atmosphere-ocean system is the El Niño/Southern Oscillation (ENSO), which has its center of action in the tropical Pacific Ocean. Before we look at the ENSO phenomenon in more detail, it is instructive to discuss first the climatological mean state of the tropical Pacific.

The tropical Pacific is characterized by warm surface water (29-30°C) in the west but much cooler temperatures in the east (22-24°C). The body of warm water in the west is known as the Pacific 'warm pool.' It is associated with the Indonesian low, a region of atmospheric heating, rising air, and therefore intense rainfall. This rising region constitutes one branch of a global zonal circulation, the so-called Walker circulation. A region of broad sinking over the eastern Pacific closes the Pacific cell of the Walker circulation. The cooler sea-surface temperatures in the eastern Pacific are the result of cold waters upwelling from below, forced by the divergent Ekman transport (see Figure 2.2.8a) due to the trade winds. The westward blowing trade winds also cause a deepening of the thermocline in the west, and a sea level difference of about 40 cm between the east and the west (Figure 2.5.3). It is the pressure gradient resulting from this sea surface height difference that gives rise to the Equatorial Undercurrent shown in Figure 2.3.14. Collectively, these east-west gradients and the associated westward blowing trade winds constitute a state of quasi-equilibrium for the ocean-atmosphere system.

Every few years, this state of equilibrium breaks down, giving rise to an ENSO event. As the name suggests, ENSO consists of two components. The first (mainly oceanic) component, El Niño, has historically been associated with a warm, weak north to south current appearing annually around December off the coasts of Ecuador and Peru (Philander [1990]). During El Niño conditions, the current is more intense than normal, and flows further to the south, greatly expanding the regional warm waters and leading to widespread mortality of fish and other marine wildlife. Very heavy rains in this region also accompany it. More recently it has been realized that these anomalies are connected to a much larger scale phenomenon in which the normally cold waters of the entire eastern tropical Pacific show dramatic warming. Therefore, the term El Niño is now used in a much broader sense to describe the warm state of the entire eastern tropical Pacific.

The second (mainly atmospheric) component of ENSO, the Southern Oscillation, is associated with large east-west shifts of mass in the tropical atmosphere between the Australian-Indonesian regions (the locus of the Indonesian low) and the southeastern tropical Pacific (the locus of the South Pacific high). Walker [1924] first described this oscillation; however, it was not until the 1960s that the connections between the El Niño...
and the Southern Oscillation phenomena were discovered. Today these phenomena are considered two aspects of one global-scale oscillation in the combined ocean-atmosphere system.

The tight coupling of the oceanic and atmospheric anomalies is evident in Figure 2.5.4, which compares an index of the Southern Oscillation with sea surface temperature anomalies in the tropical Pacific. This index is based on the normalized pressure difference between Tahiti and Darwin, Australia, and therefore measures the strength of the Indonesian Low versus the South Pacific high. Figure 2.5.4 also shows that El Niño events, i.e. the warm excursions of sea surface temperature in the tropical Pacific, are just one phase of the oscillation. The complementary phase, i.e. the cold anomaly, is often referred to as La Niña (Philander [1990]).

Cane [1992] and Philander [1990] have summarized the typical evolution of a warm ENSO event. In the early stages of an El Niño, large shifts in atmospheric mass occur so that the surface pressure decreases in the central and southeastern tropical Pacific and increases over the western Pacific and Indian oceans, resulting in a weakening of the trade winds. The onset of an El Niño is then typically marked in the western Pacific by a series of prolonged bursts of winds from the west (Webster and Peterson [1997]). They persist for one to three weeks and replace the normally weak winds from the east over the warm pool. The most important impact of these wind bursts is to severely perturb the upper ocean and excite the eastward propagation of large-scale so-called planetary waves. By analogy with large surface waves in the ocean, where the restoring force for the disturbance to the surface that is caused by the waves is gravity, planetary waves are waves for which the restoring force is the Coriolis force. The particular type of planetary wave found at the equator is a so-called equatorial Kelvin wave. These waves have their largest amplitude in the thermocline and have wavelengths of thousands of kilometers. They are trapped at the equator (until they hit the continents on the eastern side of the basin) and owe their existence to the change of sign of the Coriolis force at either side of the equator. The main effect of these Kelvin waves is to suppress the upwelling of cold water in the eastern basin and to deepen the thermocline (Figure 2.5.3b). This leads to an apparent eastward spreading of the warm water pool. However, this spreading is not a consequence of advective transport, but rather the result of the aforementioned changes in the vertical structure of the thermocline associated with the Kelvin wave.

The ocean response to the change in atmospheric forcing is not passive, however. The eastward spreading of the warm-water pool increases the heat flux from the ocean to the atmosphere in the central portions of the Pacific. The intense convective zone in the atmosphere over the western Pacific shifts eastward, further weakening the intensity of the trade winds in the central Pacific (Figure 2.5.5). This positive feedback intensifies the oceanic anomalies. In the mature state of an El Niño, the warm sea-surface
temperature anomalies extend across the entire eastern tropical Pacific from 20°S to 20°N, and even further poleward along the coast of the Americas (Figure 2.5.6). A precursor of the end of an El Niño event is the appearance of cold surface waters in the eastern equatorial Pacific by mechanisms that are only poorly understood. These low sea surface temperatures spread westward and inaugurate the opposite phase of the ENSO, La Niña. The duration of an El Niño episode is of the order of 18 months.

It is important to note that ENSO is linked to the seasonal cycle of sea-surface temperatures in the tropical Pacific. Instead of cooling early in the northern spring, the eastern Pacific continues to warm. This observation has led to the speculation that the atmosphere-ocean instability that produces ENSO results from the inability of the warm pool to export enough heat each year so that its heat content increases with time (Webster and Peterson [1997]). The warm pool is then preconditioned for El Niño, and the bursts of winds from the west may serve as triggers to release the stored energy.

The large changes in the ocean circulation, particularly the strength and source of upwelling, lead to dramatic changes in ocean biogeochemistry (Keeling and Revelle [1985; Feely et al. [1999]), biology (Barber and Chavez [1983; Lehodey et al. [1997; Chavez et al. [1999]) and ecosystem structure (Karl et al. [1995]). The impact of ENSO on the oceanic carbon cycle and atmospheric CO$_2$ will be discussed in detail in Chapter 10.

ENSO is a regional phenomenon, but its impact is global. As a direct consequence of the large shift of atmospheric weather patterns over the Pacific, the warm phases of ENSO are usually accompanied by severe drought over Australia and Indonesia, together with weakened summer monsoon rainfall over South Asia (Ropelewski and Halpert [1987]). Conversely, catastrophic flooding often occurs along the Pacific coast of South America. Moreover, coupling between ENSO and the large-scale atmospheric circulation may affect the climate of the extratropical regions remote from the Pacific (Wallace and Gutzler [1981]). Such coupling tends to change the probability of certain weather regimes in a particular region. For example, the chance of strong winter storms over the southwestern and southern United States is significantly enhanced, leading, on average, to more precipitation.

A phenomenon quite similar to El Niño exists in the tropical Atlantic (Philander [1990; Zebiak [1993]), but with a much smaller amplitude, likely a direct result of the fact that the width of the Atlantic is far smaller than that of the Pacific. In the tropical Pacific, interannual variations in the east and west are negatively correlated, whereas in the tropical Atlantic, they tend to be in phase and relatively uniform in the east-west direction. On the other hand, analyses of sea-surface temperatures in the tropical Atlantic have revealed that an important mode of variability exists in the meridional direction that appears to be symmetrical about the equator (Nobre and Shukla [1996; Chang et al.}
(1997]) (Figure 2.5.8). The fluctuations associated with this quasi-dipole are of much smaller amplitude than those in the Pacific, but nevertheless have a pronounced effect on rainfall variations in northeastern Brazil and the Sahel region. Although there is still considerable debate over the exact structure of this dipole-like variability (Enfield and Mayer [1997]) it is very likely that it represents, much like ENSO, an unstable coupled mode of the atmosphere ocean system (Zebiak [1993; Chang et al. [1997]).

**Extra-tropical variability**

Are there areas outside of the tropics that exhibit similar coherent variability on the interannual to decadal time-scale? In order to address this question Kawamura [1994] performed a so-called empirical orthogonal function (EOF) analysis of global sea-surface temperature (SST). Such an analysis attempts to split the temporal variance in the observations into sets of spatially correlated patterns (the EOF’s) and their associated magnitude of variability (the principal components or PC’s). The EOF’s that account for a large fraction of the variability are, in general, considered physically meaningful and connected with important centers of action.

Figure 2.5.7 shows the spatial pattern of the first two rotated EOF patterns and the correspondent time-series of the principal components. The spatial structure of EOF 1 shows strong positive values over the entire eastern tropical Pacific extending further northward along the coast of the Americas. The temporal variability of this pattern has a quasi periodicity of 2-5 years with little inter-decadal variability. From the resemblance of the spatial pattern with the SST anomalies seen in Figure 2.5.6 and from the similarity of the time-evolution of the EOF with the SO index shown in Figure 2.5.4, this mode can clearly be identified with the ENSO phenomenon. This confirms the notion that ENSO is the most important mode of oceanic variability on the interannual to decadal time-scale.

The spatial structure of EOF 2 shows positive values over the tropical and subtropical Indian Ocean, while pronounced negative values exist over the temperate North Pacific. This variability appears to be associated with the Pacific-Decadal Oscillation (PDO) described below. The time-series of the principal component of EOF 2 is drastically different from that of EOF 1. The time series of EOF 2 exhibits a long-term trend with some short-term interannual variability overlaid.

The spatial patterns of EOF 3 and EOF 4 and the corresponding time-series of the PCs are shown in Figure 2.5.8. By contrast with EOF 1 and EOF 2, these two patterns are confined to the Atlantic Ocean: EOF 3 to the North Atlantic, and EOF 4 to the tropical Atlantic. EOF 4 is clearly related to the tropical Atlantic variability described above, and EOF 3 is associated with the North Atlantic Oscillation (NAO) discussed below. Both time-series have a strong decadal variability.
In summary, the analysis of Kawamura [1994] reveals two centers of action for extratropical variability: the North Pacific, where variability is linked to the PDO, and the North Atlantic, where variability is dominated by the NAO. These two patterns will now be discussed briefly.

The Pacific Decadal Oscillation pattern is the leading mode of sea-surface temperature variability over the North Pacific (Mantua and Hare [2002]; see also EOF 2, Figure 2.5.7). A convenient index of this oscillation is therefore the principal component associated with this leading EOF. The PDO is associated with a deepening and southwestward shift of the Aleutian low, which causes an intensification of the Westerlies over the central North Pacific, and enhanced southerly winds along the coast of North America. These changes lead to a cooling over much of the temperate and subpolar central and western Pacific, whereas the eastern North Pacific (particularly the Gulf of Alaska) experiences anomalously warm waters. These SST changes in the temperate and subpolar central and western Pacific can be explained by increased heat loss caused by higher winds and deeper mixing, which brings cold waters to the surface. The surface anomalies propagate into the interior of the ocean and can be seen down to depths of over 400 m (Deser et al. [1996]). The mechanisms causing the PDO are only poorly understood (Mantua and Hare [2002]). On the interannual time scale, the PDO seems to be connected with the ENSO phenomenon. However, on longer timescales, the PDO is currently believed to represent an independent mode of variability.

The North Atlantic Oscillation (NAO) is the primary source of climatic variability on interannual to decadal timescales over the North Atlantic region (see also EOF 3; Figure 2.5.8). The atmospheric component of this phenomenon was discovered early in this century when meteorologists noticed that year-to-year fluctuations in wintertime air temperatures on either side of Iceland were often out of phase with one another (Walker and Bliss [1939]). More recent research has demonstrated that the NAO can be understood as the North Atlantic expression of a whole Arctic phenomenon, dubbed the Northern hemisphere Annular Mode (NAM) (Thompson and Wallace [2000]). This name originates from the observation that the primary mode of variability over the poles is zonally symmetric, and is associated with variations in the strength of the polar vortex. As the NAO and NAM are essentially the same phenomenon and our focus is on the North Atlantic, we continue to use here the historical term NAO. An index of the NAO has been defined as the normalized pressure difference between the Icelandic low and the Azores high, i.e., the meridional atmospheric pressure gradient over the North Atlantic.

A spectral analysis of the time history of the NAO index reveals that it contains significant variability at periods of 2 to 3 years and at periods of 6 to 10 years. These are significantly different from the ENSO periods. When the NAO index is high, the Icelandic low is anomalously deep and the Azores high exceptionally high. This causes a strengthening of the surface winds from the west across the Atlantic and increased
southward flow of cold polar air over the Labrador Sea. These anomalous winds lead to a drastic warming over the European continent, whereas northeastern Canada and the Labrador Sea experience anomalously cold conditions.

The anomalous patterns of the NAO must leave an imprint in the ocean. On interannual timescales, sea surface temperature anomalies across the entire Atlantic correlate very well with the NAO. Periods of high NAO lead to warm conditions in the western subtropical gyre and in the Norwegian Sea, whereas they lead to cold SST anomalies in the Labrador Sea. On these shorter timescales, the ocean response is probably just a passive response to the atmospheric forcing. However, variations associated with lower frequencies in the NAO will be reflected in subsurface water masses, because the ocean has a memory for winter conditions and integrates them over many years. Indeed, subsurface observations provide a clear depiction of the large changes that occurred in the North Atlantic region between the late 1960s, when NAO was mostly negative, to the early 1990s, when NAO was at a record high. In the late 1960s, convection in the Labrador Sea was tightly capped, whereas deep convection occurred in the Greenland Sea. By the early 1990s, conditions had completely reversed. This evolution is reflected in a substantial cooling and freshening of the mean properties of the Labrador Sea Water while the deep waters in the central Greenland Sea have warmed and become saltier. The changes are not restricted to the high latitudes. For example near Bermuda, a substantial freshening of the subtropical mode water has been detected over this period (Dickson et al. [1996]). Furthermore, wintertime mixing near Bermuda is also strongly affected by the state of the NAO, with positive phases usually associated with shallow winter mixed layers, and vice versa. We will discuss the impact of these variations on the upper ocean carbon cycle and biological productivity in chapter 10.

We conclude with a brief mention of the modes of variability in the Southern Ocean. White and Peterson [1996] recently suggested that there exists a coherent large scale variability pattern in sea-ice extent, sea surface temperature, wind speed and atmospheric pressure in the Southern Ocean (\textit{White and Peterson} [1996]). Analysis of observations over the last 20 years suggest that there are two such disturbances, one with a period of oscillation of about 3.3 years that is driven by sea-air interactions within the Southern Ocean; and a second with a period of about 5 years that may be forced by ENSO (cf., Venegas [2003] and Turner [2004]). The disturbances propagate eastward around Antarctica in a wave-like pattern. These variability patterns have been dubbed the "Antarctic Circumpolar Wave" (ACW). The putative ACW owes its existence to the fact that the Southern Ocean is the only oceanic domain encircling the globe and it is dominated by the strong eastward flow of the Antarctic Circumpolar Current. An alternative mode of Southern Ocean variability is thought to be associated with the Southern hemisphere Annular Mode (SAM), which is the counterpart to the NAM discussed above. Although this mode is the dominant mode of variability in the Southern hemisphere in the atmosphere.
(Thompson and Wallace [2000]), relatively little is known about its impact on the ocean. Hall [2002] demonstrated on the basis of an analysis of a long integration of a fully coupled climate model that this mode and not the ACW explains most of the variations in oceanic properties.
Problems

2.1 Which conservation law is responsible for the following situation: a rocket launched at 10°S in the direction of the South Pole appears to be deflected to the left.

2.2 Describe which pathway a rocket would take if it were launched in New Zealand in the direction of the North Pole.

2.3 Name two regions where upper ocean Ekman transport induced by wind leads to downwelling and two regions where this effect leads to upwelling. Explain the reasons for your choices.

2.4 Suppose the Earth were shaped like a cylinder rotating around the cylinder’s axis at the same angular velocity as earth. Would a moving body on the surface of this cylinder be subject to the Coriolis force?

2.5 During La Niñas, the trade winds in the eastern equatorial Pacific are particularly strong. Discuss the implications of this observation for upper ocean circulation and biology during these periods.

2.6 A long-term anomalous atmospheric pressure situation has led to a persistent strengthening of the westerlies and to a slackening of the trade winds by the same amount.

   (a) Discuss what happens to the strength of the equatorial upwelling and the strength of the subtropical downwelling.

   (b) How is the strength of the subtropical gyre affected?

2.7 Explain the term geostrophy. Name at least one situation in the ocean where this term is appropriate to use.

2.8 During El Niño years, the prevailing winds off the coast off Peru (Southern Hemisphere) change from being northward to being southward. Draw and explain the coastal circulation pattern before and during such years. Discuss the implications for the transport of nutrients from the deeper waters to the surface.

2.9 In 2092, nearly 100 hundred years after the first extrasolar planets were observed, two astronomers in Geneva find a solar system that rotates in opposite direction of our
They are particularly excited by the finding of a planet that looks almost identical to earth with oceans and continents, but which rotates in opposite direction of our earth.

(a) In which directions would you expect the trade winds to blow on this planet?

(b) Will you expect up- or downwelling on this planet near its equator?

(c) Assuming that the strength of the trade winds is similar on this planet to that on earth, how strong is the equatorial up- or downwelling on this planet relative to that on earth?

(d) In which direction would you expect the subtropical gyres to spin?

(e) Will we find strong boundary currents (such as the Gulf Stream) in the oceans on this planet? On which side of the oceans basins will you expect them to be found?

2.10 Many decades before the astonishing finding of the astronomers from Geneva, a team of U.S. astronomers identified a planet that rotates twice as fast as earth, yet in the same direction. In addition they found that this planet has a very similar distribution of oceans and continents as earth. Suppose that the strength and direction of the winds on this planet is similar to those on earth.

(a) How does the strength of the coastal upwelling and downwelling compare on this planet compare with that on earth?

(b) Would you expect a stronger or weaker subtropical gyre? Explain your answer.

(c) How strong do you expect the western boundary currents to be on this planet relative to the earth.

2.11 Draw a schematic latitude by depth section of salinity in the Atlantic ocean and identify the following major water masses:
- North Atlantic Deep Water
- Antarctic Intermediate Water
- Antarctic Bottom Water

2.12 What happens to the density of a seawater parcel if

(a) freshwater is added

(b) a strong wind cools it
2.13 What is meant by the expression “stratification”? What determines stratification and how does it relate to the vertical exchange of material (e.g. nutrients, etc)?

2.14 Referring to the Pacific Ocean PV and CFC-11 age contours in the Pacific Ocean shown in Figure 2.3.12, do an illustration and discuss the basic structure of PV contours on a typical isopycnal surface in the thermocline of the subtropical gyre, and discuss the flow associated with these contours, including the ventilated, recirculation, and shadow zones.

2.15 Suppose that all of the bomb produced radiocarbon and all 550 kg of the bomb produced tritium were introduced into the atmosphere in the major nuclear bomb tests of 1962.

(a) How much tritium was left on earth in the year 2002?

(b) What fraction of the bomb radiocarbon was left in the year 2002.

2.16 Redraw Figure 2.1.1 for a flow in the y-direction and show step-by-step how to obtain

\[ \frac{1}{V} \frac{\partial C}{\partial t} = -\frac{\partial C_v}{\partial y} \]

2.17 Consider an ocean in which surface waters sink into the deep ocean in the low latitudes, with the return flow from the deep ocean to the surface ocean occurring in the high latitudes. Use the integral form of the continuity equation to consider how a two box model for such an ocean should be constructed.

(a) What concentration should be used in the time derivative for the deep box?

(b) What concentrations should be used in calculating the inflows to the deep box and outflows from the deep box?
(c) All else being equal (e.g., surface ocean nutrients are depleted in low latitudes, but not in high latitudes), how would the deep ocean oxygen concentration of this ocean compare with that of the present ocean?

2.18 The Mediterranean is a closed basin except for exchange at the Straits of Gibraltar, where the inflowing water has a salinity of 36.2 ‰ and the outflowing water a salinity of 38.45 ‰. The net flux of fresh water into the Mediterranean by rivers and precipitation minus evaporation is $-40 \times 10^6$ liters per second (i.e., it loses more water through evaporation than it gains by rivers and precipitation). Its volume is $3.8 \times 10^6$ km$^3$.

(a) What is the flow of water in and out at the sill?

(b) What is the residence time of water in the Mediterranean Sea with respect to the inflow from the open ocean?

2.19 Consider a north-south continental boundary on the western side of an ocean basin in the southern hemisphere.

(a) Draw an $x$ by $z$ section that intersects the boundary and projects out into the ocean. Schematically illustrate the ocean current pattern if the wind is blowing at a constant rate from the north. Include a depth scale (i.e., show the approximate depth range of the Ekman layer).

(b) What is the total Ekman transport in Sverdrups across a 100 km long north-south section some distance from the shore? Assume a latitude of 20°S, and a $\tau_0$ of -0.4 dynes cm$^{-2}$ at the position of the north-south section.

Problems 2.19 to 2.21 refer to the following figure of a basin with an east-west dimension of 6000 km and a north-south dimension of 3300 km that extends from 20°N to 50°N:

The surface wind stress over this basin is given by
\[ \tau_0^x = -\tau_0 \cos \left( \pi \times \frac{y}{3300 \text{ km}} \right) \]

\[ \tau_0^y = 0 \]

where \( y \) is the distance in kilometers from the southern edge of the basin at 20°N, i.e.,

\[ y = (\phi - 20°N) \times 110 \text{ km} \]

with \( \phi \) being the latitude in degrees. (NOTE: this equation applies ONLY to the portion of the basin shown in the figure, not the world as a whole. As part of solving this problem, you need to calculate the \( y \) derivative of \( f \). Do this by using the Chain Rule to obtain

\[ \frac{\partial f}{\partial y} = \frac{\partial f}{\partial \phi} \frac{\partial \phi}{\partial y} \]

and using the circumference of the earth to determine the second derivative.) The value of the constant \( \tau_0 \) is

\[ \tau_0 = 1 \frac{\text{dyne}}{\text{cm}^2} = 1 \frac{\text{g}}{\text{cm} s^2} \]

The wind stress thus has the pattern:

2.20 What is the total Ekman transport across the width of the basin at 30°N? Give your answer in Svedrups.

© J. L. SARMIENTO AND N. GRUBER
September 14, 2004
2.21 What is the vertical velocity due to the Ekman divergence at 30°N?

2.22 Estimate the relative contributions of the terms containing the meridional wind stress gradient and meridional Coriolis parameter gradient to the upwelling/downwelling driven by the wind stress at 30°N.

2.23 It is observed that contours of constant PV on isopycnal surfaces can traverse 20° or more of latitude from, for example, 30°N to 10°N. By what fraction would the thickness of the isopycnal have to decrease in going from 30°N to 10°N if all of the flow continued straight south?
Panels

Panel 2.1.1

Advection contribution to tracer conservation equation

Consider the flow of water through an infinitesimally small cube such as the one depicted in Figure 2.1.1 with dimensions $\Delta x$ by $\Delta y$ by $\Delta z$ and a volume $V = \Delta x\Delta y\Delta z$. The total amount of tracer in the cube is:

$$V = \Delta x\Delta y\Delta z$$

(1)

The response of the tracer to a velocity $u$ in the $x$-direction such as that shown in Figure 2.1.1 is:

$$V \frac{\partial C}{\partial t} = C_1 u_1 \Delta y \Delta z - C_2 u_2 \Delta y \Delta z$$

(2)

Note that we have assumed that the volume is conserved, which allows us to bring it out of the time-derivative. Assume that the velocity and concentration vary continuously across the cube such that

$$u_2 = u_1 + \frac{\partial u}{\partial x} \Delta x$$

$$C_2 = C_1 + \frac{\partial C}{\partial x} \Delta x$$

(3)

Substituting (3) into (2) gives:

$$V \frac{\partial C}{\partial t} = \Delta y\Delta z \left[ C_1 u_1 - \left( C_1 u_1 + \Delta x C_1 \frac{\partial u}{\partial x} + \Delta u_1 \frac{\partial C}{\partial x} + \Delta x^2 \frac{\partial u}{\partial x} \frac{\partial C}{\partial x} \right) \right]$$

$$= -V \left( C_1 \frac{\partial u}{\partial x} + u_1 \frac{\partial C}{\partial x} + \Delta x \frac{\partial u}{\partial x} \frac{\partial C}{\partial x} \right)$$

(4)

Because of the small size of $\Delta x$ in the final term in the parentheses on the right hand side of (4), this term is much smaller than the other two. Canceling this term out and making use of the product rule to collapse the first two terms in the parentheses into a single term gives:

$$\frac{\partial C}{\partial t} = -\frac{\partial (C u)}{\partial x}$$

(5)
The advective contributions for the $y$- and $z$- directions can be derived in the same fashion, giving the final equation

$$\frac{\partial C}{\partial t} = -\frac{\partial (Cu)}{\partial x} - \frac{\partial (Cv)}{\partial y} - \frac{\partial (Cw)}{\partial z}$$

for the contribution of advection to the tracer concentration inside the cube.
Panel 2.1.2
Effect of eddies on tracer transport.

Eddies are the random motions of water parcels that do not result in any net transfer of water from one part of the ocean to another. However, eddies can carry tracer if the random motions in one direction have a different concentration than the random motions in the opposite direction. We can derive an equation for this process by starting with the advection equation (2.1.2):

\[ \frac{\partial C}{\partial t} = -\frac{\partial (Cu)}{\partial x} - \frac{\partial (Cv)}{\partial y} - \frac{\partial (Cw)}{\partial z} \]  

(1)

and separating the velocity and the concentration terms into a time average component (denoted by an overbar) defined as follows:

\[ \bar{()} = \frac{1}{T} \int_0^T () dt \]  

(2)

and deviations from the average (denoted by a prime) defined such that:

\[ u = \bar{u} + u' \]
\[ C = \bar{C} + C' \]  

(3)

Note from the definition of the deviations from the mean that the average of the deviations is equal to 0, i.e.,

\[ \bar{u'} = 0 \]
\[ \bar{C'} = 0 \]
\[ \bar{Cu'} = 0 \]
\[ \bar{C'\bar{u}} = 0 \]  

(4)

Substituting (3) into (1), taking the average, and making use of (4) to simplify the resulting equation gives:
\[
\frac{\partial \bar{C}}{\partial t} = - \left( \frac{\partial (\bar{C}u)}{\partial x} + \frac{\partial (\bar{C}v)}{\partial y} + \frac{\partial (\bar{C}w)}{\partial z} \right)
- \left( \frac{\partial (C'^u)}{\partial x} + \frac{\partial (C'^v)}{\partial y} + \frac{\partial (C'^w)}{\partial z} \right) 
\]  

(5)

The usual convention is to drop the overbar on all the terms except those involving the products of primes. This gives:

\[
\frac{\partial C}{\partial t} = - \left( \frac{\partial (Cu)}{\partial x} + \frac{\partial (Cv)}{\partial y} + \frac{\partial (Cw)}{\partial z} \right)
- \left( \frac{\partial (C'^u)}{\partial x} + \frac{\partial (C'^v)}{\partial y} + \frac{\partial (C'^w)}{\partial z} \right) 
\]

(6)

The terms in the first set of parentheses on the right-hand side represent the contribution of net advection to the tracer transport. The terms in the second pair of parentheses represent the contribution of eddies to the transport of tracer. The prime terms involve no net flow of water into or out of the box, since the average of the velocity fluctuations is zero. However, if the tracer concentration is correlated with the velocity direction, there will be a transport of tracer associated with these eddies. We assume that the eddy contribution can be represented by a Fickian type diffusion, i.e.,

\[
- \frac{\partial (C'^u)}{\partial x} - \frac{\partial (C'^v)}{\partial y} - \frac{\partial (C'^w)}{\partial z} = -D_x \frac{\partial C}{\partial x} - D_y \frac{\partial C}{\partial y} - D_z \frac{\partial C}{\partial z} 
\]

(7)

where \(D\) is an eddy diffusivity term with units of \(m^2 \text{s}^{-1}\). See text for further discussion.
Panel 2.1.3  
Mass conservation equation and incompressible flow.

Consider the flow of water through a cube of fixed volume such as that illustrated in Figure 2.1.1. The mass $m$ of water in the cube is:

$$m = \Delta x \Delta y \Delta z \rho = V \rho$$  \hspace{1cm} (1)

where $\rho$ is density. The change of the mass due to a flow in the $x$-direction is:

$$\frac{\partial m}{\partial t} = V \frac{\partial \rho}{\partial t} = \rho_1 u_1 \Delta y \Delta z - \rho_2 u_2 \Delta y \Delta z$$  \hspace{1cm} (2)

where $\rho_1 u_1 \Delta y \Delta z$ is the input from the left side of the cube, $\rho_2 u_2 \Delta y \Delta z$ is the outflow through the right side. Assume that the density and velocity change continuously across the cube such that

$$u_2 = u_1 + \frac{\partial u}{\partial x} \Delta x$$

$$\rho_2 = \rho_1 + \frac{\partial \rho}{\partial x} \Delta x$$  \hspace{1cm} (3)

Substituting (3) into (2) gives

$$V \frac{\partial \rho}{\partial t} = \Delta y \Delta z \left[ \rho_1 u_1 - \left( \rho_1 u_1 + \Delta x \rho_1 \frac{\partial u}{\partial x} + \Delta x u_1 \frac{\partial \rho}{\partial x} + \Delta x^2 \frac{\partial u}{\partial x} \frac{\partial \rho}{\partial x} \right) \right]$$  \hspace{1cm} (4)

which simplifies to:

$$\frac{\partial \rho}{\partial t} = -\rho_1 \frac{\partial u}{\partial x} - u_1 \frac{\partial \rho}{\partial x} - \Delta x \frac{\partial u}{\partial x} \frac{\partial \rho}{\partial x}$$  \hspace{1cm} (5)

As in Panel 2.1.1, we drop the final term in on the right-hand side due to its small size relative to the other terms. We make use of the product rule to collapse the first two terms in the parentheses into a single term. This gives:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial (\rho u)}{\partial x}$$  \hspace{1cm} (6)

Similar terms can be derived for the $y$- and $z$- directions, giving:
\[
\frac{\partial \rho}{\partial t} = -\frac{\partial (\rho u)}{\partial x} - \frac{\partial (\rho v)}{\partial y} - \frac{\partial (\rho w)}{\partial z} \tag{7}
\]

We can expand this expression by the product rule to give:
\[
\frac{\partial \rho}{\partial t} = -\rho \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) - \left( u \frac{\partial \rho}{\partial x} + v \frac{\partial \rho}{\partial y} + w \frac{\partial \rho}{\partial z} \right) \tag{8}
\]

and make use of the definition of the total derivative:
\[
\frac{d(\rho)}{dt} \equiv \frac{\partial (\rho)}{\partial t} + u \frac{\partial \rho}{\partial x} + v \frac{\partial \rho}{\partial y} + w \frac{\partial \rho}{\partial z} \tag{9}
\]

to obtain the following form of the mass conservation equation:
\[
\frac{d\rho}{dt} = -\rho \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) \tag{10}
\]

In dealing with tracers, we generally assume that seawater is incompressible, which is equivalent to assuming that volume is conserved. In fact, of course, it is not. Density increases by ~2% as the pressure increases from 1 atmosphere to the pressures of >400 atmospheres that are typical of the abyss. However, the effect of density variations due to compressibility is generally small compared to other tracer processes. An incompressible fluid is one in which a parcel of water conserves its density through time and as it moves around the water column. In other words, the total derivative of density is equal to zero
\[
\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + v \frac{\partial \rho}{\partial y} + w \frac{\partial \rho}{\partial z} = 0 \tag{11}
\]

From (9) we see that this condition implies that:
\[
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{12}
\]

Note that we do not need to consider diffusion of mass in the mass conservation equation, since, by definition, all net transfer of mass is included in the advection velocity terms.
### Tables

Table 2.4.1 Ventilation rates and time scales for deep water below 1500 m estimated from radiocarbon observations by *Stuiver et al.* [1983].

<table>
<thead>
<tr>
<th></th>
<th>Ventilation rate (Sv)</th>
<th>Ventilation time scale (yr)</th>
<th>Upwelling velocity (m yr(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic*</td>
<td>14</td>
<td>275</td>
<td>4</td>
</tr>
<tr>
<td>Circumpolar</td>
<td>-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>20</td>
<td>250</td>
<td>10</td>
</tr>
<tr>
<td>Pacific</td>
<td>25</td>
<td>510</td>
<td>5</td>
</tr>
<tr>
<td>World</td>
<td>55</td>
<td>500</td>
<td></td>
</tr>
</tbody>
</table>

*The NADW component of the Atlantic was specified by *Stuiver et al.* [1983]. The Atlantic upwelling velocity is based on the net inflow of 10 Sv. The ventilation time scale is based on the NADW inflow of 14 Sv summed to an estimated input of 7 Sv from the Circumpolar region, which is balanced by a loss of 11 Sv to the Circumpolar region, plus the upwelling of 10 Sv.*
References

Barber, R.T., and F.P. Chavez, Biological consequences of El Niño, Science, 222, 1203-1210, 1983.


©J. L. SARMIENTO AND N. GRUBER

September 14, 2004


Walker, G.T., Correlation in seasonal variations of weather, IX: A further study of world weather, Mem. Indian Dept., 4, 53-84, 1924.


Figure Captions

Figure 2.1.1  Schematic of an infinitesimally small cube with dimensions $\Delta x$ by $\Delta y$ by $\Delta z$ and a fixed volume $V = \Delta x\Delta y\Delta z$ with advection $u$ in the $x$ direction carrying tracer of concentration $C$.

Figure 2.2.1  Surface mean wind velocity for January and July based on the COADS data set (Deser et al. [1996]).

Figure 2.2.2  Schematic illustration of wind patterns (Pond and Pickard [1983]).

Figure 2.2.3  Streamlines of the ocean circulation in the top 50 m estimated with velocities obtained from the ECCO project (Stammer et al. [2002]). A streamline is a line that is parallel to the velocity field. The shading indicates the mean velocity in m s$^{-1}$ as indicated by the bar on the right-hand-side, with relatively higher velocities shaded a darker color.

Figure 2.2.4  Sea-surface height anomaly in m as estimated by the ECCO project (Stammer et al. [2002]).

Figure 2.2.5  Streamlines of the mean ocean circulation in the top 500 m estimated with velocities obtained from the ECCO project (Stammer et al. [2002]). The shading indicates the mean velocity in m s$^{-1}$ as indicated by the bar on the right-hand-side, with relatively higher velocities shaded a darker color.

Figure 2.2.6  Ekman transport and associated upwelling (a) and downwelling (b) resulting from wind blowing parallel to shore (Thurman [1990]).

Figure 2.2.7  (a) Sea-surface temperature from Levitus et al. [1994b]. (b) Sea-surface salinity from Levitus et al. [1994a].

Figure 2.2.8  (a) Equatorial divergence and upwelling resulting from Ekman transport driven by trade winds (Thurman [1990]). (b) Ekman downwelling resulting from convergent Ekman transport driven by the Westerlies and Trade Winds.

Figure 2.2.9  Wind driven gyre transport calculated with the model of Munk and Carrier [1950].

Figure 2.3.1  Potential density at the surface of the ocean as a function of potential temperature and salinity.
Figure 2.3.2  (a) Schematic illustration of terminology used to describe various features of the thermocline (Knauss [1997]). (b) Seasonal behavior of the mixed layer at 50°N, 145°W in the eastern North Pacific (Pickard and Emery [1990]).

Figure 2.3.3  (a) World Ocean Circulation Experiment (WOCE) sections used to depict the large scale distribution of properties in the world ocean. (b) Vertical section of potential temperature along the WOCE tracks in (a).

Figure 2.3.4  Schematic illustration of PV contours in an ocean with flat density layers.

Figure 2.3.5  Potential vorticity (units of $10^{-10}$ m$^{-1}$ s$^{-1}$) and depth (m) on two isopycnal surfaces in the North Atlantic (Sarmiento et al. [1982]). The potential vorticity, defined as $f/H$, is calculated from observations of the potential density distribution with $H$ defined as $\frac{1}{\sigma_\theta} \cdot (\frac{\partial \sigma_\theta}{\partial z})^{-1}$.

Figure 2.3.6  One year long particle trajectories on the $\sigma_\theta = 26.5$ isopycnal surface calculated with an ocean general circulation model. The point marks the end of the trajectory (Sarmiento et al. [1982]).

Figure 2.3.7  The atmospheric time history of tritium (cf., Dreisigacker and Roether [1978] and Doney et al. [1992]), two of the chlorofluorocarbons used as ocean circulation tracers (cf., Walker et al. [2000]), and bomb radiocarbon in the atmosphere (Rubin and Key [2002]). The tritium concentration is measured in rainfall and is in reported in TU units (1 Tritium Unit = 1 tritium atom per $10^{18}$ hydrogen atoms).

Figure 2.3.8  Maps of tritium on two isopycnal surfaces in the main thermocline of the North Atlantic (Sarmiento et al. [1982]). The depths of the surfaces are given in Figure 2.3.5.

Figure 2.3.9  Thermocline ventilation box model of Sarmiento [1983].

Figure 2.3.10  Thermocline ventilation rates determined from North Atlantic tritium observations by Sarmiento [1983] and Doney and Jenkins [1988], and from the combined tritium and helium-3 data set (labeled HE-3) by Doney and Jenkins [1988]. Figure taken from Doney and Jenkins [1988].

Figure 2.3.11  Isopycnal maps of ages determined by tritium/helium-3 dating (Jenkins and Wallace [1992]).
Figure 2.3.12 (a) Potential vorticity on the $\sigma_\theta = 26.4$ surface calculated from individual CTD and bottle profiles for the period 1950-present, and then objectively mapped. Heavy dashed contours mark the outcrop regions (Donxiao Zhang, personal communication). The contours are in units of $10^{-10}$ m$^1$ s$^{-1}$, with intervals of 0.5 between 0 and 5, and 2 for values greater than 5. (b) Ages calculated from CFC-11 on the $\sigma_\theta = 26.4$ surface. The locations of the data are shown on the map. The data are mostly from the WOCE period, centered in the early-to-mid 1990's, but also including data from three sections, TPS-10, TPS-24 and TPS-47, from the mid-to-late 1980's. (John Bullister, personal communication).

Figure 2.3.13 Sections of potential temperature and potential density across the Gulf Stream (Knauss [1997]).

Figure 2.3.14 A north-south section across the Equator at 140°W in the Pacific Ocean of various properties and velocity estimated by the geostrophic method. (Knauss [1997]).

Figure 2.4.1 North-south vertical sections of salinity along the WOCE sections shown in Figure 2.3.3a.

Figure 2.4.2 Basin wide zonal mean radiocarbon in ‰ for (a) the Atlantic and (b) the Pacific Oceans. The zonal means are obtained using the GLODAP gridded dataset of Key et al. [submitted] as described by Matsumoto and Key [2004]. The North Atlantic data are primarily from the Transient Tracers in the Oceans Study (~1982); the South Atlantic data are from the South Atlantic Ventilations Experiment (~1989); and the Pacific data from the World Ocean Circulation Experiment (~1994). The data have not been corrected for the Suess effect and bomb radiocarbon input.

Figure 2.4.3 Radiocarbon based age estimates in years with respect to the pre-industrial atmosphere at a depth of 3500 m based on the same observations as in Figure 2.4.2 (Key et al. [submitted]; and Matsumoto and Key [2004]).

Figure 2.4.4 A map of $PO_4^+$ at 3000 m taken from Broecker et al. [1998]. $PO_4^+$ is defined by equation (2.4.2).

Figure 2.4.5 The radiocarbon deficiency due to decay in the Atlantic Ocean. The contours are at 10‰ intervals, and are labeled by age using 80 years per 10‰ interval (taken from Broecker et al. [1991] and Broecker et al. [1998]).

Figure 2.4.6 Box model diagram of the deep ocean circulation below 1500 m according to Stuiver et al. [1983].
Figure 2.4.7  Space and time scales of oceanic processes depicted in a frequency (one over the time scale of the motion) versus wave number (2π over the length scale) diagram. Short time and space scale processes are in the upper right hand corner of the diagram, whereas longer time and space scale processes are in the lower left hand corner of the diagram. The shaded areas depict some of the relevant features of ocean circulation such as: 1, eddies, fronts, and western boundary currents; 2, deep-ocean convective overturning; and 3, turbulent motions including the very small scale vertical motions. Present OGCM’s of the large scale ocean circulation cover approximately the area depicted by “T.” Global eddy-resolving models cover the area depicted by “E,” and small scale models of ocean convection and eddy simulations of the ocean mixed layer cover the areas depicted by “C” and “L,” respectively. The point of this diagram is that present OGCM’s are not capable of resolving major features of the ocean circulation that are thought to be of importance. The dotted extensions of the boxes are an estimate of the expected gain in resolution over the next 6 years as computers improve. This figure is taken from Willebrand and Haidvogel [2001]. The thin lines in the upper part of the diagram are “dispersion curves” for linear gravity waves, and in the lower part of the diagram they are for so-called planetary waves (see discussion in Section 2.5).

Figure 2.4.8  Radiocarbon and chlorofluorocarbon simulations by 19 different OGCM’s (Matsumoto et al. [2004]) that participated in phase 2 of the Ocean Carbon Model-Intercomparison Project (OCMIP-2; Dutay et al. [2001]). The P2A model we discuss in the text is identified by the number 19 in this figure. Also shown are observations indicated by a dot and 2 σ error bars for radiocarbon and 15% uncertainty for the CFC-11 inventory estimate. The various water types are defined as follows: North Atlantic Deep Water (NADW) is Equator-60°N, 1000-3500 m; North Pacific Deep Water (NPDW) is Equator-60°N, 1500-5000 m; and Circumpolar Deep Water (CDW) is 90°S-45°S, 1500-5000 m. These boundaries have been applied to both observation and models, except that in models, a smaller NADW depth range of 1500-2500 m was used, because most models produce too shallow a NADW. (a) shows both NADW and NPDW on the vertical axis. The NADW is at the top of the figure (i.e., younger water). The lines have a 1:1 slope. (b) and (c) show the CFC-11 inventory in two regions of the ocean.

Figure 2.4.9  (a) Global meridional overturning stream function in the P2A OGCM described in the text. The contours are in Sverdrups (1 Sv=10⁶ m³ s⁻¹). North is to the left. (b) and (c) give a basin breakdown of the global meridional stream function for the Atlantic and Pacific Oceans, respectively.
Figure 2.4.10  A schematic depiction of the meridional overturning circulation based on the analysis of Toggweiler and Samuels [1993b].

Figure 2.4.11  A schematic depiction of the deep circulation based on the great conveyor circulation depiction of Broecker [1991] as modified by Gnanadesikan and Hallberg [2002].

Figure 2.5.1  Model simulated snapshots of sea surface temperature (SST) and chlorophyll $a$ off of West Coast of North America showing the influence of mesoscale eddies on both of these properties. The results are obtained with a eddy-resolving coupled physical-ecological model. Adapted from Gruber et al. [submitted].

Figure 2.5.2  Schematic representation of the impact of (a) cyclonic and (b) anticyclonic eddies on surface ocean productivity. Adapted from McGillicuddy et al. [1998].

Figure 2.5.3  Response of the thermal structure of the equatorial Pacific to changes in surface winds (Peixoto and Oort [1992]). (a) Under normal easterly trade wind conditions sea level rises to the west and the thermocline deepens. (b) When the trade winds relax, equatorial Kelvin waves propagate eastward, which lead to a rise in sea level and a deepening of the thermocline near the South American coast (El Niño conditions). (c) The normal situation is amplified during strong trade winds (La Niña conditions).

Figure 2.5.4  (a) Time series of sea-surface temperature anomalies in the Niño 3.4 region in the Pacific region (120˚W – 170˚W, 5˚S – 5˚N). The anomalies have been calculated relative to a base climatology of 1950 – 1979 using data from Trenberth and Hoar [1997]. (b) Time series of the Southern Oscillation Index. This index has been computed using monthly mean sea level pressures at Tahiti and Darwin, Australia. Actually shown is a spline fit through the monthly data with a stiffness chosen to filter out variability on timescales smaller than 4 months. Arrows indicate major El Niño events as defined Trenberth and Hoar [1997].

Figure 2.5.5  Schematic view of sea surface temperature and tropical rainfall in the equatorial Pacific Ocean during normal, El Nino, and La Nina conditions. The sea-surface temperature is shaded: blue-cold and orange-warm. The dark arrows indicate the direction of air movement in the atmosphere: upward arrows are associated with clouds and rainfall and downward-pointing arrows are associated with a general lack of rainfall. Taken from: http://www.cpc.noaa.gov/products/analysis_monitoring/ensocycle/enso_cycle.htm

©J. L. SARMIENTO AND N. GRUBER

September 14, 2004
Figure 2.5.6 Sea surface temperature anomalies (degree Celsius) during a typical El Niño obtained by averaging data for the episodes between 1950 and 1973 (Rasmusson and Carpenter [1982]). (a) March to May after the onset, (b) the following August to October, (c) the following December to February, and (d) May to July more than a year after the onset.

Figure 2.5.7 (a) Spatial patterns of the leading two rotated EOF modes of monthly mean SST anomalies based on data for 34 years from January 1955 to December 1988. Contour interval is 2.0 in relative units. (b) Time series of the principal components of the first two EOF modes shown in (a). These figures are from Kawamura [1994].

Figure 2.5.8 (a) Spatial patterns of the third and forth leading rotated EOF modes of monthly mean SST anomalies based on data for 34 years from January 1955 to December 1988. Contour interval is 2.0 in relative units. (b) Time series of the principal components of the second two EOF modes shown in (a). These figures are from Kawamura [1994].
Figure 2.1.1 Schematic of an infinitesimally small cube with dimensions $\Delta x$ by $\Delta y$ by $\Delta z$ and a fixed volume $V = \Delta x \Delta y \Delta z$, with advection $u$ carrying tracer of concentration $C$ in the $x$-direction.