8. Waves and other eddies

Introduction

Earlier we introduced the concept of the zonally averaged circulation, and defined "eddies" as zonal inhomogeneities. In atmospheric science, the term "eddy" is often used synonymously with "wave," but actually waves are special kinds of eddies. The Glossary of Meteorology of the American Meteorological Society defines a wave as "Generally, any pattern with some roughly identifiable periodicity in time and/or space." I don't find this definition very satisfactory. Alternatively, a "wave" can be defined as a process that transports energy without transporting mass. Waves can also transport momentum. Familiar examples are sound waves, and water waves seen at the beach; the latter are examples of what meteorologists call "gravity waves."

Because waves do not transport mass, they cannot spatially redistribute any property that is "attached" to the mass and transported only by moving mass, i.e., any property that is conserved under adiabatic frictionless processes. Examples of such quantities include the mixing ratios of water vapor and other minor constituents, the potential temperature (under dry adiabatic processes), and the potential vorticity. *Waves can spatially redistribute both momentum and energy, because both of these can be exchanged between fluid particles through the pressure-gradient force.* Virtually all waves are eddies, but only a subset of eddies are waves.

Waves can be produced by a variety of mechanisms (Fig. 8.1). Orographic forcing, heating that is localized at particular longitudes, and baroclinic and barotropic instability are among the mechanisms that can generate large-scale waves in the atmosphere. In the case of a wave forced by heating, the wave itself may or may not alter the heating.

Waves arise from a variety of physical mechanisms, and appear on many spatial and temporal scales. The atmosphere supports Rossby waves, Kelvin waves, inertia-gravity waves, mixed Rossby-gravity waves and sound waves. Wavelengths range from the circumference of the Earth, to the width of North America, to a few millimeters. Periods range from thousandths of a second to weeks. The spectral distributions of energy and fluxes over these various scales are complicated and interesting.

From elementary physics, we know that a particle does not undergo any net displacement as a small-amplitude wave passes by; it moves in a more or less circular path, and returns to its

starting point. This is consistent with the definition of "wave" given above. We also know, however, that it is possible for a surfer to ride on a large-amplitude water wave. We can think of the surfer as a particle of mass that is, in fact, transported by the wave. This illustrates that large-amplitude (often called "finite-amplitude") waves can transport mass. A wave that can transport mass can also transport any intensive property that is "attached" to the mass, such as potential temperature or water vapor mixing ratio. If a wave is defined, as above, as a process that transports energy but not mass, then beyond some limit a large-amplitude wave is not a wave at all. Note, however, that finite-amplitude waves tend to "break," i.e., they disintegrate into turbulence.

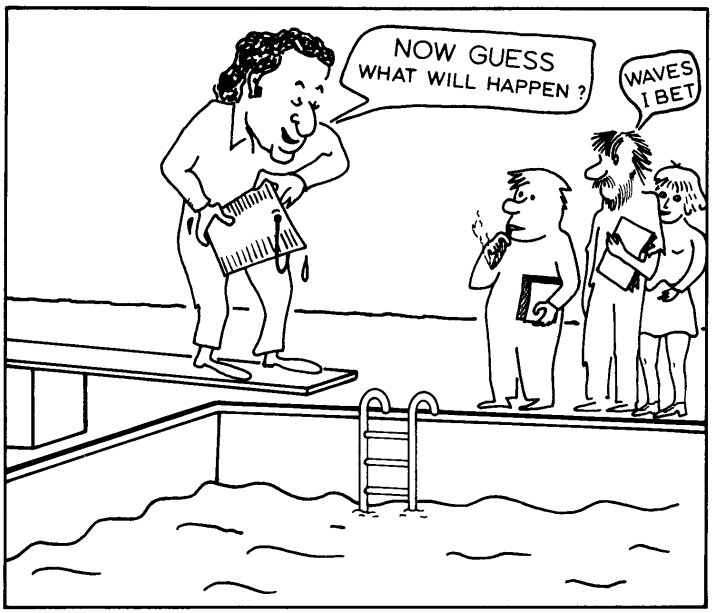

Figure 8.1: A cartoon that appeared in Morel (1973). The man on the diving board is Jule Charney. The other three people are presumably M.I.T. graduate students of the early 1970s.

The word "eddy" can serve as a generic term to denote zonal inhomogeneities that may or may not be waves.

As discussed earlier, it is useful to distinguish between "stationary" eddies, which are anchored to features (such as mountain ranges) on the Earth's surface and so appear in time-averaged (e.g. monthly mean) maps, and "transient" eddies that move and so are hidden in time averages.

Eddies appear in the winds, temperatures, geopotential heights, surface pressure, water vapor mixing ratio, and all of the other fields that characterize the circulation. They are important aspects of the circulation in their own right, and in addition they are of interest because they can affect the zonally averaged flow by producing fluxes and flux convergences. Such feedbacks of the eddies on the zonally averaged flow are the subject of the next chapter.

Rossby waves, which arise from conservation of potential vorticity on the sphere, are often referred to as "planetary waves," because they depend for their existence on rotation in the presence of spherical geometry, and so are characteristic of planetary atmospheres and oceans, although of course they also occur in stars. Rossby waves propagate westward relative to the mean flow, so that it is possible for them to be stationary (with respect to the surface) in a westerly regime. They can be excited in many ways, including interactions of the mean flow with mountains, convective events, and instabilities of various kinds. The most energetic Rossby waves have very large horizontal scales. To the extent that they are excited at low levels, their

energy propagation can be upward. Upward-propagating Rossby waves are believed to play an important role in stratospheric sudden warmings, as discussed in the next chapter. Meridionally propagating Rossby waves carry angular momentum between the tropics and middle latitudes.

Gravity waves arise through the action of the buoyancy force under stable stratification. They occur on many scales. When their scale is larger than the radius of deformation, they are strongly influenced by rotation, and are called "gravity-inertia waves." They can be produced by many mechanisms, including topographic forcing and convection. Vertically propagating gravity waves are thought to produce important vertical momentum transports that strongly affect the large-scale circulation, especially in the stratosphere and above. Today, the role of gravity waves in the general circulation is a "hot" topic.

The tropical atmosphere is home to two special classes of equatorially trapped waves. These are the "mixed Rossby-gravity waves," also called Yanai waves, which propagate westward; and Kelvin waves, which propagate eastward. As described later, these two types of waves have been implicated in the physical mechanism that drives the Quasi-Biennial Oscillation, although gravity waves are now believed to be quite important. Kelvin waves are also believed to play important roles in El Niño and in the Madden-Julian Oscillation. There are also equatorially trapped Rossby waves, which are for the most part similar to mid-latitude Rossby waves.

The purpose of the present chapter is to describe the observed climatological distribution of eddy activity, and to offer some theories of the mechanisms that produce both stationary and transient eddies.

Free and forced small-amplitude oscillations of a thin spherical atmosphere

Perturbation equations

Pierre-Simon Laplace (Fig. 8.2) was a French mathematician and astronomer whose name is now associated with many important ideas. His remarkably prescient study of the free and forced oscillations of a thin atmosphere on a spherical planet was originally published in French in 1799; an English translation followed in 1832. The 200-year-old paper is still very relevant today. Here we briefly outline his work, omitting the mathematical details. A *QuickStudy* giving those details and discussing the applications of Laplace's work to the problem of atmospheric tides is available. You can also find a more extensive discussion of this topic in Lindzen's (1990) book.

Figure 8.2: A portrait of Pierre-Simon Laplace, who lived from 1749 to 1827.

Laplace considered a spherical planet without mountains, and with a highly idealized basic state:

$$\bar{\mathbf{V}}_h = 0, \bar{\omega} = 0, \frac{\partial \bar{\phi}}{\partial p} = -\bar{\alpha}, p\bar{\alpha} = R\bar{T}(p), p_S = p_0 = \text{constant}. \tag{1}$$

Here $\bar{T}(p)$ is an arbitrary function of p. Note that $\bar{T}$ does not depend on latitude. This basic state has no meridional temperature gradient and no mean flow. It is, of course, in balance.

The linearized governing equations are

$$\frac{\partial u'}{\partial t} = (2\Omega \sin\varphi)v' - \frac{1}{a\cos\varphi}\frac{\partial \phi'}{\partial \lambda}, \tag{2}$$

$$\frac{\partial v'}{\partial t} = -(2\Omega \sin\varphi)u' - \frac{1}{a}\frac{\partial \phi'}{\partial \varphi}, \tag{3}$$

$$\frac{1}{a\cos\varphi}\left[\frac{\partial u'}{\partial \lambda} + \frac{\partial}{\partial \varphi}(v'\cos\varphi)\right] + \frac{\partial \omega'}{\partial p} = 0, \tag{4}$$

$$\frac{\partial}{\partial p}\left(\frac{\partial \phi'}{\partial t}\right) + S_p\omega' = -\frac{R}{c_p}\frac{Q}{p}, \tag{5}$$

where

$$S_p \equiv -\frac{\bar{\alpha}}{\bar{\theta}}\frac{\partial \bar{\theta}}{\partial p} \tag{6}$$

is the static stability, which depends only on p, and Q is the heating. Friction has been neglected in the momentum equations. Also,

$$\phi' = gz' + \Phi(\lambda, \varphi, t) \tag{7}$$

where $\Phi(\lambda,\varphi,t)$ is the external gravitational tidal potential, due to the moon and/or sun. In (7), we recognize that the atmosphere experiences gravitational accelerations due to the pulls of the moon and sun, in addition to that of the Earth. The variation of Φ with p is negligible, because the atmosphere is thin compared to the distances to the sun and moon. Note that these equations are valid only for atmospheres that are shallow compared to the planetary radius, a.

We look for *separable* solutions of the form

$$\begin{bmatrix} u' \\ v' \\ \omega' \\ \phi' \\ Q \\ \Phi \end{bmatrix} = \sum_n \left\{ \begin{bmatrix} U_n^{\sigma,m}(p) \\ V_n^{\sigma,m}(p) \\ W_n^{\sigma,m}(p) \\ Z_n^{\sigma,m}(p) \\ J_n^{\sigma,m}(p) \\ G_n^{\sigma,m}(p) \end{bmatrix} \Theta_n^{\sigma,s}(\varphi) \right\} \exp\underbrace{\left[i(m\lambda+\sigma t) \right]}_{phase}, \tag{8}$$

where the $\Theta_n^{\sigma,m}(\varphi)$ are as-yet-undetermined functions of latitude only, and

$$\begin{aligned} m &= \text{ zonal wave number } = 0,1,2\ldots \\ \sigma &= \text{ frequency, } \sigma < 0 \rightarrow \text{eastward moving} \\ &\qquad\qquad \sigma > 0 \rightarrow \text{westward moving} \end{aligned} \tag{9}$$

The superscripts (σ,m) simply denote the particular frequency and zonal wave number associated with each mode. The subscript n is introduced to recognize the possibility of multiple solutions, and the summation over n represents a superposition of these solutions. It can be shown that the set $\left\{\Theta_n^{\sigma,s}(\varphi)\right\}$ for all n is complete for $\frac{-\pi}{2} \le \varphi \le \frac{\pi}{2}$. At this point, we do not know what meridional structures are represented by the $\Theta_n^{\sigma,m}(\varphi)$.

Several pages of manipulation lead to the following two equations:

$$F\left(\Theta_n^{\sigma,m}\right) = -\varepsilon_n \Theta_n^{\sigma,m}, \tag{10}$$

$$\frac{d^2 W_n^{\sigma,m}}{dp^2} + \frac{S_p}{gh_n} W_n^{\sigma,m} = -\frac{R}{gh_n c_p}\left(\frac{J_n^{\sigma,m}}{p}\right). \tag{11}$$

Here F is a linear operator, defined by

$$F \equiv \frac{d}{d\mu}\left(\frac{1-\mu^2}{\upsilon^2-\mu^2}\frac{d}{d\mu}\right)-\frac{1}{\left(\upsilon^2-\mu^2\right)}\left[\frac{m}{\upsilon}\left(\frac{\upsilon^2+\mu^2}{\upsilon^2-\mu^2}\right)+\frac{m^2}{1-\mu^2}\right]; \tag{12}$$

$\upsilon \equiv \frac{\sigma}{2\Omega}$ is the normalized frequency; and $\mu \equiv \sin\varphi$, so that $d\mu \equiv \cos\varphi d\varphi$. We have introduced the nondimensional quantity

$$\varepsilon_n \equiv \frac{4\Omega^2 a^2}{g h_n} \ . \tag{13}$$

The quantity h_n , which appears in (11) and (13), is a "separation constant," because it arises during the separation of variables. It is called the "equivalent depth," for reasons that will become clear later.

Not surprisingly, Eq. (10) is called the meridional structure equation, and Eq. (11) is called the vertical structure equation. Recall that we have derived these equations using the assumptions that the basic state is at rest, and that the temperature depends on pressure (i.e., height) only. Separation of variables is not possible if the basic state is made more realistic, e.g., if the observed zonally averaged temperature and winds are used.

Eq. (10) was derived by Laplace about 200 years ago. It is often called the Laplace Tidal Equation, or LTE. It is a second-order ordinary differential equation, and so two boundary conditions are needed. It suffices to assume that the $\Theta_n^{\sigma,m}$ are bounded at the poles, i.e. at $\mu = -1$ and 1.

Note that (10) and its boundary conditions are satisfied quite nicely by the trivial solution $\Theta_n^{\sigma,m} \equiv 0$. Non-trivial solutions do exist, but only for particular choices of the parameters υ and/or h_n (or ε_n). If these parameters are chosen "at random," the only solution of (10) that satisfies the boundary conditions is the trivial solution $\Theta_n^{\sigma,m} \equiv 0$. A problem of this type is called an "eigenvalue" problem. The frequencies and/or equivalent depths that allow non-trivial solutions are called the eigenvalues, and the nontrivial $\Theta_n^{\sigma,m}$ are called the eigenfunctions or eigensolutions. For this particular problem, the eigenfunctions are the so-called *Hough functions*, which are discussed below.

All information about the planetary radius, rotation rate, and gravity is "buried" in the parameters ε_n and υ . The parameter ε_n is sometimes called by the imposing name "the terrestrial constant." Because it contains only two non-dimensional parameters characterizing the

planet, the LTE does not "know" or "care" very much about the particular planet to which it is being applied. For given υ and ε_n, the eigenvalues and eigenfunctions of (10) are the same for all planets, provided that the atmosphere in question is shallow compared to the planetary radius. This means that the solutions of (10) have a very broad applicability.

The vertical structure equation, (11), is a second-order ordinary differential equation for $W_n^{\sigma,m}(p)$. For now we regard $J_n^{\sigma,m}$ as known, so that (11) contains the single unknown $W_n^{\sigma,m}$. The assumption that $J_n^{\sigma,m}$ is known makes sense only if $J_n^{\sigma,m}$ is at least approximately independent of the motion. Such an assumption would be reasonable for heating due to absorption of solar radiation by ozone. It would be completely inappropriate for cumulus heating.

At the top of the atmosphere we apply the boundary condition

$$W_n^{\sigma,m} = 0 \text{ at } p = 0 . \tag{14}$$

This is exact. The exact lower boundary condition (in the absence of mountains) is $w \equiv \frac{Dz}{Dt}$ at $p = p_S(\lambda,\varphi,t)$. We use the linearized lower boundary condition

$$\frac{Dz'}{Dt} \cong \left(\frac{\partial z'}{\partial t}\right)_p + \omega' \frac{\partial \bar{z}}{\partial p} = 0 \text{ at } p = p_0 . \tag{15}$$

Here p_0 is the spatially and temporally constant value of p_S in the basic state. Because

$$gz' = \phi' - \Phi , \tag{16}$$

where $\Phi(\lambda,\varphi,t)$ is known, and using the hydrostaticity of the basic state, as expressed by

$$g\left(\frac{\partial z}{\partial p}\right)_{p=p_0} = -\bar{\alpha}_0 = -\frac{R\bar{T}_0}{p_0} \equiv -g\frac{H_0}{p_0} , \tag{17}$$

we can rewrite the linearized lower boundary condition (15) as

$$\frac{\partial \phi'}{\partial t} - \omega' g \frac{H_0}{p_0} = \frac{\partial \Phi}{\partial t} \text{ at } p = p_0 . \tag{18}$$

This involves both ϕ' and ω'. After some additional algebra to eliminate ϕ' in (18), we can finally express the lower boundary condition entirely in terms of $W_n^{\sigma,m}$, as

$$\boxed{\frac{dW_n^{\sigma,m}}{dp} - \frac{H_0}{h_n}\frac{W_n^{\sigma,m}}{p_0} = \frac{i\sigma}{gh_n}G_n^{\sigma,m} \text{ at } p = p_0} \,. \tag{19}$$

Note that the gravitational forcing enters the problem through the lower boundary condition on the vertical structure equation. The thermal forcing enters through the vertical structure equation itself. Neither type of forcing appears in the LTE.

Free oscillations of the first and second kinds

A free oscillation is one for which there is no thermal or gravitational forcing. It is, therefore, a mode of oscillation that can persist in the absence of forcing, something like the continuing tone emitted by a bell.

When there is *no thermal forcing* the vertical structure equation (11) reduces to

$$\frac{d^2W}{dp^2} + \frac{S_p}{g\hat{h}}W = 0 \,. \tag{20}$$

Here the superscripts (σ,m) and the subscript n have been dropped for simplicity. When there is no gravitation forcing the surface boundary condition (19) can be simplified to

$$\frac{dW}{dp} - \frac{H_0}{\hat{h}}\frac{W}{p_0} = 0 \text{ at } p = p_0 \,. \tag{21}$$

We also have

$$W = 0 \text{ at } p = 0 \,. \tag{22}$$

The system (20)-(22) has non-trivial solutions only for special values of h. These eigenvalues are denoted by $\hat{h}$. For $h \neq \hat{h}$, only the trivial solution (i.e. $W(p) \equiv 0$) exists. In order to find the $\hat{h}$ and the corresponding solutions for $W(p)$, we have to specify the static stability S_p as a function of height. Different choices for S_p will give different $\hat{h}$ and $W(p)$.

For the case of free oscillations, the solution procedure is summarized in Fig. 8.3. We have two eigenvalue problems: one from the vertical structure equation, and a second from the LTE. In the vertical structure problem, the eigenvalues are the equivalent depths. In the meridional structure problem, the eigenvalues are the frequencies.

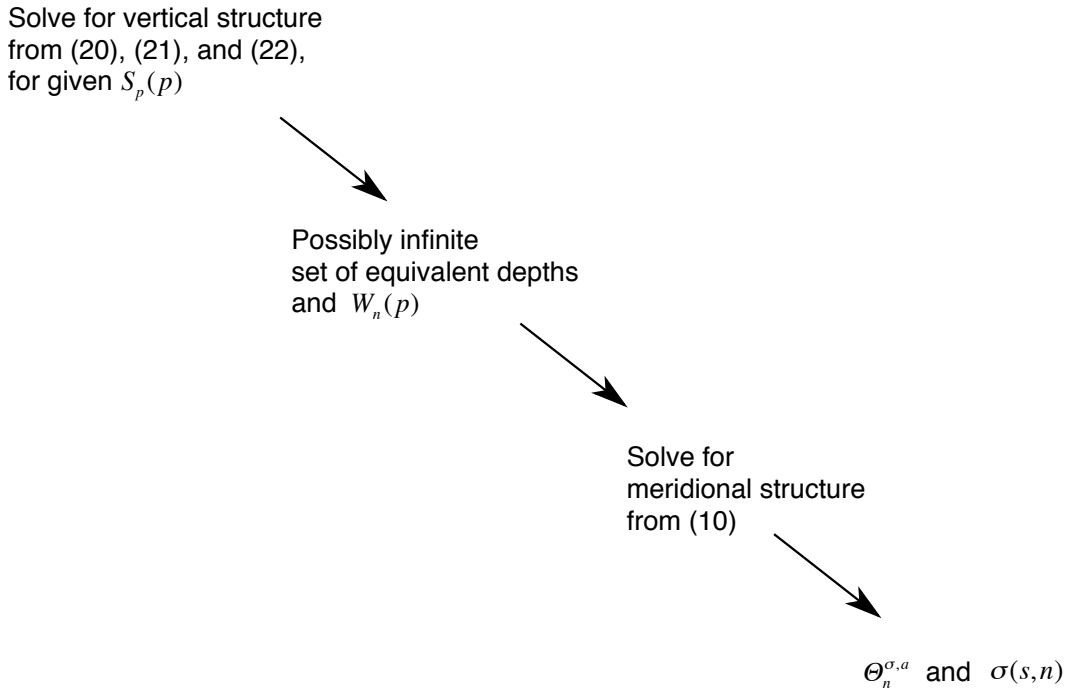

Figure 8.3: The solution procedure for free oscillations. The vertical structure equation is solved first, yielding the equivalent depths as eigenvalues. The frequencies are then obtained as eigenvalues of the LTE.

As a very simple example, suppose that $S_p = 0$. This means that the potential temperature is uniform with height. Then we find from (20) that

$$\frac{d^2W}{dp^2} = 0 \,. \tag{23}$$

A solution of (23) that is consistent with the upper boundary condition (22) is

$$W = Ap \,, \tag{24}$$

where A is an arbitrary constant. Use of (24) in the lower boundary condition (21) gives

$$\hat{h} = H_0 \,. \tag{25}$$

This is the only possible equivalent depth for free oscillations of an isentropic atmosphere. With more general stratifications there can be many (infinitely many) equivalent depths. As mentioned earlier, this is why we need the subscript n, i.e. we use n to denote an individual solution. The procedure used to find the equivalent depth in this simple example can also be used for other stratifications.

With $\hat{h}$ given by (25), nontrivial solutions of the LTE, (10), exist only when there is a special relation (called a dispersion relation) among υ, m, and n. The parameter n is sometimes called the "wave type."

The Hough functions (i.e., the eigen-solutions of the LTE) have been tabulated by Longuet-Higgins (1968) and others. Here we consider only some limiting cases. First suppose that there is no rotation, so that $\upsilon = \dfrac{\sigma}{2\Omega} \to \infty$. We continue to assume that $S_p = 0$ so that (25) applies. For this case, we find that

$$\upsilon^2 F \to \frac{d}{d\mu}\left[\left(1-\mu^2\right)\frac{d}{d\mu}\right] - \frac{m^2}{1-\mu^2}, \tag{26}$$

and

$$\upsilon^2 \varepsilon \to \frac{\sigma^2 a^2}{gh} = \frac{\sigma^2 a^2}{gH_0}. \tag{27}$$

Then the LTE reduces to

$$\frac{d}{d\mu}\left[\left(1-\mu^2\right)\frac{d\Theta}{d\mu}\right] + \left(\frac{\sigma^2 a^2}{gH_0} - \frac{m^2}{1-\mu^2}\right)\Theta = 0. \tag{28}$$

It can be shown that (28) has solutions that are bounded as $\mu \to \pm 1$ only for

$$\frac{\sigma^2 a^2}{gH_0} = n(n+1), \quad n = 1,2,3\ldots \tag{29}$$

This is a dispersion equation that relates frequency and wave number. Frequencies that satisfy (29) are eigenvalues. The eigenfunctions, i.e., the nontrivial solutions of (28), are called associated Legendre functions of order n and rank s, denoted by

$$\Theta_n = P_n^m(\mu), \text{ for } n \geq m .$$

(30)

What this means is that, for the case of no rotation, the Hough functions reduce to the associated Legrendre functions; in other words, the Hough functions are associated Legendre functions modified by rotation. Note that n and m are both integers such that $n \geq m$. The P_n^m are discussed in the *QuickStudy* on spherical harmonics.

When (30) is combined with the longitudinal structure shown in (7), we find that the two-dimensional horizontal structure of the waves is given by

$$Y_n^m(\mu,\lambda) = P_n^m(\mu)\exp(im\lambda) .$$

(31)

These are the spherical harmonics. Here n is the total number of nodal circles, s is the zonal wave number, and $n-m$ is the number of nodes in the meridional direction, also known as the "meridional nodal number." In summary, for the case of no rotation, the horizontal structures of the free waves are spherical harmonics.

The solutions found here are external gravity waves. They are called "external" because they have no nodes in the vertical and do not propagate vertically. An isentropic atmosphere cannot support internal gravity waves, because when θ is uniform with height a vertically displaced particle does not experience a gravitational restoring force. A stratified (i.e., non-isentropic) atmosphere can support both external and internal gravity waves.

The frequencies of the external gravity waves can be written as

$$\sigma = \pm\frac{\sqrt{n(n+1)gH_0}}{a} ;$$

(32)

they depend on the wave's horizontal scale through the two-dimensional index, n , but they are independent of m . For example, when $n=1$, m can be either 0 or 1 (because $n \geq m \geq 0$), but both modes have the same frequency. This is not true when rotation is present, because with rotation the zonal direction (in which scale is measured by m) becomes physically "different" from the meridional direction.

Now consider $\Omega \neq 0$, still for an isentropic atmosphere, and neglect all details. Define a stream function ψ and a velocity potential χ , so that

$$u' = -\frac{1}{a}\left(\frac{\partial \psi}{\partial \varphi}\right)_p + \frac{1}{a\cos\varphi}\left(\frac{\partial \chi}{\partial \lambda}\right)_p,$$
$$v' = \frac{1}{a\cos\varphi}\left(\frac{\partial \psi}{\partial \lambda}\right)_p + \frac{1}{a}\left(\frac{\partial \chi}{\partial \varphi}\right)_p \quad . \tag{33}$$

The vorticity is then $\xi_p = \mathbf{k}\cdot\left(\nabla_p \times \mathbf{V}_h\right) = \nabla_p^2 \psi$, and the divergence is $\delta_p = \nabla_p \cdot \mathbf{V}_h = \nabla_p^2 \chi$. The equation of horizontal motion can be differentiated to obtain

$$\frac{\partial}{\partial t}\nabla_p^2 \psi + \beta v' + f\nabla_p^2 \chi = 0 , \tag{34}$$

(the vorticity equation) and

$$\frac{\partial}{\partial t}\nabla_p^2 \chi + \beta u' - f\nabla_p^2 \psi = -g\nabla_p^2 z' \tag{35}$$

(the divergence equation), where

$$\beta \equiv \frac{2\Omega\cos\varphi}{a} = \frac{1}{a}\frac{df}{d\varphi} \tag{36}$$

is the rate of change of the Coriolis parameter with latitude. We can also show (see the problems at the end of this chapter) that for an isentropic atmosphere

$$\frac{\partial z'}{\partial t} + H_0 \nabla_p^2 \chi = 0 . \tag{37}$$

Equations (33) through (36) form a closed set that can be solved for ψ , χ , and z' , as well as u' and v' .

From (34) and (37), we see that when $\Omega \neq 0$ (so that $\beta \neq 0$), stationary motion cannot exist unless $v' = 0$. For nontrivial stationary motion with $v' = 0$, it follows from the equation of zonal motion, (2), that $m = 0$, i.e., *the motion must be both purely zonal (because $v' = 0$) and zonally uniform*.

Margules (1893) and Hough (1898) showed that the LTE has two classes of solutions, which they named Free Oscillations of the First and Second Classes. For the case of $\varepsilon_n \equiv \dfrac{4\Omega^2 a^2}{g h_n}$ small (weak rotation or large equivalent depth), approximate solutions of (34) and (35) can be obtained by expanding in spherical harmonics (see Longuet-Higgins, 1968). These Free Oscillations of the First Class (FOFC) are essentially gravity waves, satisfying

$$\chi \cong A_n^m P_n^m(\mu) e^{i(m\lambda+\sigma t)}, \ \psi \cong \text{constant (irrotational)}, \tag{38}$$

the dispersion equation (32), which was obtained for $\Omega = 0$. Haurwitz (1937) derived a more accurate expression for the frequencies of the FOFC:

$$\sigma \cong \frac{\Omega m}{n(n+1)} \pm \sqrt{\frac{\Omega^2 m^2}{n^2(n+1)^2} + n(n+1)\frac{gH_0}{a^2}}\ . \tag{39}$$

This should be compared with (32). For large n, (39) reduces to (32). The additional terms in (39) involve Ω, and vanish for $\Omega = 0$. For $n \geq 4$ the error that comes from using (32) instead of (39) is less than 1%. From (39) we see that eastward propagating inertia gravity waves have frequencies slightly different from those of westward propagating inertia gravity waves. The difference is due to rotation.

The Free Oscillations of the Second Kind (FOSC) are also called Rossby-Haurwitz waves. They are purely nondivergent and so satisfy $\psi \cong B_n^m P_n^m(\mu) e^{i(m\lambda+\sigma t)}$, $\chi \cong \text{constant}$. The dispersion equation for Rossby-Haurwitz waves is

$$\boxed{\sigma \cong \frac{2\Omega m}{n(n+1)} > 0}\ . \tag{40}$$

Eq. (40) shows that for Rossby-Haurwitz waves, in contrast to pure gravity waves, σ does depend explicitly on the zonal wave number, m. Because $\sigma > 0$, the FOSC always propagates westward, with zonal phase speed $\dfrac{-2\Omega}{n(n+1)}$. "Bigger" waves, with smaller values of n and/or m, propagate faster.

Rossby-Haurwitz waves are almost nondivergent. The wave solutions can be found by assuming $\chi = \text{constant}$ (i.e., nondivergence) from the beginning. For nondivergent flow, Eq. (34) reduces to

$$\frac{\partial}{\partial t}\nabla^2\psi + \frac{2\Omega}{a^2}\frac{\partial\psi}{\partial\lambda} = 0 \,. \tag{41}$$

Using the form of the Laplacian in spherical coordinates, this can be written as

$$i\sigma\left\{\frac{1}{a^2\cos^2\varphi}\left[\cos\varphi\frac{d}{d\varphi}\left(\cos\varphi\frac{\partial}{\partial\varphi}\right) - m^2\right]\right\}\hat{\psi} + \frac{2\Omega i m}{a^2}\hat{\psi} = 0 \tag{42}$$

or

$$\left\{\frac{d}{d\mu}\left[\left(1-\mu^2\right)\frac{d}{d\mu}\right] + \left(\frac{2\Omega m}{\sigma} - \frac{m^2}{1-\mu^2}\right)\right\}\hat{\psi} = 0 \,. \tag{43}$$

This is another eigenvalue problem. Nontrivial solutions of (43) exist when $\frac{2\Omega m}{\sigma} = n(n+1)$, which leads directly to (40). The nontrivial solutions take the form $\hat{\psi} = B_n^m P_n^m(\mu)$.

The westward propagation of Rossby-Haurwitz waves is due to the Earth's sphericity, which implies variations of the Coriolis parameter with latitude. To see this, rewrite (41) as

$$\frac{\partial\zeta}{\partial t} = -\beta v \,, \tag{44}$$

where

$$\beta \equiv \frac{2\Omega\cos\varphi}{a} = \frac{1}{a}\frac{df}{d\varphi} \geq 0 \,. \tag{45}$$

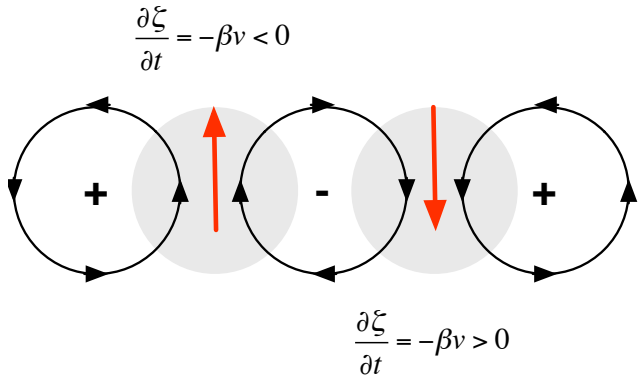

Figure 8.4: Chain of vortices along a latitude circle, illustrating the westward propagation of Rossby waves.

Consider a chain of vortices along a latitude circle, as shown in Fig. 8.4. Where $v > 0$, to the west of the place where $\zeta < 0$, we have $\beta v > 0$, so we get $\frac{\partial \zeta}{\partial t} < 0$. This occurs to the west of the place where $\zeta < 0$. Similarly, where $v < 0$, to the west of the place where $\zeta > 0$, we have $\beta v < 0$, so we get $\frac{\partial \zeta}{\partial t} > 0$. This is a simple way to see why Rossby waves propagate westward relative to the mean flow. All other things being equal, the waves propagate westward more quickly near the Equator, where β is large, than at higher latitudes. As a result, a Rossby wave that propagates towards the Equator will tend to curve westward as it moves.

m, n	Observed phase speed	Computed phase speed
1, 2	70	115
2, 3	40	53
3, 4	20	28
1, 4	20	28
2, 5	12	16
3, 6	8	9

Table 8.1: A comparison of the observed (middle column) and computed (right column) westward phase speeds of the planetary waves, from Eliassen and Machenhauer (1965). The units are degrees of longitude per day. The computed phase speeds are for non-divergent Rossby waves.

A direct test of the theory of non-divergent Rossby waves was made by Eliassen and Machenhauer (1965) and Deland (1965). They used spherical-harmonics to analyze the 500 mb stream function, isolating transient waves by taking the difference in 24 hours. Their results,

illustrated in Fig. 8.5, show westward propagation. Table 8.1 compares the computed and observed phase speeds, in degrees of longitude per day. *The model over-predicts the westward phase speeds*, due to the neglect of the effects of divergence.

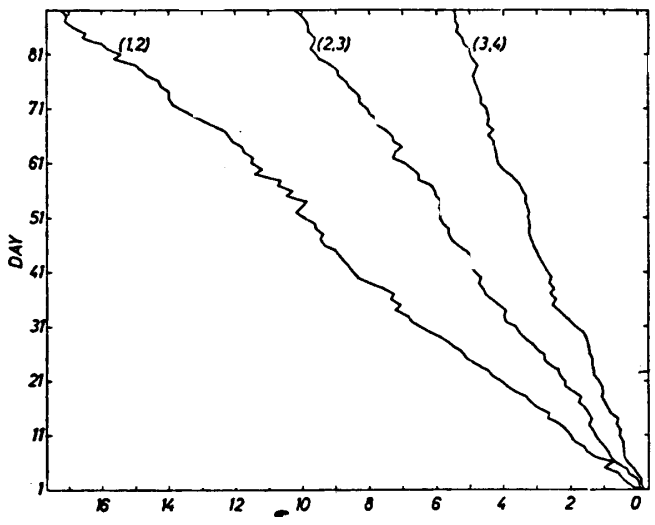

Figure 8.5: Successive daily values of the phase angle for the 24 hour tendency field, at the 500 mb level for the components $(m,n)=(1,2),(2,3),(3,4)$, during the 90 day period beginning 1 December 1956. The ordinate is time, in days. The abscissa represents the number of westward circulations around the Earth after the first passage of the Greenwich meridian, and so it is a measure of longitudinal phase propagation. From Eliassen and Machenhauer (1965).

The atmospheric tides are forced oscillations, in which the forcing is either thermal or gravitational. The mathematical problem is quite different from the case of free oscillations, even though the equations are (nearly) the same. The key point is that the frequency and zonal wave number of the forcing are known. Within the limits of applicability of the linear equations, the frequency and zonal wave number of the atmospheric response must be the same as those of the forcing. The LTE (the meridional structure equation) must then be solved as an eigenvalue problem for the equivalent depth -- quite different from the case of free oscillations. Once the equivalent depth has been determined, the vertical structure equation can be solved for the response to the forcing.

Observations of stationary and transient eddies in middle latitudes

Blackmon (1976) discussed the observed eddy activity in the Northern Hemisphere, as seen in the 500 mb geopotential height. He used a ten-year record, and considered both summer and winter conditions. The data were available twice per day, at 00 Z and 12 Z.

Blackmon filtered the data in both space and time, in order to isolate particular space-time scales. He expanded the height fields into spherical harmonics Y_n^m (see the *QuickStudy* on spherical harmonics), where superscript m denotes the zonal wave number, and subscript n denotes the "two dimensional index." The number of nodes in the meridional direction, i.e. the "meridional nodal number" (try saying that five times, fast) is $n-m\geq 0$; note that $m\leq n$ is required. The largest value of n considered by Blackmon was $n=18$. Note that $m=18$ corresponds to a zonal wave length of 20° of longitude, which in middle latitudes corresponds to a wavelength of roughly 1000 km.

The spherical harmonics form a complete orthonormal basis that can be used to represent an arbitrary function on the sphere (see the *QuickStudy*). Using the spherical harmonics, the data can be expanded, with "triangular truncation," as follows:

$$Z(\lambda,\varphi) = \sum_{m=-M}^{M} \left(\sum_{n=|m|}^{M} C_n^m Y_n^m \right). \tag{46}$$

Here the C_n^m are the expansion coefficients, and M is a suitably chosen positive integer. Larger values of M allow a more detailed representation of $Z(\lambda,\varphi)$. The zonal wave number is m, and the parameter n is called the "two-dimensional index." For further discussion, see the *QuickStudy*.

Blackmon defined three categories of spatial scales, based on the two-dimensional index:

Regime I: $0 \le n \le 6$*, or "long waves;"*

Regime II: $7 \le n \le 12$*, or "medium-scale waves;" and*

Regime III: $13 \le n \le 18$*, or "short waves.*"

Because Blackmon's truncation scheme is based on the two-dimensional index, a particular wave can have nodes in the zonal direction or in the meridional direction, most have both. Note that for all three regimes $0 \le m \le n$. All three regimes therefore contain modes with small values of m, i.e. long zonal scales.

The expansion coefficients for each set of waves can be determined for each observation time. They were filtered in time, using three filters:

Low pass: Admits periods in the range longer than or equal to 10 days;

Medium-pass: Admits periods in the range 2.5 days to 6 days;

High-pass: Admits periods in the range 1 day to 2 days.

Note that a period of one day represents the most rapidly fluctuating wave that can be captured by the twice-a-day data that was used in Blackmon's study.

The lower panel of Fig. 8.6 shows the time-averaged 500 mb height averaged over nine winters. The features seen here are stationary waves. Note the two prominent troughs, one near the east coast of North America, and the other near Japan. These same features can be seen in the figures of Chapter 2, of course.

The upper panel of Fig. 8.6 shows the total root-mean-square (rms) geopotential height, for winter, without time or space filtering. There are three prominent "centers of action," in the North Pacific, the North Atlantic, and over Siberia.

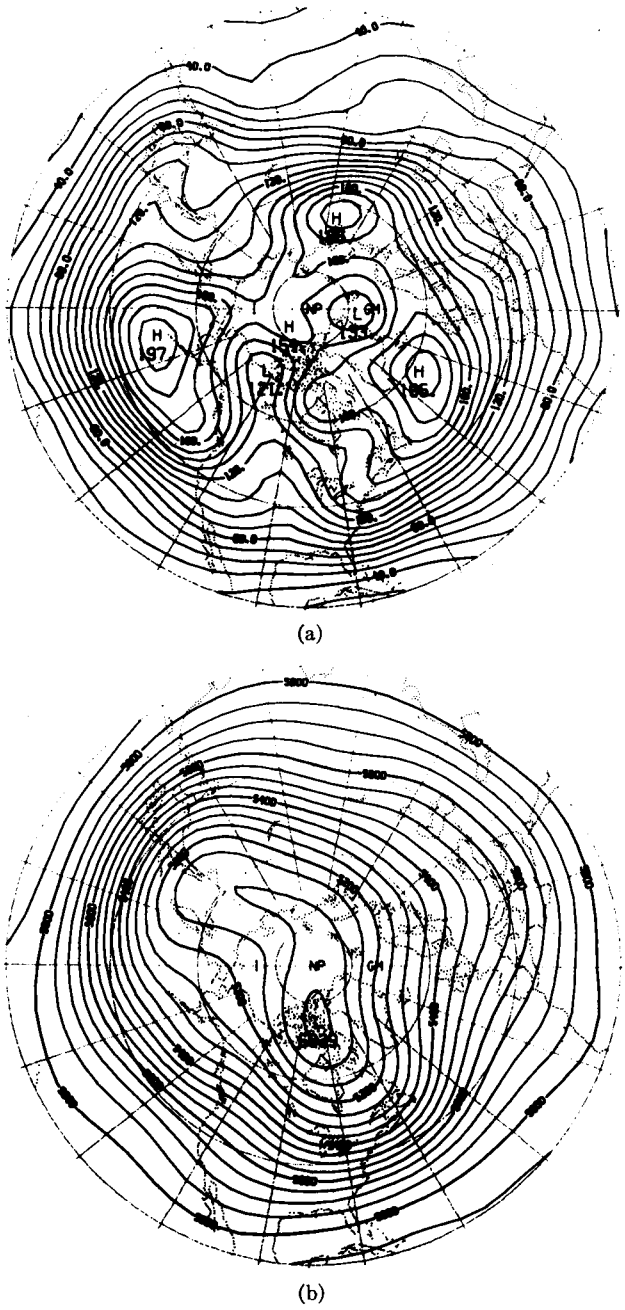

Figure 8.6: Maps of the geopotential height for nine winters: (a) rms unfiltered, contour interval 10 m; (b) average, contour interval 50 m. From Blackmon (1976).

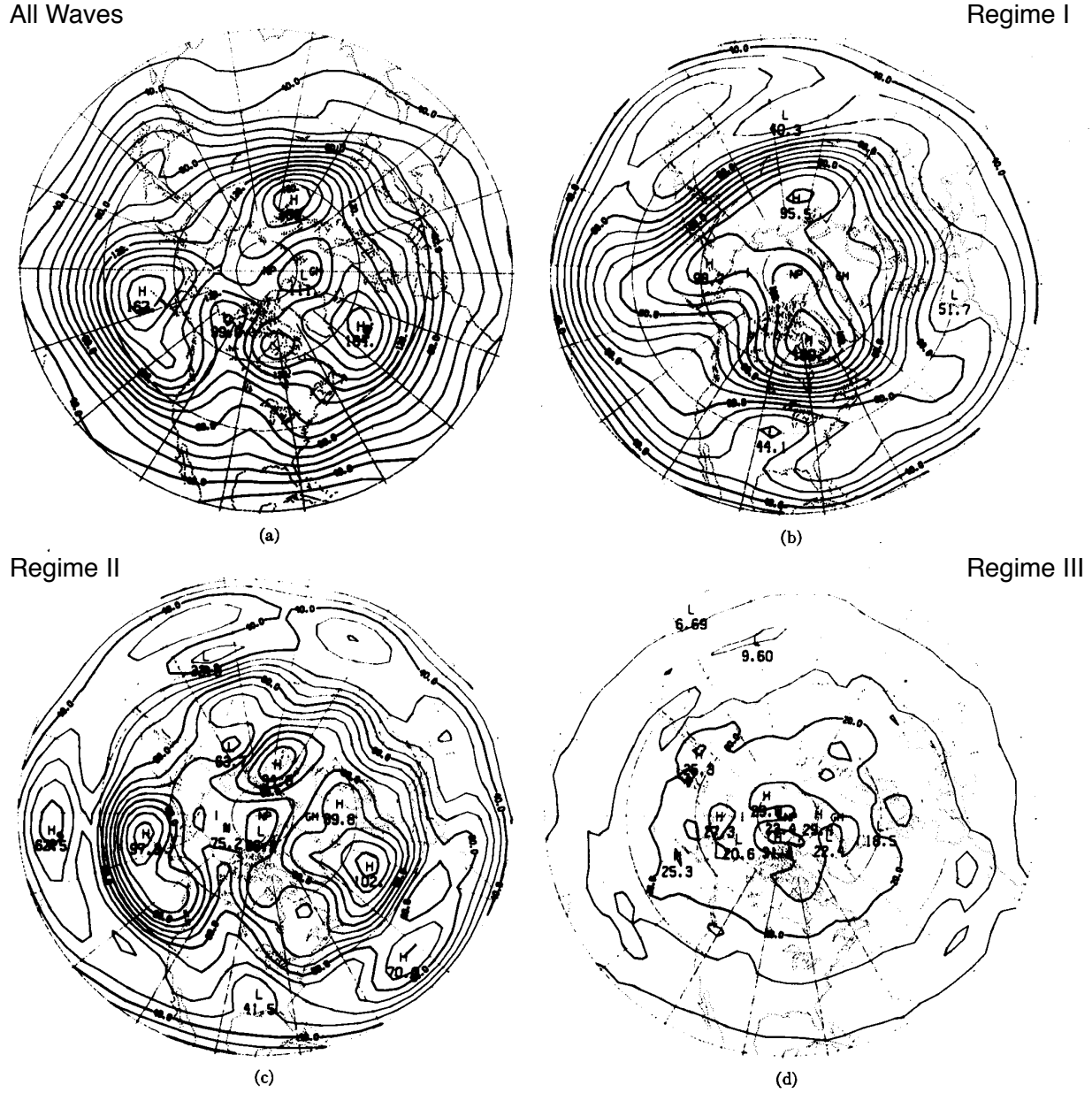

Figure 8.7: Maps of the low-pass filtered rms fields (winter): (a) all waves, contour interval 10 m; (b) waves in Regime I, contour interval 5 m; (c) waves in Regime II, contour interval 5 m; (d) waves in Regime III, contour interval 5 m. From Blackmon (1976).

Fig. 8.7 shows low-pass filtered (long period) rms winter heights, for all spatial scales (top left), for Regime I (top right), for Regime II (bottom left), and for Regime III (bottom right). It is clear that very little contribution comes from Regime III. Both Regime I and Regime II contribute significantly. The Regime II contribution shows maxima in regions where blocking commonly occurs. Blocking will be discussed later.

Fig. 8.8 is similar to Fig. 8.7, but for medium-pass waves, i.e. those of "synoptic" periods in the range of 2.5 to 6 days. For this range of periods, most of the action comes from Regime II and Regime III; the long waves do not contribute much. Note, however, that the contour interval is smaller than that used in Fig. 8.7, so that the total amount of activity indicated in Fig. 8.8 is less.

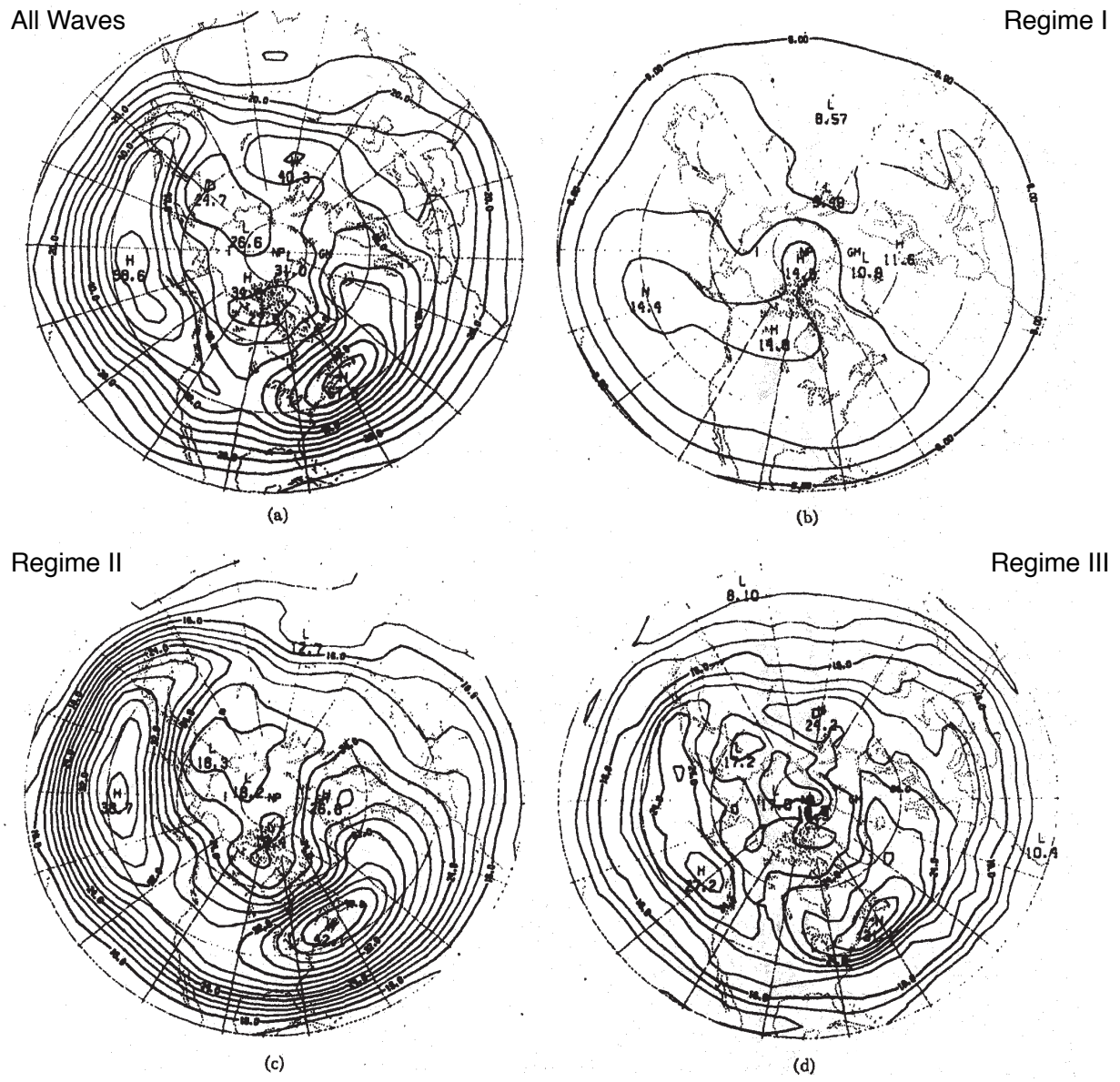

Figure 8.8: Maps of the medium-pass filtered rms fields (winter): **(a) all waves, contour interval 5 m; (b) waves in Regime I, contour interval 2 m; (c) waves in Regime II, contour interval 2 m; (d) waves in Regime III, contour interval 2 m. From Blackmon (1976).**

Even less power resides in high-pass waves (not shown).

n	Frequency (days)$^{-1}$															
	0	1/15	2/15	3/15	4/15	5/15	6/15	7/15	8/15	9/15	10/15	11/15	12/15	13/15	14/15	15/15
18	51.7	49.8	45.8	39.3	37.0	26.7	21.2	17.3	13.7	10.1	7.2	5.3	4.1	3.4	3.2	3.4
17	68.7	66.5	58.6	48.1	40.1	34.0	27.9	22.1	15.8	11.0	7.6	5.4	4.3	3.1	3.2	3.3
16	98.5	97.5	82.4	69.6	56.5	45.0	34.8	26.0	18.4	12.2	7.9	5.5	4.1	3.3	3.3	3.5
15	141.6	131.0	108.1	85.9	70.3	58.6	75.6	32.3	19.7	11.6	7.3	5.1	4.1	3.3	3.3	3.4
14	210.1	195.5	158.9	122.3	99.5	79.2	53.3	32.3	18.2	10.8	6.9	4.9	3.5	3.3	3.3	3.4
13	302.8	272.5	213.2	164.7	129.0	92.6	56.1	31.6	16.8	9.4	5.7	4.4	3.2	2.9	2.9	3.3
12	502.6	435.2	311.1	225.5	159.9	99.8	53.5	26.8	13.6	8.0	5.3	3.9	3.2	3.0	3.0	3.4
11	746.3	639.7	443.3	290.6	176.5	94.4	43.8	21.1	10.6	6.0	4.1	3.6	3.0	2.8	3.0	3.4
10	1190.1	943.3	545.9	322.8	181.0	80.5	33.6	15.6	8.3	5.4	3.9	3.3	3.0	2.9	3.1	3.6
9	1815.2	1417.6	739.3	323.8	135.8	55.0	23.2	11.4	6.1.	4.1	3.1	2.8	2.4	2.7	2.9	3.6
8	3305.7	2175.9	732.7	245.2	104.5	42.0	17.3	9.1	4.3	3.6	2.3	2.6	2.2	2.7	2.5	3.4
7	3732.6	2370.4	698.2	191.2	71.2	29.3	12.9	7.7	3.8	3.6	2.3	2.9	1.9	2.6	2.5	4.3
6	3731.5	2374.2	657.5	149.3	51.0	20.6	9.0	5.8	2.4	2.8	1.3	2.3	1.5	2.4	2.3	4.2
5	3563.5	2192.6	522.5	114.0	39.1	16.2	7.1	4.9	3.4	2.7	2.2	2.2	2.2	2.2	2.6	4.8
4	2553.1	1600.8	420.2	95.1	34.5	13.8	5.7	4.2	2.1	2.6	1.5	2.1	1.3	2.0	2.1	3.8
3	1448.2	992.6	329.2	79.7	27.0	9.4	4.6	3.3	2.1	1.8	1.4	1.5	1.2	1.5	5.3	4.6
2	1115.3	642.0	146.3	41.8	18.5	8.3	3.3	2.3	1.6	1.6	1.1	1.3	1.2	1.4	2.2	3.7
1	294.5	221.7	105.9	45.1	15.8	4.1	2.0	1.7	1.2	1.0	0.9	1.0	1.0	1.1	6.0	12.4
0	168.6	92.6	17.1	4.7	2.5	1.4	1.0	0.8	0.6	0.6	0.6	0.6	0.6	0.6	1.8	3.5

Table 8.2: Power as a function of wave number and frequency for the winter season. Units: m^2 rad^{-1} (15 days)$^{-1}$.

Table 8.2 shows that in winter the greatest power is found along a "ridge" running from lower left (small n, low frequency) to upper right (large n, high frequency). By far the largest power occurs in the low frequencies and for two-dimensional index n on the order of 6. The diurnal tide is also apparent near the bottom right, with $n = 2$.

The corresponding results for the summer season are omitted here for brevity. In summer, the wave amplitudes are greatly reduced, as is the strength of the mean flow, and the centers of action are shifted towards the pole. As in winter, the long waves and low frequencies dominate.

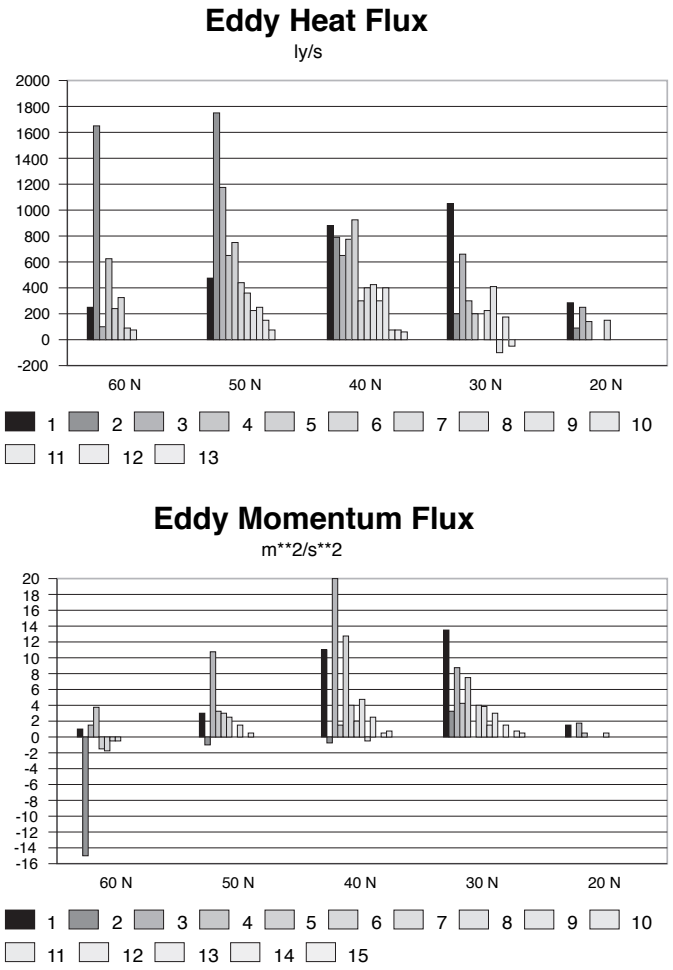

Figure 8.9: Upper panel: The total eddy heat transport across a latitude circle as a function of wave number for selected latitudes, for January 1962. Lower panel: The eddy momentum transport averaged with respect to pressure, as a function of zonal wave number, for selected latitudes, based on observations for January 1962. In both panels the gray scale at the bottom denotes the zonal wave number. From Wiin-Nielsen et al. (1963).

One conclusion from Blackmon's study is that most of the transient eddy energy resides at the low frequencies and long wavelengths. This has been known for a long time. For example, Wiin-Nielsen et al. (1963) analyzed the total heat and momentum transport across a latitude circle as a function of wave number for selected latitudes. As shown in Fig. 8.9, by far the strongest contributions come from zonal wave numbers less than 10. For middle latitudes, this corresponds to wave lengths longer than 2000 km. The results show that the waves of largest

spatial scale do most of the work of transporting energy and momentum poleward; the "synoptic-scale" and mesoscale eddies that we think of as "weather" play only a relatively minor role. They are fleas on the back of the general circulation.

Theory of orographically forced stationary waves

Stationary waves are forced by mechanical and/or thermal effects that are anchored to the Earth's surface. Mountain ranges can produce waves either by orographic forcing (i.e., by blocking the flow) or by acting as elevated heat sources. Thermal forcing is also associated with land-sea contrasts, sea surface temperature gradients, and the like.

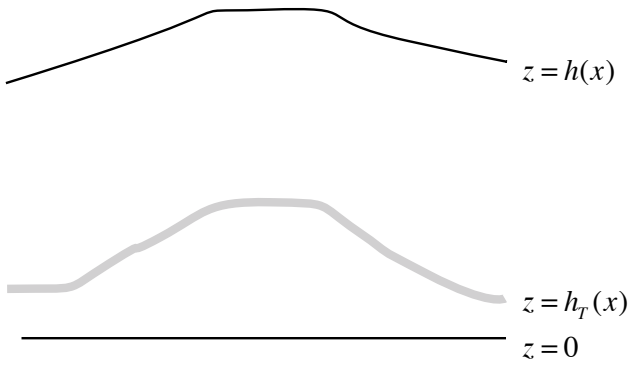

Figure 8.10: Schematic illustrating shallow water flowing over a "mountain."

Held (1983) summarized the work of Charney and Eliassen (1949), who tried to understand the effects of orographic barriers on stationary waves in middle latitudes. The starting point is consideration of the conservation of potential vorticity in shallow water (see the *QuickStudy* on the shallow water equations), with the quasi-geostrophic approximation, i.e.

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \cdot \nabla\right) q = -\left(\frac{r\zeta}{h - h_T}\right), \tag{47}$$

where

$$q \equiv \frac{\zeta + f}{h - h_T}. \tag{48}$$

Here h is the height of the free surface, $[h]$ is the zonally averaged value of h, h_T is the height of the topography, $h^* = h - [h]$, $\mathbf{V}_g$ is the geostrophic wind, and r is a non-negative Rayleigh friction coefficient, whose role will be discussed later. See Fig. 8.10. We use the quasi-geostrophic approximation, with

$$\mathbf{V}_g = \frac{g}{f_0}\mathbf{k} \times \nabla h \quad \text{and} \quad \zeta = \frac{g}{f_0}\nabla^2 h. \tag{49}$$

We linearize about the zonally averaged state, which is assumed to be geostrophically balanced. Using (48), we can show that the basic-state zonal wind satisfies

$$[u] = \frac{g}{f_0}\frac{\partial [h]}{\partial y}.$$

(50)

Here the square brackets denote a zonal mean. We assume that $[u]$ is independent of y as well as x and t. The potential vorticity gradient of the zonally averaged flow is given by

$$[h]\frac{\partial}{\partial y}[q] = \beta + \frac{f_0^2[u]}{g[h]}.$$

(51)

In linearizing about the basic state given by (49) and (50), we assume not only that the perturbation vorticity and height are small, but also that h_T is small compared to $[h]$, so that we can neglect any products of h_T with a perturbation quantity. The linearized version of (46) is then

$$\left(\frac{\partial}{\partial t} + [u]\frac{\partial}{\partial x}\right)q^* + v^*\frac{\partial [q]}{\partial y} = -\frac{r\zeta^*}{[h]}.$$

(52)

where all perturbation quantities can be expressed in terms of h^*, i.e., the departure of h from its zonal mean, as follows:

$$v^* = \frac{g}{f_0}\frac{\partial h^*}{\partial x},$$

(53)

$$\zeta^* = \frac{g}{f_0}\nabla^2 h^*,$$

(54)

$$[h]q^* = \varsigma^* - f_0\frac{\left(h^* - h_T\right)}{[h]}$$
$$= \frac{g}{f_0}\nabla^2 h^* - f_0\frac{\left(h^* - h_T\right)}{[h]}$$

(55)

By substitution from (50) and (52)-(54), we can rewrite the perturbation potential vorticity equation, (51), as

$$\left(\frac{\partial}{\partial t}+[u]\frac{\partial}{\partial x}\right)\left\{\frac{g}{f_0}\nabla^2 h^* - f_0\frac{\left(h^*-h_T\right)}{[h]}\right\}+\frac{g}{f_0}\frac{\partial h^*}{\partial x}\left\{\beta+\frac{f_0^2[u]}{g[h]}\right\}=-r\frac{g}{f_0}\nabla^2 h^*,$$

(56)

or, after rearranging,

$$\frac{\partial}{\partial t}\left(\frac{g}{f_0}\nabla^2 h^* - \frac{f_0}{[h]}h^*\right)+\frac{g}{f_0}[u]\frac{\partial}{\partial x}\left(\nabla^2 h^*\right)+\beta\frac{g}{f_0}\frac{\partial h^*}{\partial x}+r\frac{g}{f_0}\nabla^2 h^* = \frac{[u]f_0}{[h]}\frac{\partial h_T}{\partial x}.$$

(57)

Here some nice cancellation has occurred, and we have placed the "topographic forcing" term, involving $\frac{\partial h_T}{\partial x}$, on the right-hand side, to set it apart. The waves described by (56) are "forced" by topography, which enters mathematically through the inhomogeneous term on the right-hand side, i.e. $-[u]\frac{f_0}{[h]}\frac{\partial h_T}{\partial x}$. In the presence of such forcing, a non-zero h^* is demanded by (56).

Assume that the perturbations have the form

$$h^* = \mathrm{Re}\left\{\hat{h}\exp\left[i\left(kx+ly-\sigma t\right)\right]\right\},$$

(58)

where $\hat{h}$ is a constant, and also that the topography satisfies

$$h_T = \mathrm{Re}\left\{\hat{h}_T\exp\left[i\left(kx+ly\right)\right]\right\}.$$

(59)

With (57), a positive value of indicates eastward propagation. Note that we could choose $l=0$, which would give us the special case of no meridional variations. Substitution of (57) and (58) into (56) gives:

$$\left\{-\sigma\left(K^2+\lambda^2\right)+k\left([u]K^2-\beta\right)-irK^2\right\}\hat{h}e^{-i\sigma t}=[u]\lambda^{-2}k\hat{h}_T.$$

(60)

Here K is the total wave number, which is defined by

$$K^2 \equiv k^2+l^2,$$

(61)

and λ is the radius of deformation, which is defined by

$$\lambda^2 \equiv g[h] / f_0^2 .$$

(62)

Note that λ is defined in terms of the basic state. We see directly from (59) that the amplitude of the waves, as measured by $\hat{h}$, is proportional to the amplitude of the forcing, as measured by $\hat{h}_T$. The proportionality factor is rather complicated, however.

First consider the special case in which a free wave exists in the absence of topographic forcing. Then (59) reduces to a dispersion formula, which can be written as

$$-\sigma\left(K^2 + \lambda^{-2}\right) + k\left([u]K^2 - \beta\right) - irK^2 = 0 ,$$

(63)

or

$$\sigma = \frac{k\left([u]K^2 - \beta\right) - irK^2}{K^2 + \lambda^{-2}} .$$

(64)

Eq. (64) describes a damped, free Rossby wave in a balanced mean flow. The wave dies out after a finite time, because there is no forcing to sustain it against the frictional damping. To see how the friction leads to damping, write

$$\sigma = \sigma_0 - i\left(\tau_f\right)^{-1},$$

(65)

where

$$\sigma_0 \equiv \frac{k\left\{[u]K^2 - \beta\right\}}{K^2 + \lambda^{-2}} ,$$

(66)

and

$$\left(\tau_f\right)^{-1} \equiv \frac{rK^2}{K^2 + \lambda^{-2}} .$$

(67)

Then (57) can be rewritten as

$$h^*(x,y,t) = e^{-t/\tau_f} \operatorname{Re}\left\{\hat{h} \exp\left[i(kx + ly - \sigma_0 t\right]\right\} .$$

(68)

The wave-like solution, with period σ_0 , decays with e-folding time τ_f . The free waves are killed off by friction.

A stationary wave is one for which $\mathrm{Re}\{\sigma\}=0$. Under what conditions can a free Rossby wave be stationary? It can be seen from (66) that a stationary free wave is possible only for $[u]>0$. The reason is that the Rossby wave propagates toward the west relative to the mean flow, so the mean flow must be from the west to hold the wave steady relative to the Earth's surface. In such a case, the total wave number of the stationary wave is

$$K^2=\beta/[u]\equiv K_S^2. \tag{69}$$

Here the subscript S on K_S stands for "stationary."

With this preparation, we return now to the topographically forced case, and assume a stationary, neutral wave, i.e., $\omega=0$. For the case of no friction $(r=0)$, we find from (59) that

$$\hat{h}=\frac{\hat{h}_T}{\lambda^2\left(K^2-K_S^2\right)}. \tag{70}$$

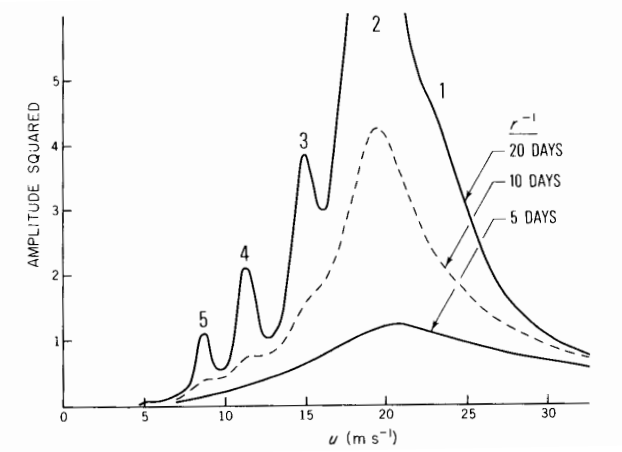

Figure 8.11: The mean-square height response, in the Charney-Eliassen model, as a function of $[u]$, for different values of the Rayleigh friction coefficient. The units are 10^4 m^2. The integers written above the curve indicate the values of $[u]$ for which particular zonal wave numbers resonate. From Held (1983).

Note that in the absence of friction the forced wave has "infinite amplitude" for $K^2=K_S^2$. This is the phenomenon of resonance, familiar from introductory physics. The infinity can be avoided by turning on the friction, i.e. allowing r to be positive. With friction, the amplitude of the stationary neutral wave is

$$\hat{h}=\frac{\hat{h}_T}{\lambda^2\left(K^2-K_S^2-i\dfrac{rK^2}{k[u]}\right)}. \tag{71}$$

Eq. (70) shows that, mathematically, friction makes the amplitude complex. Clearly the denominator of (70) cannot become zero, so long as . For $K^2=K_S^2$, (70) simplifies to

$$\hat{h} = \hat{h}_T \left(\frac{ik[u]^2}{\lambda^2 r \beta} \right).$$

(72)

Fig. 8.11 shows how the squared amplitude of the steady-state wave response varies as the zonal wind speed changes, for different values of r . Resonance occurs for a wind speed close to 20 m s^{-1}. Sufficiently strong damping leads to smooth behavior near resonance.

As seen in Fig. 8.12, the Charney-Eliassen model, despite its extreme simplicity, can explain reasonably well the observed zonal structure of the 500 mb height field at 45° N in January. This strongly suggests that the observed midlatitude stationary eddies in winter are forced primarily by topography. A similar conclusion was reached by Manabe and Terpstra (1974) through numerical experiments with a general circulation model.

Fig. 8.13 shows the observed time-latitude sections of the zonal and meridional contributions to the stationary wave kinetic energy, i.e. $\left[u^{*2}\right]$ and $\left[v^{*2}\right]$, respectively. In winter, the zonal component is strongest in middle latitudes, and we can think of this as

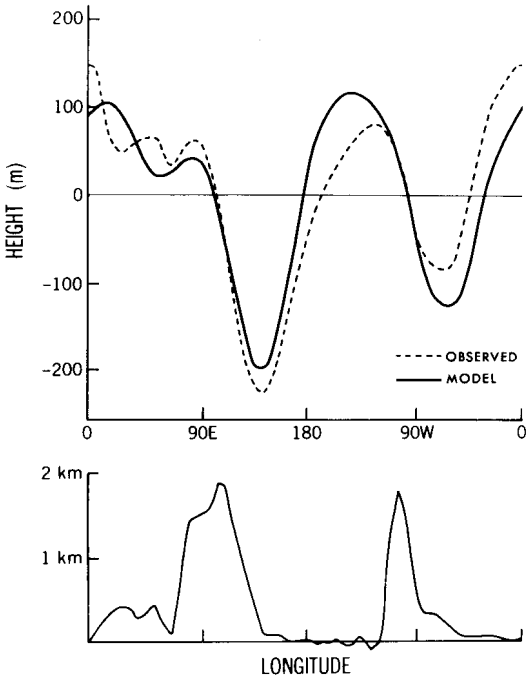

Figure 8.12: Upper panel: The height response as a function of longitude in the Charney-Eliassen model for days and (solid line) and the observed climatological 500 mb eddy heights at 45° N in January (dashed line). The lower panel shows the topography used. From Held (1983).

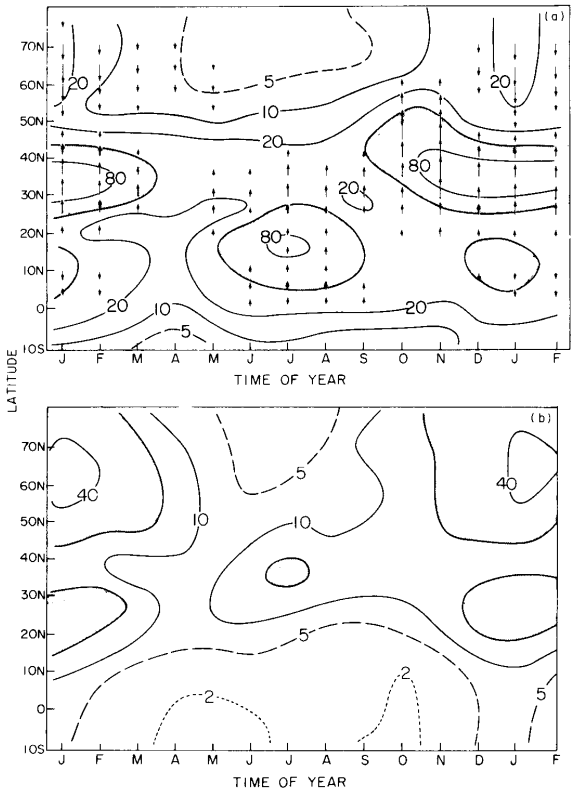

Figure 8.13: Time-latitude sections of Northern Hemisphere climatological mean stationary wave kinetic energy at the 200 mb level, based on data of Oort and Rasmusson (1971). a) Zonal wind component; b) meridional wind component, in units of $m^2\ s^{-2}$. Arrows in a) denote the direction and relative magnitude of the meridional flux of zonal momentum by the stationary waves. From Wallace (1983).

corresponding to the orographically forced waves analyzed above, although of course there is also a thermally forced component. In summer, there is a subtropical maximum of stationary eddy kinetic energy associated with the monsoons, which will be discussed later.

Tropical waves

In his doctoral thesis, Taroh Matsuno (1966) studied the linearized shallow water equations (see the *QuickStudy* on the shallow water equations) applied to an equatorial β-plane (Fig. 8.14). When Matsuno began this work, his motivation was to investigate to what extent near-equatorial motions are geostrophic. In the process, however, he discovered two new classes of tropical waves, which soon after were detected in the observations. These same waves are actually among the solutions found by Laplace, but this was not recognized until later (Lindzen, 1967). The model studied by Matsuno turns out to be relevant to a wide variety of phenomena, including monsoons, the Madden-Julian oscillation, the Quasi-Biennial Oscillation, and El Niño. This wide applicability is amazing, in view of the model's extreme simplicity.

Figure 8.14: A photograph of Prof. Taroh Matsuno, taken as he gave a lecture at UCLA in January 1998.

The shallow-water equations on an equatorial β-plane, linearized about a state of rest, are:

$$\frac{\partial u}{\partial t} - fv + g\frac{\partial h}{\partial x} = 0 ,$$

$$\frac{\partial v}{\partial t} + fu + g\frac{\partial h}{\partial y} = 0 ,$$

$$\frac{\partial h}{\partial t} + H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0 . \qquad (73)$$

Here $f \equiv \beta y$, where y is distance in the meridional direction, measured from $y = 0$ at the Equator (i.e., $y = a\varphi$), and $\beta \equiv \frac{df}{dy}$ is approximated by a constant value. Matsuno defined a time scale, $T \equiv \sqrt{\frac{1}{c\beta}}$, and a length scale, $L \equiv \sqrt{\frac{c}{\beta}}$. Here $c \equiv \sqrt{gH}$ is the phase speed of a pure gravity wave. With these length and time scales, the velocity scale is simply c. The length L can be interpreted as the "Equatorial radius of deformation." For $c = 10$ m s^{-1}, we find that $L = 1000$ km and $T \cong 1$ day. Non-dimensionalizing the governing equations using T and L, we obtain

$$\begin{aligned}
\frac{\partial u}{\partial t} - yv + \frac{\partial \phi}{\partial x} &= 0 \quad , \\
\frac{\partial v}{\partial t} + yu + \frac{\partial \phi}{\partial y} &= 0 \quad , \\
\frac{\partial \phi}{\partial t} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= 0 \quad .
\end{aligned} \tag{74}$$

Here ϕ is the non-dimensional form of gh.

As a side comment, we note that these equations can actually apply to a model with vertical structure (e.g., McCreary, 1981), and so are more readily applicable to the real atmosphere than one might guess. For example, consider a two-level model:

$$\begin{aligned}
\frac{\partial \mathbf{V}_1}{\partial t} + f\mathbf{k} \times \mathbf{V}_1 + \nabla \phi_1 &= 0 \quad , \\
\frac{\partial \mathbf{V}_3}{\partial t} + f\mathbf{k} \times \mathbf{V}_3 + \nabla \phi_3 &= 0 \quad , \\
\frac{\partial}{\partial t}\left(\phi_3 - \phi_1\right) + S\Delta p \omega_2 &= 0 \quad .
\end{aligned} \tag{75}$$

As shown in Fig. 8.17, subscript 1 denotes the upper level and subscript 3 denotes the lower level. The vertical velocity is defined in between, at level 2. We use $\Delta p \equiv p_3 - p_1$ to denote the pressure thickness between the two layers, and $S \equiv -\frac{\alpha}{\theta}\frac{\partial \theta}{\partial p}$ is the static stability of the basic state. Let

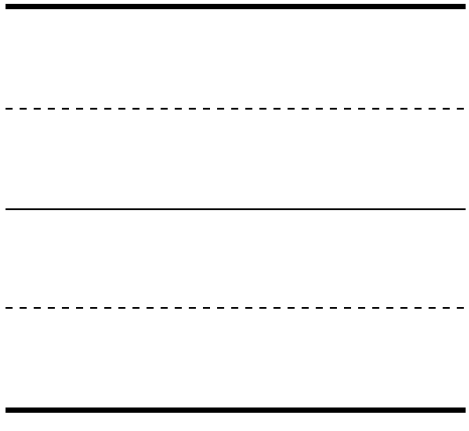

Figure 8.17: Schematic used to explain the two-level model represented by Eqs. (75).

$$\mathbf{V}_d \equiv \mathbf{V}_3 - \mathbf{V}_1 , \tag{76}$$

$$\phi_d \equiv \phi_3 - \phi_1 , \tag{77}$$

be the vertical shear (actually, difference) of the horizontal wind between the two layers, and the thickness between the two layers, respectively. Then (74) implies that

$$\frac{\partial \mathbf{V}_d}{\partial t} + f\mathbf{k} \times \mathbf{V}_d + \nabla \phi_d = 0 \,, \tag{78}$$

and

$$\frac{\partial \phi_d}{\partial t} + \frac{S\Delta p^2}{2} \nabla \cdot \mathbf{V}_d = 0 \,, \tag{79}$$

which are identical to the shallow water equations, and we can identify

$$c_i \equiv \Delta p \sqrt{\frac{S}{2}} \tag{80}$$

as the phase-speed of the internal gravity waves.

We return now to our discussion of (73). Assume solutions of the form

$$\begin{aligned} u &= \hat{u}(y) e^{i(kx+\sigma t)}, \\ v &= \hat{v}(y) e^{i(kx+\sigma t)}, \\ \phi &= \hat{\phi}(y) e^{i(kx+\sigma t)} \,. \end{aligned} \tag{81}$$

If we adopt the convention that k is positive, then $\sigma > 0$ corresponds to westward propagation, and $\omega < 0$ to eastward propagation. Substitution into (73) gives

$$\begin{aligned} i\sigma \hat{u} - y\hat{v} + ik\hat{\phi} &= 0, \\ i\sigma \hat{v} + y\hat{u} + \frac{d\hat{\phi}}{dy} &= 0, \\ i\sigma \hat{\phi} + ik\hat{u} + \frac{d\hat{v}}{dy} &= 0 \quad . \end{aligned} \tag{82}$$

We can solve the first of these equations for $\hat{u}$ in terms of $\hat{v}$ and $\hat{\phi}$, and the result can be used to eliminate $\hat{u}$ in the other two equations. Then the system (81) can be rewritten as

$$\begin{aligned}
\sigma\hat{u} &= -k\hat{\phi} - iy\hat{v}, \\
ky\hat{\phi} - \sigma\frac{d\hat{\phi}}{dy} &= i\left(\sigma^2 - y^2\right)\hat{v}, \\
\left(\sigma^2 - k^2\right)\hat{\phi} &= i\left(ky\hat{v} + \sigma\frac{d\hat{v}}{dy}\right) \quad .
\end{aligned}$$

(83)

Before considering the general case, we discuss a special solution, called the equatorial Kelvin wave, for which the meridional wind is identically zero. In that case, (83) reduces to

$$\begin{aligned}
\sigma\hat{u} &= -k\hat{\phi} \,, \\
ku\hat{\phi} - \sigma\frac{d\hat{\phi}}{dy} &= 0 \,, \\
\left(\sigma^2 - k^2\right)\hat{\phi} &= 0 \,.
\end{aligned}$$

(84)

The third of these implies that, for non-trivial solutions,

$$\sigma = \pm k \,.$$

(85)

With the use and inclusion of (85), our system to describe the Kelvin wave is simplified to

$$\begin{aligned}
\pm\hat{u} &= -\hat{\phi} \,, \\
-y\hat{\phi} \pm \frac{d\hat{\phi}}{dy} &= 0 \,, \\
\sigma &= \pm k \,.
\end{aligned}$$

(86)

The middle equation above obviously determines the meridional structure of ϕ. The solution is

$$\hat{\phi} = e^{\pm y^2/2} \,.$$

(87)

Here the plus and minus signs correspond to those used in (85). If we choose the plus sign, we get solutions that grow exponentially away from the Equator, which is unacceptable, especially since the Equatorial beta plane approximation is only useful near the Equator. We therefore choose the minus sign in (85)-(87), which gives a "bell-shaped" solution that has a maximum on the Equator and decays strongly away from the Equator. The solution for the equatorial Kelvin wave can now be written as

$$\hat{u} = e^{-y^2/2} ,$$
$$\hat{\phi} = e^{-y^2/2} ,$$
$$\sigma = -k .$$

(88)

Note that $\sigma = -k$ implies eastward propagation. An equatorial Kelvin wave always propagates towards the east.

Returning now to the general case, we can eliminate $\hat{\phi}$ between the second and third equations of (83) to obtain

$$ky\left(ky\hat{v} + \sigma \frac{d\hat{v}}{dy} \right) - \sigma \frac{d}{dy}\left(ky\hat{v} + \sigma \frac{d\hat{v}}{dy} \right) = \left(\sigma^2 - k^2 \right)\left(\sigma^2 - y^2 \right)\hat{v} ,$$

(89)

which can be simplified to

$$\boxed{\frac{d^2\hat{v}}{dy^2} + \left(\sigma^2 - k^2 + \frac{k}{\sigma} - y^2 \right)\hat{v} = 0} .$$

(90)

Because the substitution used to obtain (89)-(90) is only valid for $\sigma^2 - k^2 \neq 0$, (90) does not apply to the Kelvin wave, for which $\sigma^2 - k^2 = 0$.

We expect the solutions of (90) to have "wavy" behavior for $\sigma^2 - k^2 + \frac{k}{\sigma} - y^2 > 0$, and exponential behavior for $\sigma^2 - k^2 + \frac{k}{\sigma} - y^2 < 0$. Because of the $-y^2$ term, exponential behavior is guaranteed to emerge sufficiently far from the Equator, and, as with the Kelvin wave, we want this to be exponential decay rather than exponential growth. Therefore, as boundary conditions, we use

$$\hat{v} \to 0 \text{ as } y \to \pm\infty .$$

(91)

It can be shown that nontrivial solutions satisfying these boundary conditions exist when

$$\boxed{\sigma^2 - k^2 + \frac{k}{\sigma} = 2n + 1 \text{ for } n = 0,\ 1,\ 2\ldots} .$$

(92)

This is a dispersion equation. The expression on the right-hand side of (92) generates all positive odd integers, so that Eq. (92) is equivalent to the statement that $\sigma^2 - k^2 + \frac{k}{\sigma}$ is a positive odd integer. You can confirm by substitution that the solutions of (90) are

$$\hat{v}(y) = Ce^{-\frac{1}{2}y^2} H_n(y), \tag{93}$$

where C is an arbitrary real constant and $H_n(y)$ is the nth Hermite polynomial, which is given by

$$H_n(y) \equiv (-1)^n e^{y^2} \frac{d^n}{dy^n}\left(e^{-y^2}\right) \tag{94}$$

(see the QuickStudy on Hermite polynomials). Just as for the equatorial Kelvin wave, the factor $e^{-\frac{1}{2}y^2}$ ensures that these modes decay rapidly away from the Equator. For realistic values of the parameters, the e-folding distance is about 1000 km.

Because the y-dependence of each solution governed by (92) takes the form of a Hermite polynomial, and because the different Hermite polynomials are distinguished by their values of the parameter n, we can think of the meridional shapes of the Equatorial waves as being determined by the value of n. Different solutions with the same value of n will have the same meridional shape. Although (92) does not apply to the Kelvin wave, and although (92) is only valid for non-negative values of n, it turns out that $\sigma = -k$ (the dispersion relation for the Kelvin wave) is a solution of (92) if we set $n = -1$. For this reason, the Kelvin wave is often called the $n = -1$ solution of Matsuno's model.

The dispersion equation, (92), is cubic in σ, and so there are three σ s for each (k,n) pair. Two of these correspond to inertia-gravity waves. For large k, they can be approximated by

$$\sigma_{1,2} \cong \pm\sqrt{k^2 + 2n + 1}\,. \tag{95}$$

These expressions can be compared with (32). The third root corresponds to a Rossby wave. For large k, it can be approximated by

$$\sigma_3 \cong \frac{k}{\sqrt{k^2 + 2n + 1}}\,. \tag{96}$$

For the special case $n = 0$, the dispersion equation (92) can be *factored*, to give

$$(\sigma - k)\left(\sigma^2 + k\sigma - 1\right) = 0 . \tag{97}$$

Matsuno showed that for this case the three roots can be interpreted as follows:

$$\text{Eastward gravity wave: } \sigma_1 = -\frac{k}{2} - \sqrt{\left(\frac{k}{2}\right)^2 + 1} , \tag{98}$$

$$\text{Westward gravity wave: } \sigma_2 = \begin{cases} \sqrt{\left(\frac{k}{2}\right)^2 + 1} - \frac{k}{2} & \text{for } k \le \frac{1}{\sqrt{2}} \\ k & \text{for } k \ge \frac{1}{\sqrt{2}} \end{cases} , \tag{99}$$

$$\text{Rossby wave: } \sigma_3 = \begin{cases} k & \text{for } k \le \frac{1}{\sqrt{2}} \\ \sqrt{\left(\frac{k}{2}\right)^2 + 1} - \frac{k}{2} & \text{for } k \ge \frac{1}{\sqrt{2}} \end{cases} . \tag{100}$$

For reasons discussed just below Eq. (89), the root $\sigma = k$ has to be thrown out in (98) and (99). Notice that (with $n = 0$) the westward gravity wave and the Rossby wave are not really distinct. They coincide for $k = 1/\sqrt{2}$. Matsuno concluded, therefore, that for $n = 0$ there are actually only two waves: an eastward moving gravity wave, and a "mixed Rossby-gravity wave," which is also known as the Yanai wave[1]. The Yanai wave behaves like a gravity wave for $k < 1/\sqrt{2}$, and like a Rossby wave for $k > 1/\sqrt{2}$. The dispersion relation for the Yanai wave is

$$\sigma = \sqrt{\left(\frac{k}{2}\right)^2 + 1} - \frac{k}{2} . \tag{101}$$

Because $H_0(y) = 1$, Eq. (93) reduces to

[1] Michio Yanai and Taroh Matsuno shared an office in graduate school.

$$\hat{v}(y) = Ce^{-\frac{1}{2}y^2} \tag{102}$$

for the Yanai wave. This shows that in Yanai waves the meridional velocity has the same sign on both sides of the Equator and is a maximum on the Equator.

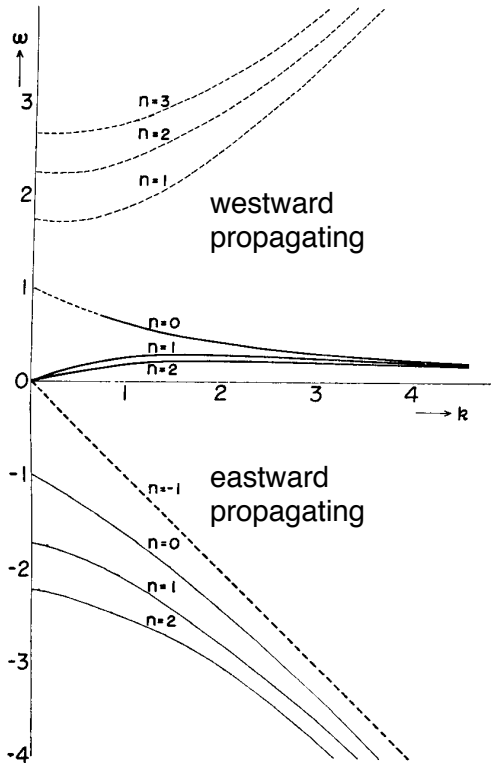

Figure 8.18: Frequencies as functions of wave number. Positive frequencies correspond to westward propagation. Thin solid line: eastward propagating inertia-gravity waves. Thin dashed line: westward propagating inertia-gravity waves. Thick solid line: Rossby (quasi-geostrophic) waves. Thick dashed line: The Kelvin wave. The westward moving wave with n=0 is the mixed Rossby-gravity or Yanai wave. It is denoted by a dashed line for, and by a solid line for. From Matsuno (1966).

The dispersion relations associated with the various-solutions of Matsuno's model are portrayed in Fig. 8.18. Recall that positive values of the frequency correspond to westward propagating waves, and negative values to eastward propagating waves. The thick solid curves arcing upward from the origin represent Rossby waves, with positive values of σ . The dashed curves in the upper part of the diagram correspond to westward propagating inertia-gravity waves, and the thin solid curves in the lower part of the diagram correspond to eastward propagating inertia gravity waves.

The westward propagating wave represented by the curve that is partly solid and partly dashed is the mixed Rossby-gravity wave, or Yanai wave. The dashed portion of this curve, plotted for $k < 1/\sqrt{2}$, represents those wave numbers for which the Yanai wave behaves like a westward propagating gravity wave. The solid portion of the curve, for $k > 1/\sqrt{2}$, represents those wave numbers for which the Yanai wave behaves like a Rossby wave.

The thick dashed line proceeding downwards towards the right from the origin represents the Kelvin wave.

For $n = 0$, the eastward moving inertia-gravity wave and westward moving Yanai wave

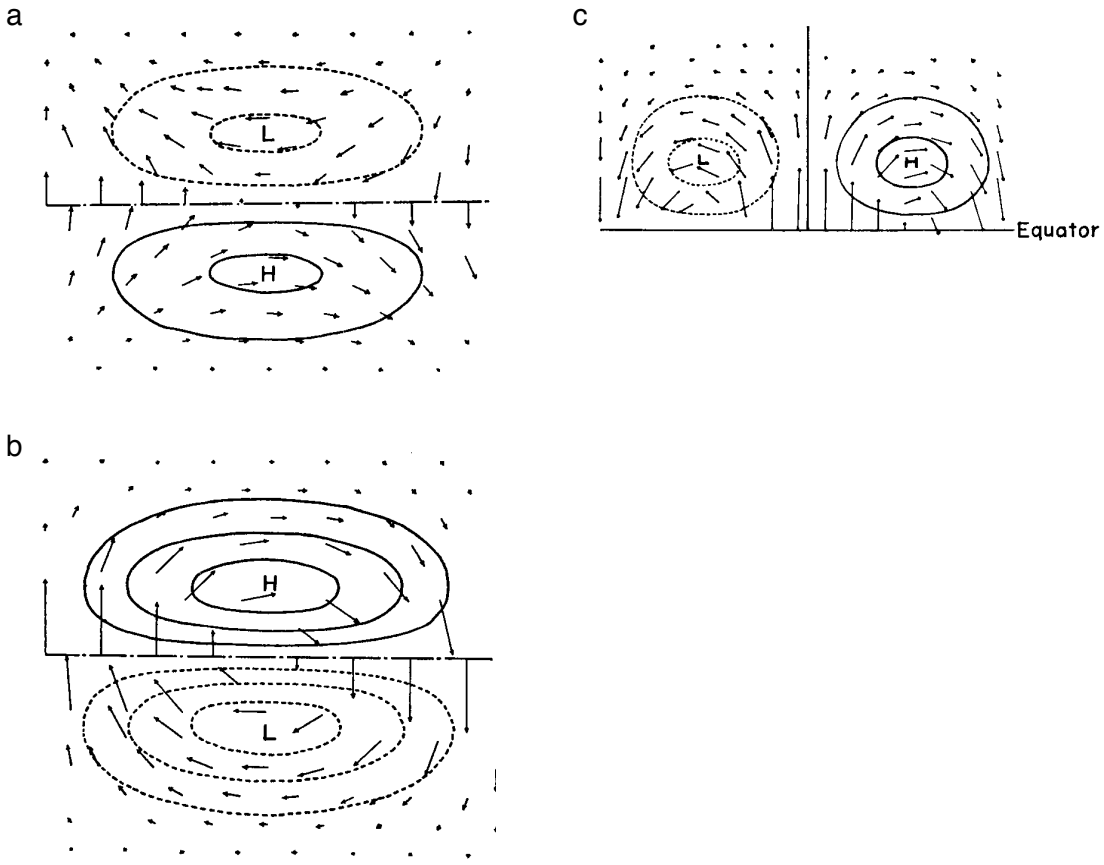

Figure 8.19: Pressure and velocity distributions of solutions for $n = 0$ and $k = 0.5$. a) Eastward moving inertia-gravity wave. b) Westward moving Yanai wave, which for this value of k behaves like an inertia-gravity wave. c) The structure of the Yanai wave for $n = 0$ and $k = 1$, in which case the Yanai wave acts like a Rossby wave. For each mode, v is a maximum on the Equator and does not pass through zero anywhere. This is characteristic of $n = 0$. From Matsuno (1966).

have the structures shown in the upper and middle panels of Fig. 8.19. For a pure gravity wave we expect the winds to be perpendicular to the isobars. When rotation is dominant, the winds are parallel to the isobars. The waves shown look like pure gravity waves near the Equator. For $n = 0$ and $k = 1$, the Yanai wave takes on the characteristics of a Rossby wave, as shown in the lower panel.

Solutions for $n = 1$ are shown on the left side of Fig. 8.20. The corresponding results for

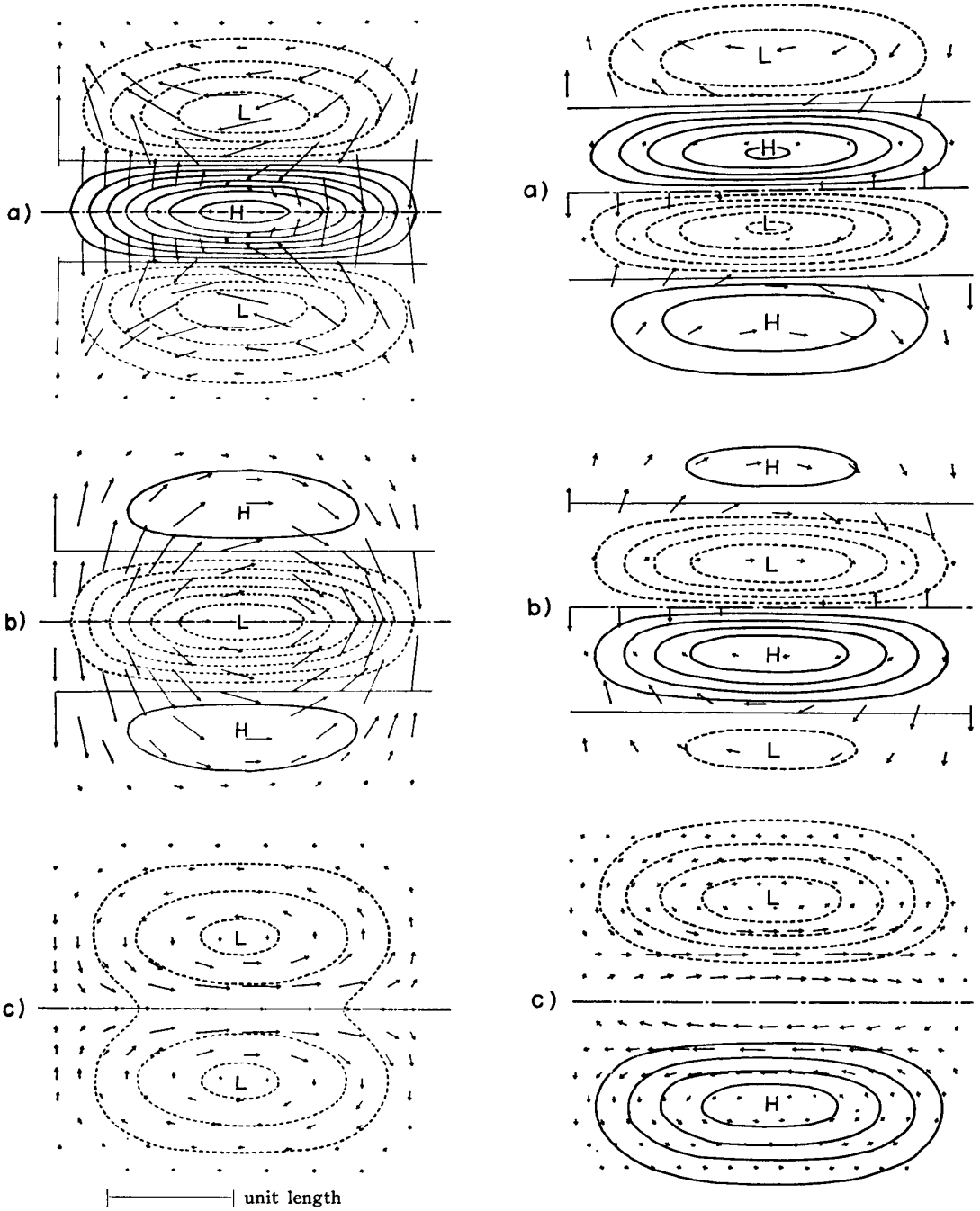

Figure 8.20: Left side: Pressure and velocity distributions of solutions for $n = 1$. a) Eastward propagating inertia-gravity wave. For each mode, $u = 0$ on the Equator, as we expect for $n = 1$. b) Westward propagating inertia-gravity wave. c) Rossby wave. Right side: Corresponding results for $n = 2$. For each mode, u is symmetrical across the Equation, as we expect for $n = 2$. From Matsuno (1966).

$n = 2$ are shown on the right side of the figure. Recall that the subscript n denotes the solution whose meridional structure is described by the n th Hermite polynomial. As can be seen in the figure, higher values of correspond to more nodes in the meridional direction.

The structure of the Kelvin wave is shown in Fig. 8.21. The velocity vectors are purely zonal, as expected, and the zonal wind is in phase with the pressure, as in a gravity wave.

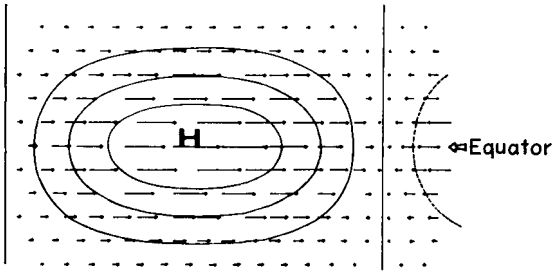

Figure 8.21: Pressure and velocity distributions for $n = -1$ and $k = 0.5$. This is the Kelvin wave. From Matsuno (1966).

The vertically continuous (i.e., non-shallow-water) version of Matsuno's analysis was presented by Lindzen (1967).

There have been many observational studies of Equatorial waves. Maruyama and Yanai (1966) observed the Yanai wave not long after Matsuno had predicted its existence, and Wallace and Kousky (1968) found the Kelvin wave soon thereafter. Matsuno's theoretical discoveries were confirmed by observations in remarkably short order!

Wheeler and Kiladis (1999) examined the space-time variability of the tropical outgoing long wave radiation.In Fig. 8.22, the data have been separated into modes that are symmetric

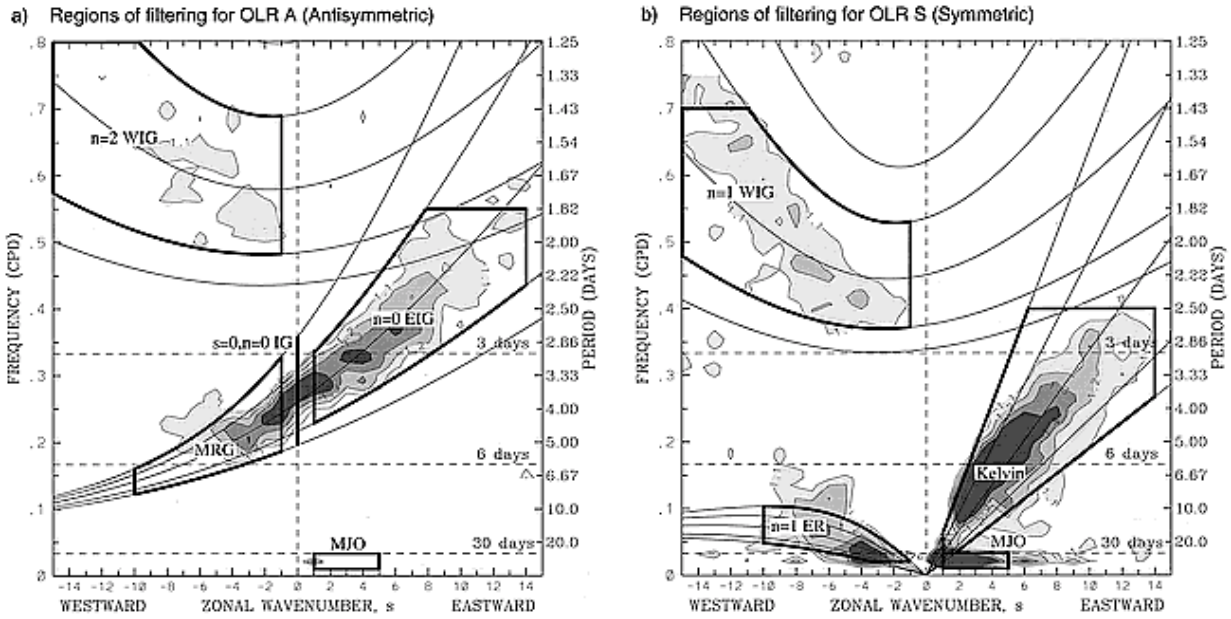

Figure 8.22: The variability of the outgoing longwave radiation (OLR) as a function of frequency and zonal wave number for modes that are symmetric across the Equator (right panel) and anti-symmetric (left panel). Eastward propagation is associated with positive wave numbers, and vice versa. The boxes select particular wave types. From Wheeler and Kiladis (1999).

across the Equator (right panel), such as the Kelvin wave, and modes that are anti-symmetric

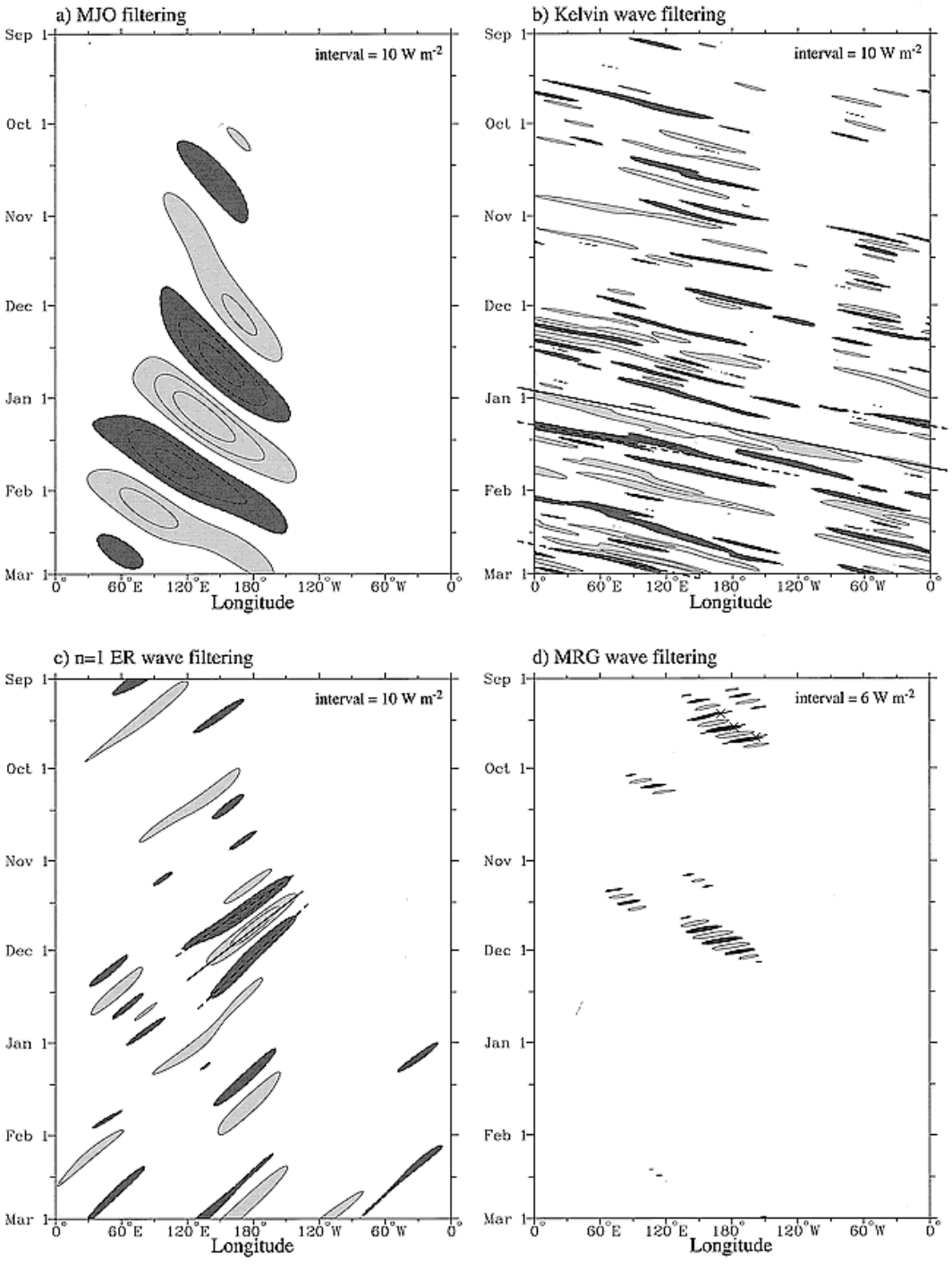

Figure 8.23: Longitudinal propagation of the Madden-Julian Oscillation (MJO; discussed later), Kelvin waves, equatorial Rossby (ER) waves, and mixed-Rossby-gravity (MRG) waves, as seen in the OLR. The zero contour has been omitted. The various modes are selected by including only the contributions from wave numbers and frequencies that fall in the corresponding boxes in Fig. (8.22). This is what is meant by "filtering." From Wheeler and Kiladis (1999).

across the Equator (left panel), such as the mixed-Rossby-gravity wave. Through the use of additional filtering procedures motivated by Matsuno's results, Wheeler and Kiladis were able to show the longitudinal propagation of various types of equatorially trapped disturbances (see Figs. 8.23 and 8.24.).

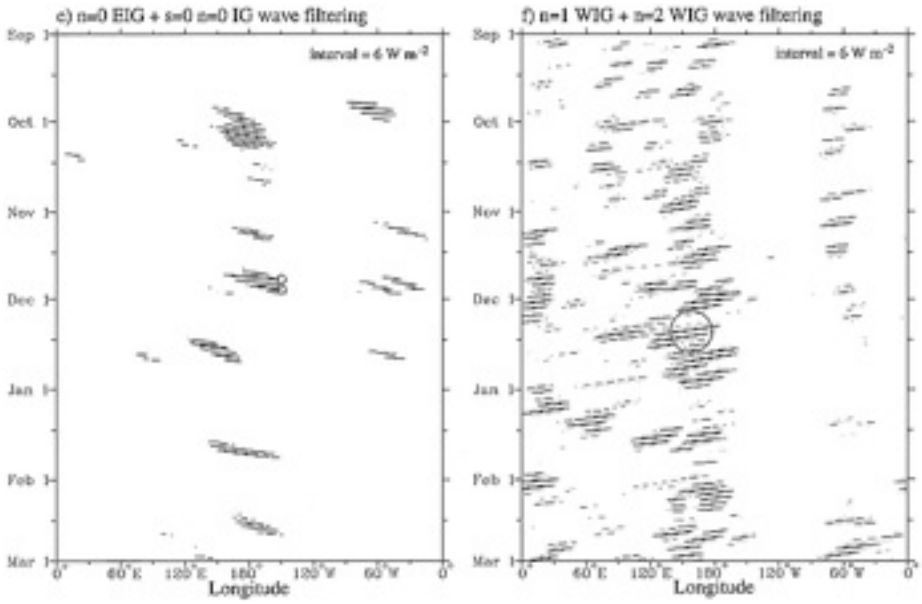

Figure 8.24: Longitudinal propagation of eastward- and westward-propagating inertia gravity waves, as seen in the OLR. The zero contour has been omitted. As in Fig. (8.23), the various modes are selected by including only the contributions from wave numbers and frequencies that fall in the corresponding boxes in Fig. (8.22). From Wheeler and Kiladis (1999).

The response of the tropical atmosphere to stationary heat sources and sinks

The discussion above is all about "free waves." Forced solutions to Matsuno's model are also of great importance. Fig. 8.25, which is taken from Matsuno (1966), shows the stationary circulation driven by a mass source and sink on the Equator. Think of this figure in terms of the low-level flow. The mass sink can be interpreted as a region of rising motion, where the air is converging at low levels, e.g., in the western equatorial Pacific. The mass source can be interpreted as a region of sinking motion, where the air is diverging at low levels, e.g., in the eastern equatorial Pacific. (Unfortunately the mass sink is plotted on the east side and the mass source is plotted on the west side, but this does not really matter because the solution is periodic in the zonal direction anyway.) The model predicts strong westerlies converging (from the west, of course) at low levels into the region of rising motion, and low-level easterlies converging on the east side of the region of low-level convergence. The easterlies can be interpreted as the trades, and as the lower branch of the Walker circulation. The westerlies can be interpreted as a "monsoon-like" westerly inflow to a region of heating. Further discussion is given later.

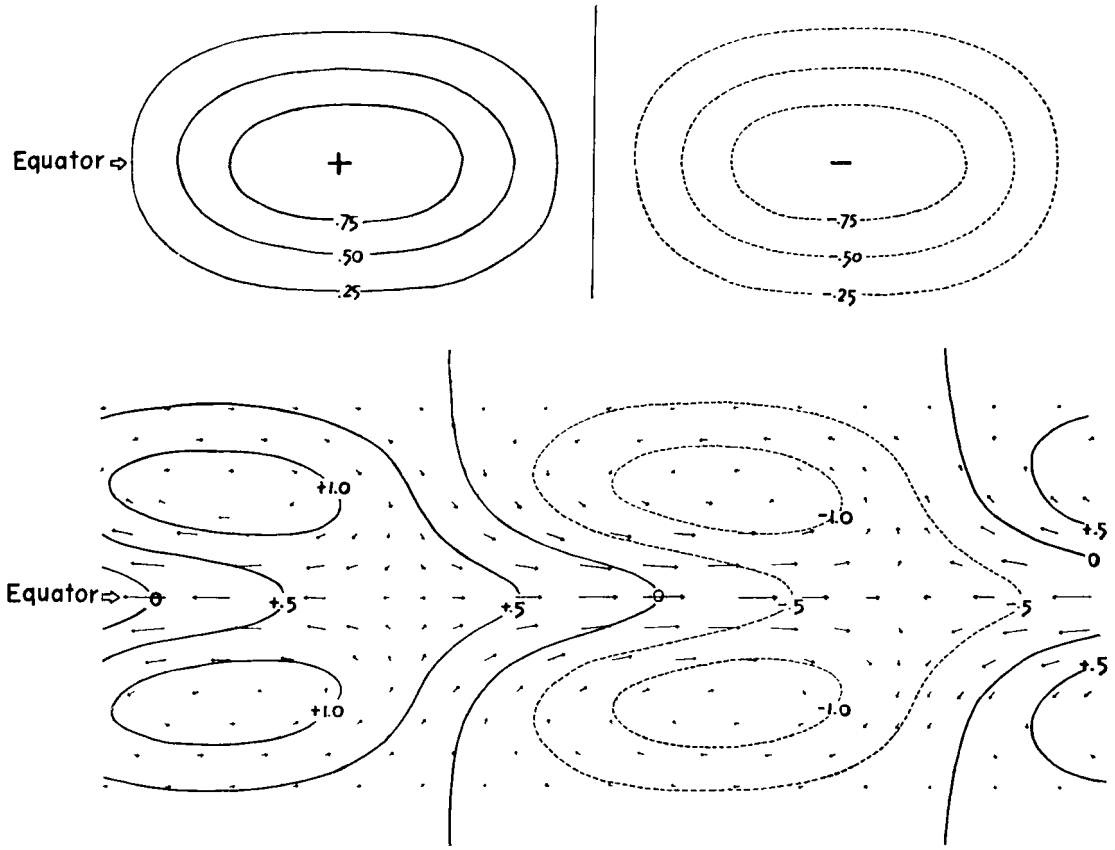

Figure 8.25: Stationary circulation pattern (lower panel) forced by the mass source and sink shown in the upper panel. From Matsuno (1966).

Webster (1972) and Gill (1980) followed Matsuno's lead by developing simple analytic models of the response of a resting tropical atmosphere to heat sources and sinks. Since much of the convective heating in the tropics is confined over three relatively small land regions (Africa, South America, and the Indonesian region), Gill examined the atmospheric response to a relatively small-scale heating source that is centered on the equator. If the atmosphere is abruptly heated at some initial time, Kelvin waves propagate rapidly eastward and generate easterly trade winds to the east of the heating. Thus the easterly trade winds in the Pacific could result from Kelvin waves produced by convective heating over Indonesia. Similarly, Rossby waves propagate westward and generate westerlies to the west of the heating. Because the fastest Rossby wave travels at only one-third the speed of the Kelvin wave, the effects of the Rossby waves would be expected to reach only one-third as far those as those of the Kelvin wave. Gill interpreted the westerlies over the Indian ocean as a response to Rossby waves generated by convective heating over Indonesia.

Gill (1980) studied what amounts to a steady-state version of Matsuno's model, and introduced forcing in the form of mass sources and sinks, along with very simple damping. Corresponding to (73), we have

$$\varepsilon u - yv + \frac{\partial \phi}{\partial x} = 0 \,,$$
$$yu + \frac{\partial \phi}{\partial y} = 0 \,,$$
$$\varepsilon \phi + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = -Q \,.$$

(103)

and, as a purely diagnostic relation,

$$w = \varepsilon \phi + Q \,.$$

(104)

The wind components u and v represent the lower-tropospheric variables. In (102) and (103), ε^{-1} is a dissipation time scale, and Q is a "heating rate" that must be specified. The variables ϕ, w, and Q are defined in the middle troposphere. Gill included dissipation in the form of Rayleigh friction and Newtonian cooling, and for simplicity assumed that the time scales, given by ε^{-1}, are equal. Rayleigh friction is a simple parameterization of friction in which the velocity is divided by a frictional time scale. The friction term is neglected in the meridional momentum equation of (102); see Gill (1980) for an explanation.

Gill focused primarily on cases for which the heating is symmetric or anti-symmetric about the equator. The solution for symmetric heating resembles a Walker circulation, with lower-tropospheric inflow into the heating region and upper-tropospheric outflow. The Walker Circulation is discussed in detail later in this Chapter. The surface easterlies cover a larger area than the surface westerlies because the phase speed of the eastward-propagating Kelvin wave is three times faster than that of the westward-moving Rossby wave.

By using (103) to form a vorticity equation for the case of no damping, and then substituting from the continuity equation, Gill found that

$$v = yQ \,.$$

(105)

This is closely related to what is sometimes called "Sverdrup balance," in which the "meridional advection of the Coriolis parameter," i.e., the so-called β -term of the vorticity equation, is balanced by the divergence term, which is represented by the heating rate on the right-hand side of (105). According to (105), v changes sign across the equator in a region where Q has a single sign. In a region of cooling ($Q < 0$), the flow is towards the Equator on both sides, and in a region of heating it is away from the Equator on both sides.

For $Q > 0$, (104) implies poleward motion in the lower layer and equatorward motion in the upper layer. This suggests that in regions of heating, e.g. the western Pacific, the Walker circulation produces a north-south circulation that opposes the Hadley circulation. Geisler (1981) found the same result. For $Q < 0$, the low-level motion is Equatorward; this is what we see in the subtropical highs, e.g,. in the eastern Pacific.

The solution for anti-symmetric heating consists of a mixed Rossby-gravity wave and a Rossby wave. There is no Kelvin-wave response because the Kelvin wave is intrinsically symmetric across the Equator. Long mixed Rossby-gravity waves do not propagate, and so the response of this wave type is largely confined to the region of heating. Due to the westward propagation of Rossby waves, no response is generated to the east of the forcing region. To the west, the region of westerly flow into the heating region is limited because the Rossby modes

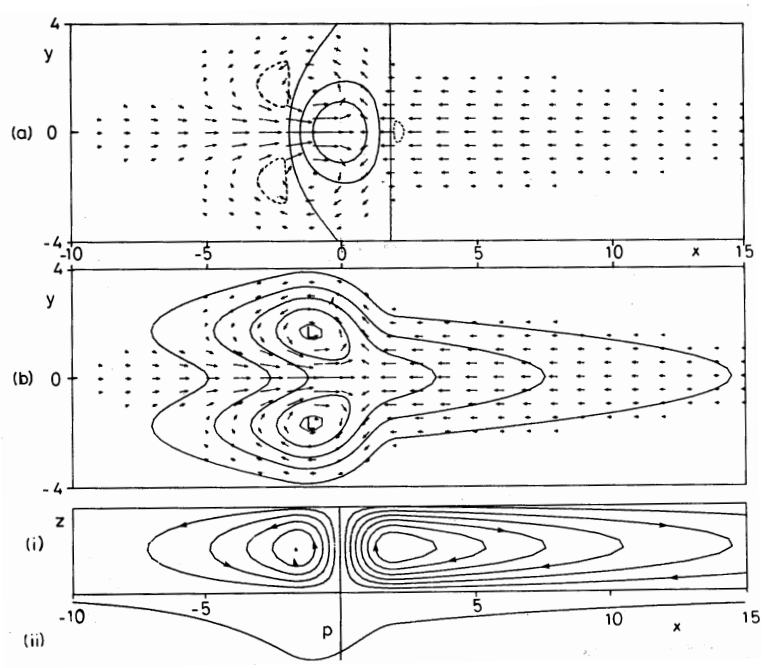

Figure 8.26: Solution of Gill's model for the case of heating symmetric about the Equator. The upper panel shows the heating field and the low-level wind field. The center panel shows the perturbation pressure field, which features low pressure along the Equator generally, with twin cyclones slightly off the Equator. The bottom panel shows the implied vertical motion and the zonal variation of the pressure along the Equator. From Gill (1980).

travel slowly, and so are dissipated before they can propagate far to the west.

Gill interpreted the symmetric case as a simulation of the Walker circulation, and the asymmetric case as a simulation of the Hadley circulation.

For heating centered on the Equator, as in Fig. 8.26, Gill found strong westerlies on the west side, and strong easterlies on the east side, combining to give strong zonal convergence on the heating. The westerlies can be interpreted as the time-averaged response to westward -

propagating Rossby waves excited by the heating, and the easterlies can be interpreted as the time-averaged response to eastward-propagating Kelvin waves excited by the heating. This implies a Walker circulation, as indicated in Fig. 8.26, and a surface pressure field with a minimum pressure slightly to the west of the heating.

When the heating is antisymmetric across the Equator, as in Fig. 8.27, the model produces something like a Hadley circulation, with a low-level cyclonic circulation on the side

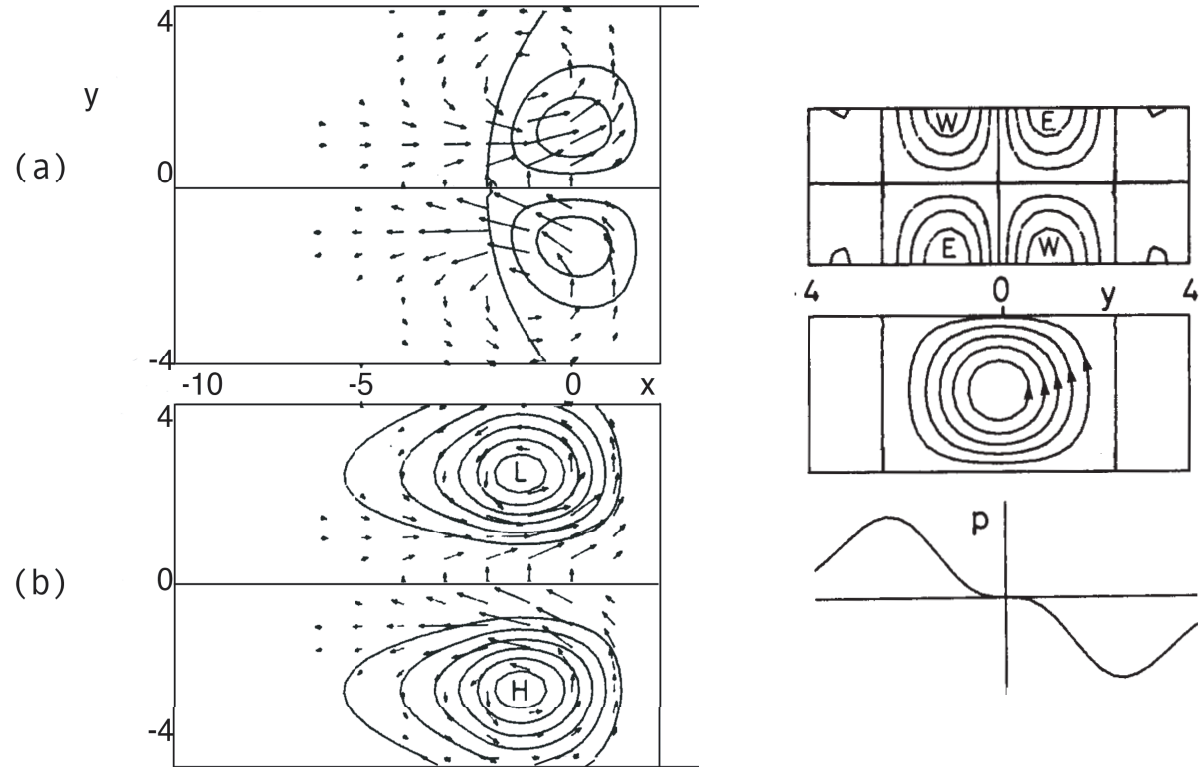

Figure 8.27: The response to antisymmetric heating. On the left side, the panel a) shows contours of the mid-level vertical velocity superimposed on the horizontal wind vectors for the lower layer. Panel b) shows contours of the perturbation surface pressure, again with the lower-layer horizontal wind field superimposed. The right-hand panels show the zonally integrated solution corresponding to the results in the left-hand panels. The upper right-hand panel shows the latitude-height distributions of the zonal velocity and the stream function of the mean meridional circulation, as well as the meridional profile of the surface pressure. From Gill (1980).

with positive heating, and a low-level anticyclone on the other side.

When the symmetric and antisymmetric heatings are combined, as in Fig. 8.28, the model produces a circulation that looks remarkably similar to that of the Asian summer monsoon, as discussed in the next subsection.

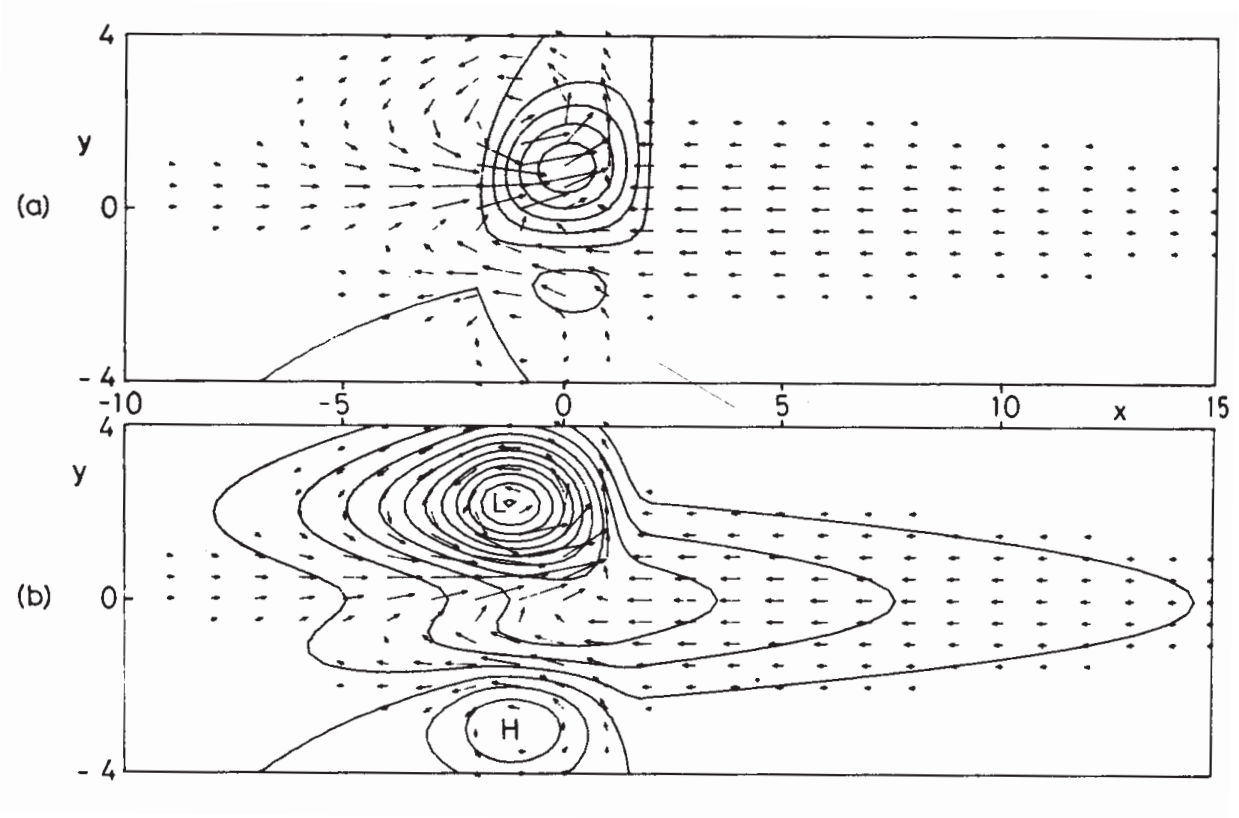

Figure 8.28: The response of Gill's model to a combination of symmetric and antisymmetric heating. From Gill (1980). Panel a) shows the heating, and panel b) shows the surface pressure. The low-level winds are shown in both panels.

Although Gill plausibly demonstrated that heating of limited extent generates tropical waves that produce broad wind and pressure fields resembling the observations, his results must be viewed with caution due to several limitations. First, results were generated for a specified rather than predicted heating of the tropical troposphere, and so ocean-atmosphere interactions and feedbacks involving moist convection were excluded. Second, the model includes neither a moisture budget nor cloud radiative effects. Last and most important, the model was linearized about a resting atmosphere. Although linearization is useful, Gill's results must be interpreted as the response of the tropical atmosphere to a perturbation of a basic state, not as a prediction of the basic-state climate. Gill demonstrated the sensitivity of the tropical troposphere to the spatial distribution of heating, but his study does not really address the basic-state climate.

Monsoons

Monsoons occur in many parts of the world. They can be viewed as thermally forced stationary planetary waves associated with land-sea contrast. The most spectacular monsoon on Earth is the one associated with Asia, our largest continent.

The Asian monsoon can be defined in a number of ways. By definition, a monsoon is a dramatic seasonal reversal of the low-level prevailing winds (Lighthill and Pearce, 1981). In the

dramatic case of the Asian monsoon, the winds near the surface reverse from the northeast in winter to the southwest in summer, as seen in Fig. 8.29. The 15 m s^{-1} low-level southwest wind

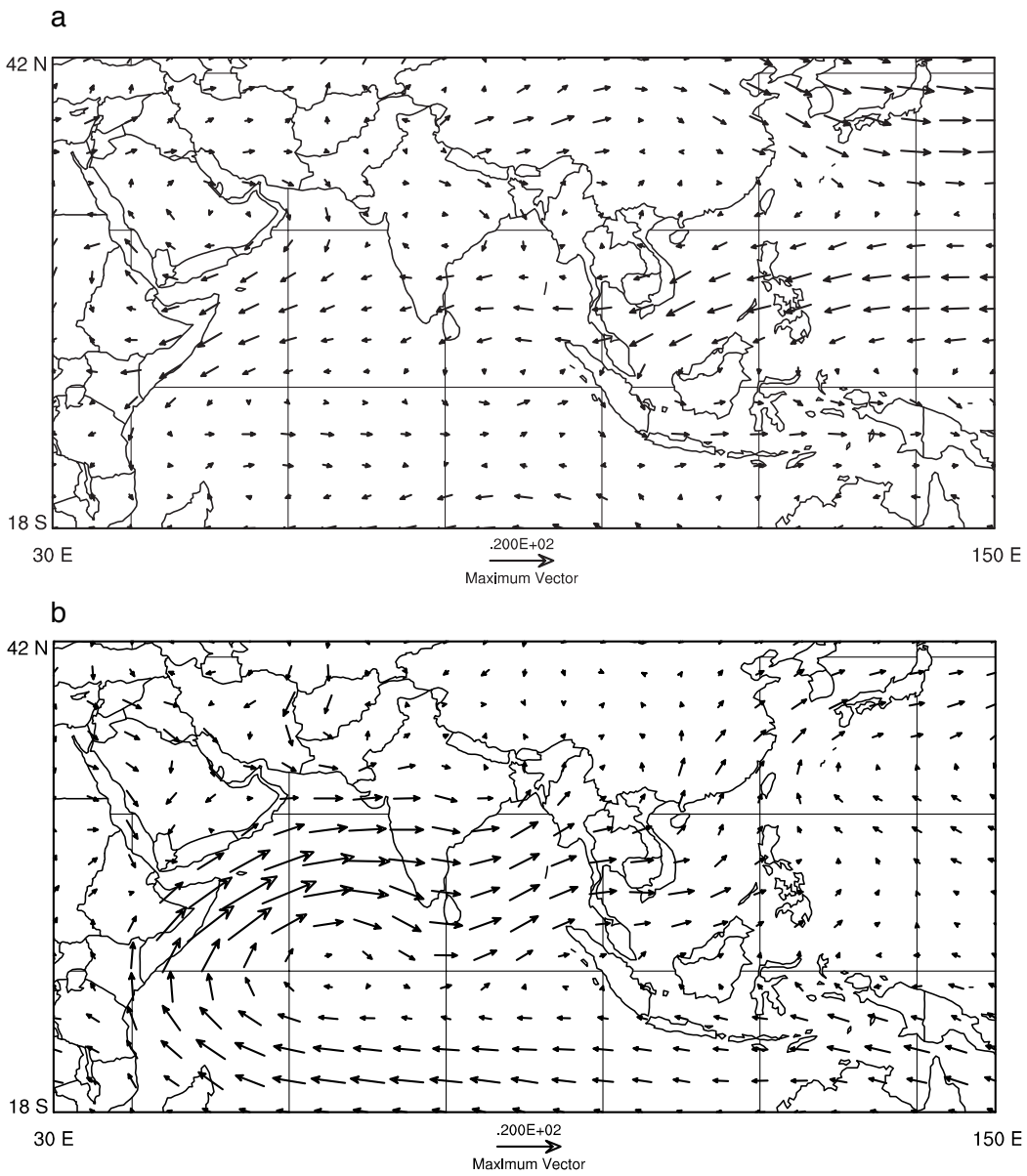

Figure 8.29: Observed 850 mb wind vectors for a) January, and b) July.

that crosses the Equator and flows from the coast of east Africa to the shores of India is known as the Somali jet. It is one of the strongest low-level jets in the world. There are other seasonal changes of wind in the world, but none have the geographical scope or the socioeconomic impact of the Asian monsoon.

The basic trigger for the monsoon is a contrast between the temperature of the Asian continent and that of the surrounding ocean. The thermal anomaly in the middle troposphere is enhanced by the spectacular topography of the Tibetan Plateau (Fig. 8.30), which extends

upward to about the 500 mb level. Much of the "surface" heating associated with the summer monsoon actually occurs in the middle troposphere, because it is located on the Tibetan Plateau. The Tibetan Plateau towers above the surrounding land surface, with average elevations over the central Plateau of over 3000 m. The observed JJA mean 500 mb temperature for the monsoon

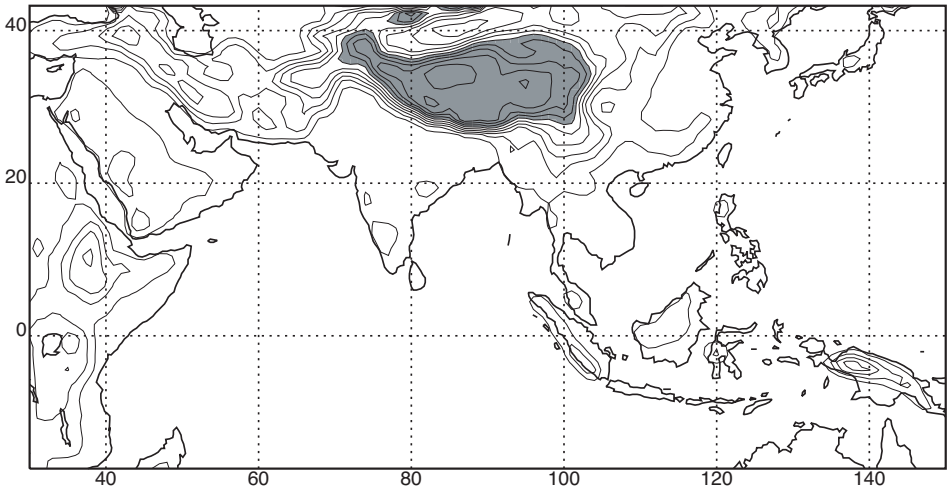

Figure 8.30: Average elevation of the monsoon region. Data was averaged to 1º x 1º, and then 9-point smoothed. Terrain over 3000 m high is shaded.

region is shown in Fig. 8.31. An island of warm air is centered over the Tibetan plateau.

After the non-permanent snow on the Plateau has melted in late spring and early summer,

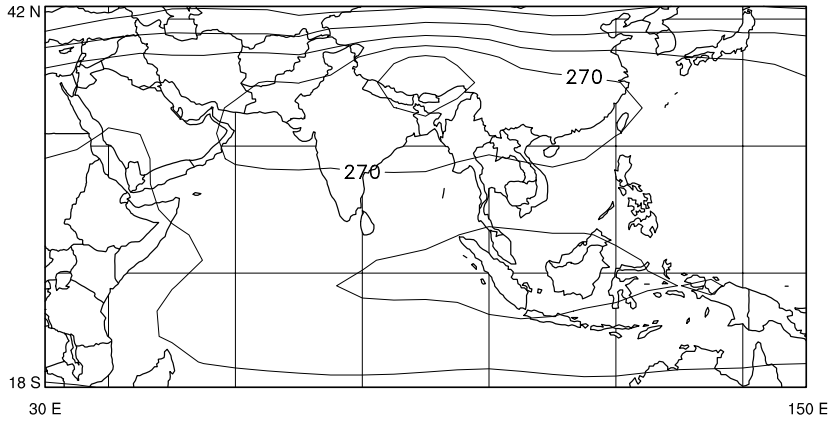

Figure 8.31: Observed JJA climatological 500 mb temperatures. Contour interval is 2 K.

the surface and the air above it are heated to a temperature higher than that of the surrounding atmosphere. Rising motion balances this heating, and this forces convergence in the lower and middle troposphere, and compensating divergence aloft. As the seasonal heating builds, a trough forms over southern India in late May, and subsequently moves north and west. In some cases, the progression of the monsoon trough is expedited by the passage of a tropical or extratropical

cyclone to the north (Mooley and Shukla, 1987). Cloudiness and precipitation begin to increase at the southern tip of India in late May. Nearly the entire country has begun to receive monsoon precipitation by the end of June.

The onset of the summer monsoon brings cooler surface temperatures to India and other areas that receive monsoonal precipitation, due to the increase in clouds as well as the increase in soil moisture that accompanies the precipitation. The data shows, as one might expect, that the lowest surface air temperatures in the monsoon region occur at the highest elevations. There is a large area of low (less than 280 K) surface air temperature on the Tibetan Plateau, and a large area of high (greater than 305 K) surface air temperatures on the Arabian Peninsula.

From an agricultural standpoint, the beginning of the precipitation associated with the Asian summer monsoon is probably one of the most anticipated events in the world. The onset of the monsoon is generally defined as the beginning of consistent rainfall of the monsoon season. Fig. 8.32 shows a plot of the dates of onset of the Asian monsoon from Krishnamurti et al. (1990). The Ganges valley receives copious amounts of rain from monsoon depressions that form in the Bay of Bengal and propagate northward and westward. An example is shown in Fig. 8.33 a. Many areas within the monsoon region also receive significant amounts of precipitation from tropical cyclones. Fig. 8.33 b shows the daily precipitation totals at an average of several stations on the southwest coast of India, and gives a sense as to the observed intraseasonal variability of monsoon precipitation. Much of the precipitation of the Asian monsoon is forced by southwest winds flowing over the western shores of India and Southeast Asia, as well as the foothills of the Tibetan Plateau (Johnson and Houze, 1987).

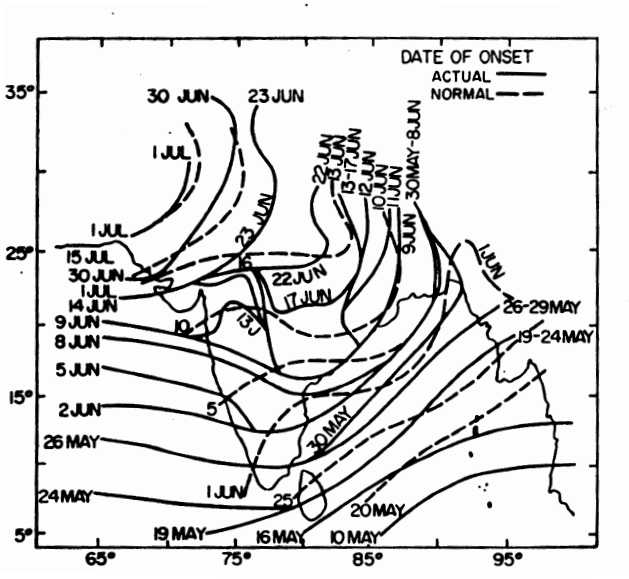

Figure 8.32: Dates of onset of the Asian monsoon near India in 1988 (actual) and mean (normal). From Krishnamurti *et al.* (1990).

Fig. 8.34 shows JJA means of the precipitation across the monsoon region, according to the climatological data set of Legates and Wilmott (1990). There are two major precipitation maxima. One is west of the southwest coast of India, and the other is west of southern Myanmar. Both of these areas receive strong onshore flow, showing that the ocean is the source of the moisture. Minima occur near Sri Lanka and the east coast of Vietnam. These areas appear to be in orographic rain shadows. The northern and western parts of the monsoon region are quite dry, receiving less than 2 mm day-1 of rain. The observed 850 mb wind analysis shows that these areas do not receive much moisture from the Indian Ocean during JJA. Only a small portion of the monsoon region receives more than 20 mm day-1 of precipitation.

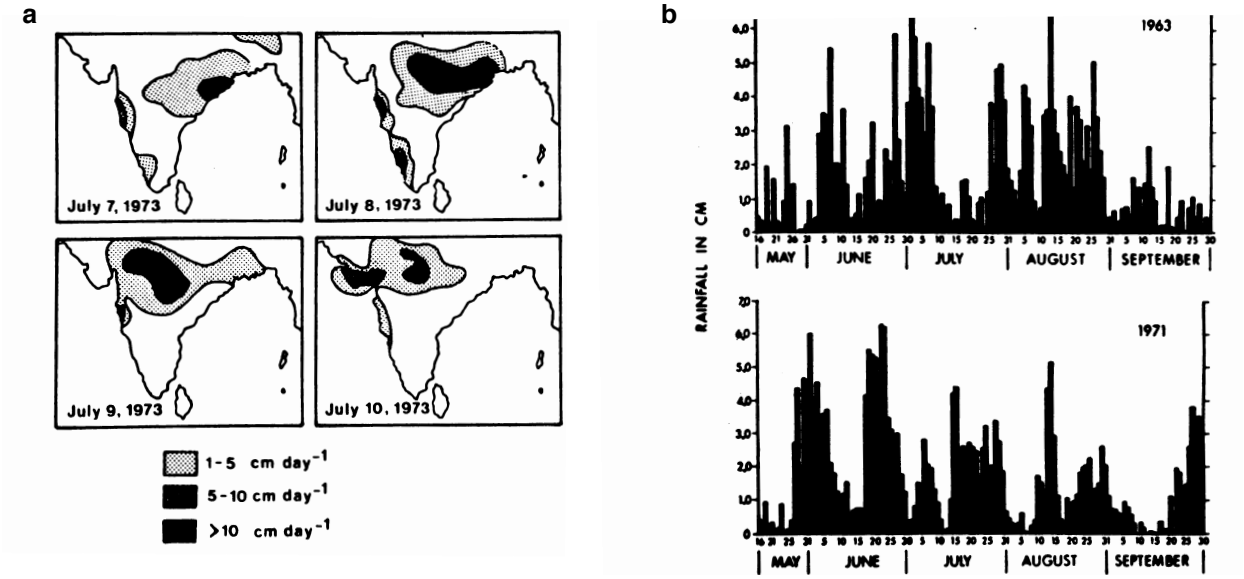

Figure 8.33: a) Progress of a monsoon depression across India. Many such depressions occur throughout the summer monsoon. b) Observed daily rainfall along the southwest coast of India for the summer monsoon seasons of 1963 and 1971. From Webster (1987).

The precipitation in much of the monsoon region varies on many time scales. The individual disturbances that cause the precipitation associated with the Asian monsoon last only a few days at any single location. However, there are also prominent variations known as "breaks," which occur at periods of approximately 10-20 days and 40-50 days (Webster, 1987). The 10-20 day variation is related to the periodic northward propagation of the Intertropical Convergence

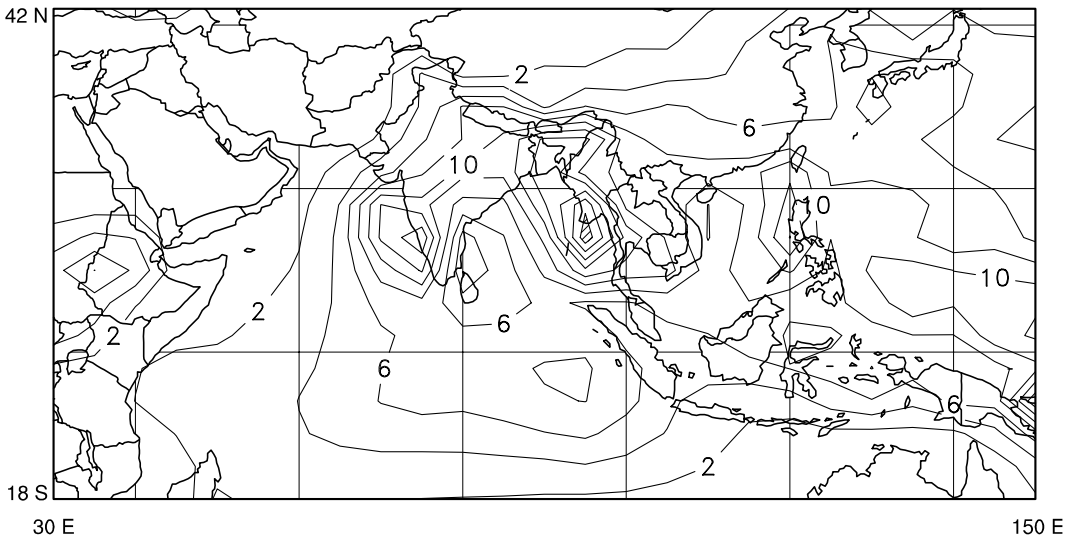

Figure 8.34: The observed JJA climatological precipitation (Legates and Willmott, 1990). The contour interval is 2 mm day^{-1}.

Zone (ITCZ), which begins near the Equator, and progresses to the foothills of the Tibetan Plateau in roughly 15 days, as seen in Fig. 8.35. The ITCZ usually reforms in the south after it has progressed to the foothills of the Tibetan Plateau, but occasionally it stays near the Plateau for an "extended break" period of 40-50 days. These extended break periods have been linked by some (Webster, 1987) to the 40-50 day oscillation discovered by Madden and Julian (1972),

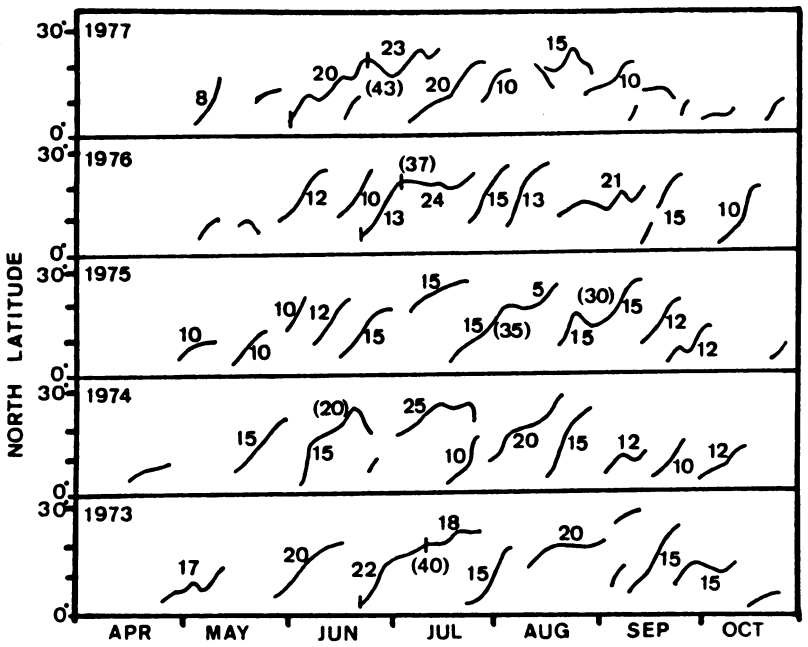

Figure 8.35: Mean latitudinal position of the monsoon trough in the Indian Ocean for the summers of 1973-1977, as obtained from the maximum cloudiness zone and the 700 mb trough. Numbers refer to longevity of a particular cloudiness zone, with extended break periods indicated by parentheses. From Webster (1987), Webster (1983), and Sikka and Gadgil (1980).

which is discussed in detail later.

One of the most prominent signatures of the Asian monsoon is the monsoon trough, which extends from Bangladesh to the Arabian Peninsula. The sea level pressure field is dominated by the monsoon trough, and an area of high pressure over the Tibetan Plateau.

The heating of the middle troposphere by the Tibetan plateau induces convergence in the middle and lower troposphere (Yanai, et al., 1992), largely through the Somali jet. To balance the convergence at the lower levels of the atmosphere, there must be large-scale rising motion, and divergence aloft. The observed JJA mean 500 mb vertical velocity for the monsoon region is shown in Fig. 8.36. The strongest areas of rising motion are over southwestern India and the Bay of Bengal. These areas both receive copious amounts of precipitation in JJA. Two of the strongest areas of sinking motion are over the eastern Mediterranean Sea and northern China.

Both areas are quite dry in JJA, receiving less than 2 mm day^{-1} of rain. The area mean vertical

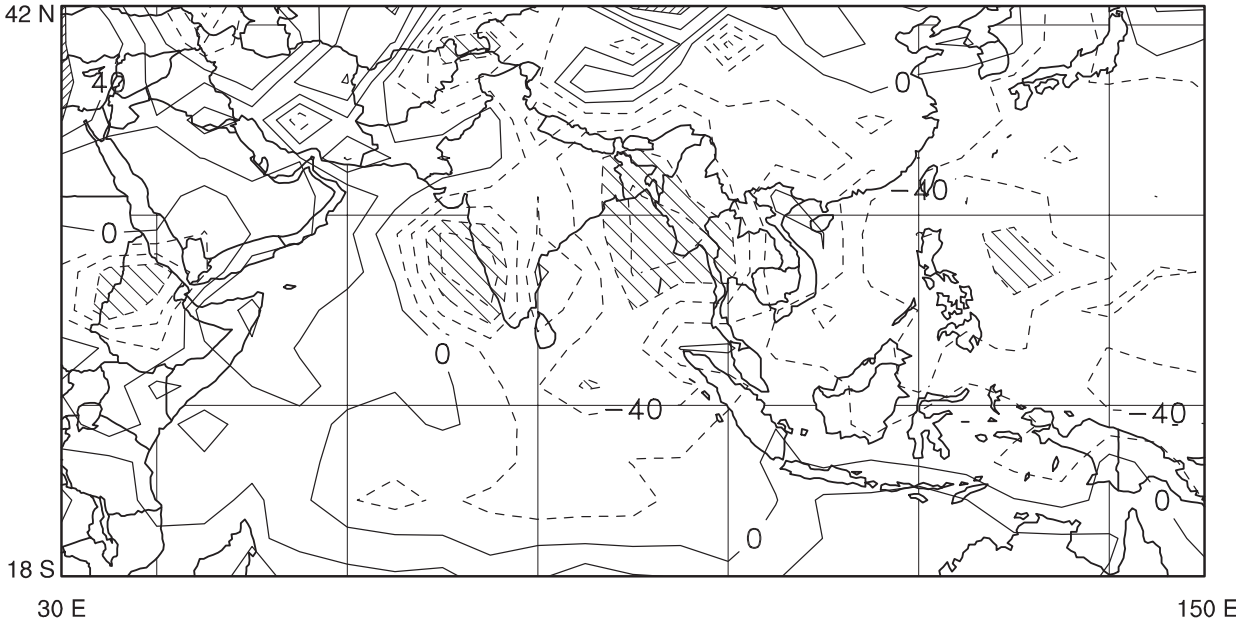

Figure 8.36: Observed JJA climatological 500 mb vertical velocity. Contour interval is 20 mb day^{-1}. Areas with vertical velocities more negative than -60 mb day^{-1} have light shading.

velocity is upward, at -10 mb day^{-1}.

The upper-level divergence is associated with a broad, strong anticyclonic circulation over the Plateau at 200 mb, as shown in Fig. 8.37. The strong, upper-level easterlies () are consistent with the thermal wind relationship, because the lower and middle tropospheric temperatures actually increase towards the north in the Northern Hemisphere (Yanai et al., 1992; Murakami, 1987; Yanai and Li, 1993). There is a slight northerly component to the 200 mb winds at the Equator, especially on the eastern side of the region. The winds shift abruptly from easterlies to westerlies at the northern and southern fringes of the monsoon region, with westerlies of up to 30 m s^{-1} at about 35 °N. Fig. 8.38 shows a latitude-pressure cross-section of the observed zonal wind at 77.5° E.

From the above discussion, it is clear that we can interpret the monsoon as a zonally localized Hadley circulation (Webster 1987), with a meridional low-level branch that moves towards a warm, convectively active ascending branch, and an upper-level return flow that feeds a relatively cool descending branch. In fact, in the northern summer the Hadley circulation is more or less contained within the longitudes of the monsoon region. The monsoon is thus a direct circulation, which converts potential energy to kinetic energy.

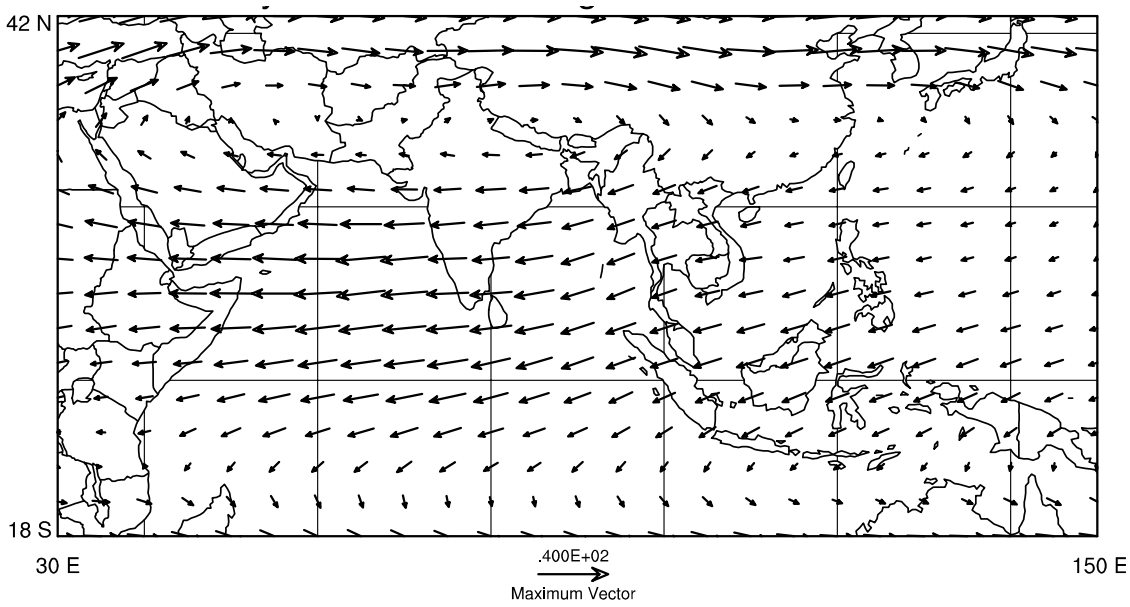

Figure 8.37: JJA climatological 200 mb winds. The scale vector is 50 m s^{-1}.

The left-hand column of Fig. 8.39 shows the hourly record of surface pressure and its Fourier spectrum at 10°N, 77.5°E, near the southern tip of India, at 18°N, 77.5°E, and at 26°N, 77.5°E in northern India. All three locations show a large peak at 12 hours, which corresponds to

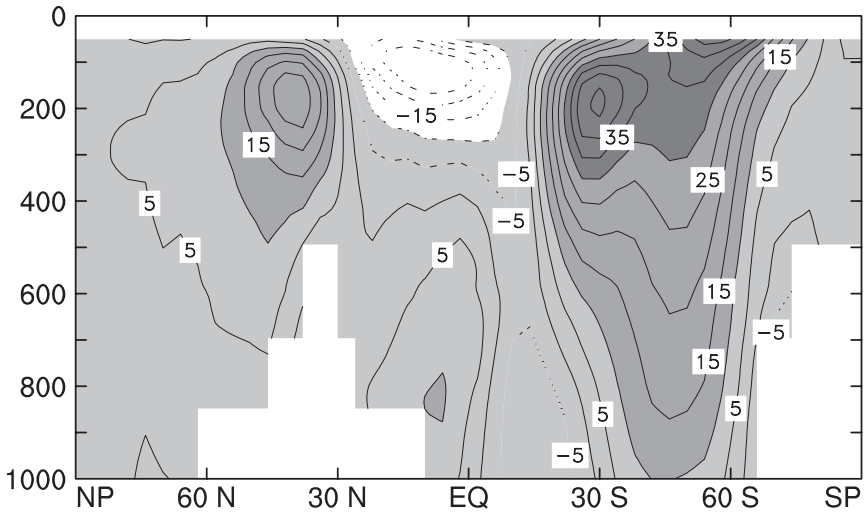

Figure 8.38: Latitude-pressure plot of the JJA climatological zonal winds at 77.5 ºE. The contour interval is 5 m s^{-1}.

the semidiurnal tide. There is also a noticeable diurnal tide. The right-hand column of Fig. 8.39 shows similar observations for the same three locations, but for longer periods. Beyond 24 hours, all three locations show oscillations with time scales of a few days, which is the time scale of

synoptic weather systems. There is also a spectral peak at 10-20 days, which may be associated with the period of time it takes for the monsoon trough to move from the southern tip of India to

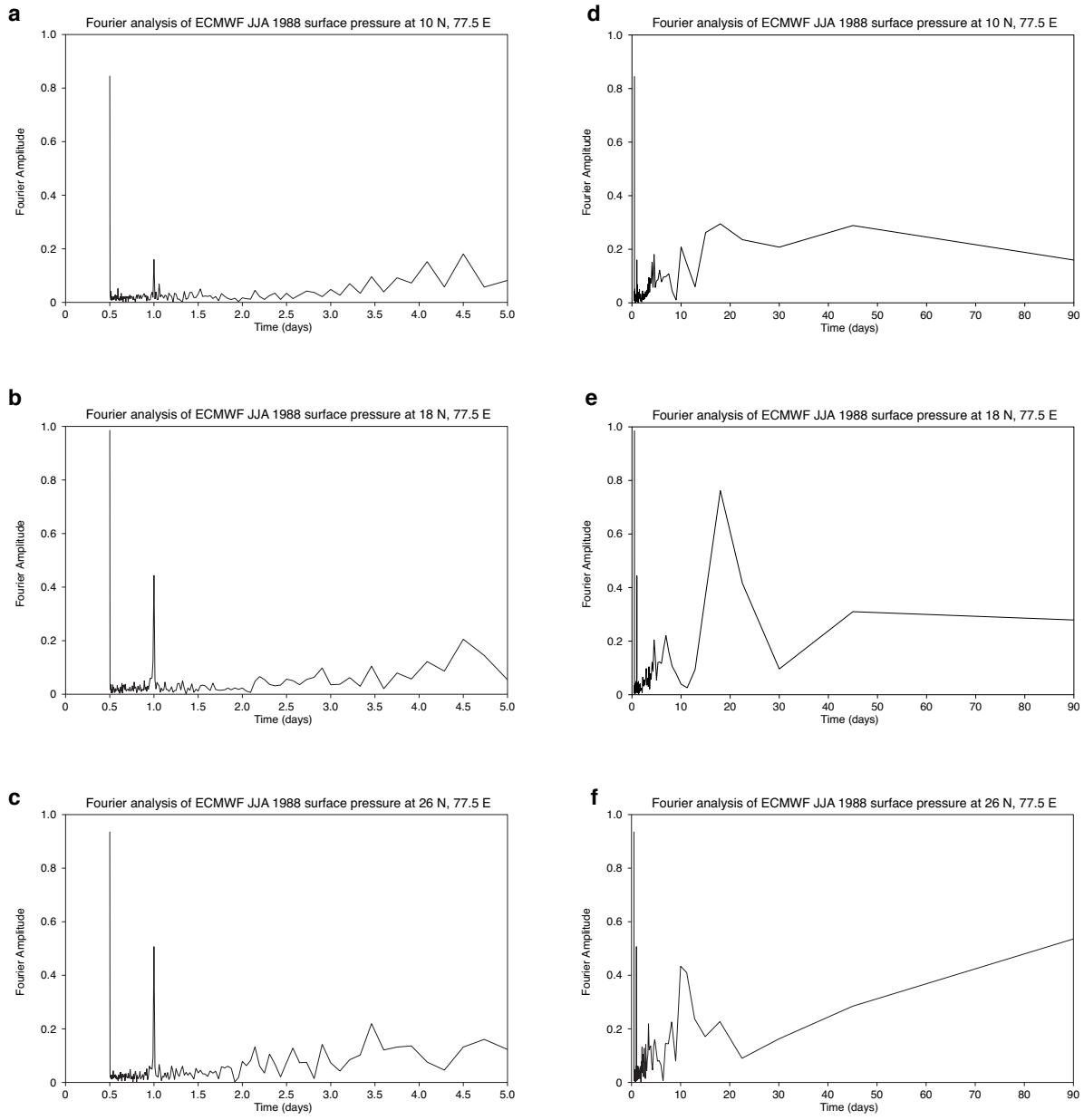

Figure 8.39: Observed Fourier power spectra of surface pressure. a) at 10ºN, 77.5ºE, near the southern tip of India, for oscillations with a period of 0-5 days; b) at 18ºN, 77.5ºE, for oscillations with a period of 0-5 days; c) at 26ºN, 77.5ºE, for oscillations with a period of 0-5 days; d) at 10ºN, 77.5ºE, for oscillations with a period of 0-90 days; e) at 18ºN, 77.5ºE, for oscillations with a period of 0-90 days; and f) at 26ºN, 77.5ºE, for oscillations with a period of 0-90 days.

the Himalayan foothills and back again (Webster, 1987). The 10-20 day peak for 26° N is at about 18 days, while the peak at 10° N is only 13 days. The difference between the two periods may be due to the monsoon trough not making it as far north as 26° N during some of its north-

south oscillations, which would cause the surface pressure dip to be less frequent there. There is also an oscillation with a period of 90 days that might be associated with the onset and departure of the monsoon itself.

Precipitation in the monsoon region has also been observed to vary systematically on time scales of 2-5 days, 10-20 days, and 40-50 days (Webster, 1987). Due to the convective nature of much of the precipitation that takes place in the Asian monsoon, there are also significant variations on time scales of only a few hours. For periods on the order of 10-20 days, the spectral analyses of precipitation are similar to those of surface pressure. This is understandable, because the heaviest precipitation is generally associated with the monsoon trough. There do not appear to be any semidiurnal or diurnal variations in precipitation, although there is a large amount of high-frequency spectral energy, as might be expected with convective rainfall. There are some indications of spectral peaks of precipitation at 25-35 days. These might be associated with the Madden-Julian oscillation.

There have been many efforts to understand the interannual variability that occurs with the Asian monsoon. The 1987 summer monsoon was remarkably dry over most of India (Krishnamurti et al., 1989) whereas the 1988 summer monsoon was quite wet in the same region (Krishnamurti et al., 1990). Charney and Shukla (1981) argued on the basis of modeling studies that the largest variations in monsoon rainfall and sea level pressure are due to changes in the boundary conditions. They speculated that if these boundary conditions could be specified without error, then seasonal predictions of quantities such as monthly-mean monsoon precipitation would be possible. Dry Asian summer monsoons have been statistically correlated to anomalously warm SSTs in the eastern Pacific (El Niño), and wet monsoons have been correlated to anomalously cold SSTs in the eastern Pacific (La Niña) by some researchers (Shukla and Paolino, 1983; Rasmusson and Carpenter, 1983), but there is some dissension on this point (Webster and Yang, 1992). The interannual variability of monsoons has also been linked to variations in snow cover (Yanai and Li, 1994; Barnett et al., 1989) and the thermal contrast between the equatorial Pacific and the Tibetan Plateau (Fu and Fletcher, 1985).

The Walker Circulation

The Walker Circulation (named by Bjerknes, 1966) is an east-west overturning of the atmosphere above the tropical Pacific Ocean, with rising motion on the west side, over the so-called “Warm Pool,” and sinking motion on the east side. The Walker Circulation can be viewed as a thermally excited stationary eddy. Although the Walker Circulation is driven by the east-to-west sea surface temperature gradient, it also helps to maintain that gradient through mechanisms to be discussed later. For this reason, the Walker Circulation is best understood as a coupled ocean-atmosphere phenomenon. It undergoes strong interannual variability. Fig. 8.40 is a schematic illustration of the Walker circulation and its relation to the surface wind field in the southern Hemisphere. The equatorward flow just west of South America can be viewed as the inflow to the ITCZ (which is generally north of the Equator in this region), and so it is in a sense a portion of the lower branch of the Hadley circulation. Fig. 8.41 shows the observed longitude-height cross sections of the zonal wind and vertical velocity, for January.

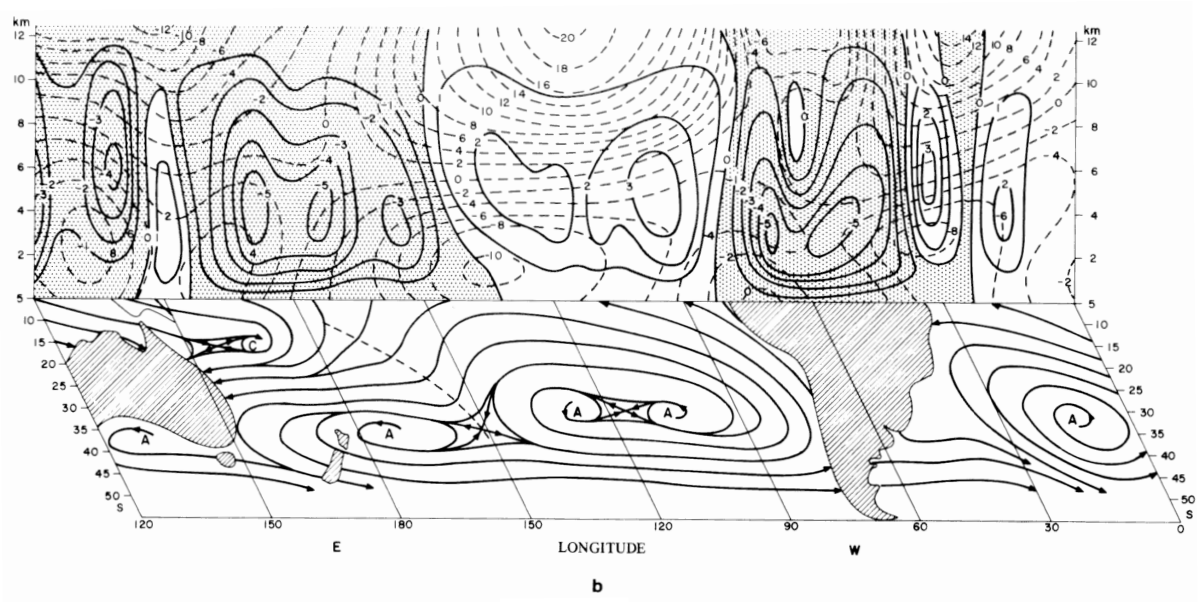

Figure 8.40: The Walker and Hadley Circulations. The foreground shows the lower-tropospheric flow for January. The vertical cross-section in the background shows the vertical motion (solid lines) and the zonal wind component at 5° S (dashed lines). From Philander (1990).

The Hadley Circulation is defined in terms of zonal averages, and so a particle participating in the Hadley Circulation through motions in the latitude-height plane cannot "escape" by moving to a different longitude. In contrast, the Walker Circulation is restricted to a narrow band of tropical latitudes, so that a particle participating in the Walker Circulation can escape by moving off to a different latitude; in fact, such meridional escapes are to be expected in view of the strong meridional motions associated with the Hadley Circulation. For this reason, we should not think of the Walker Circulation as a closed "race track;" it is better to view the Hadley and Walker Circulations as closely linked. For instance, a parcel may travel westward across the tropical Pacific in the lower branch of the Walker Circulation, ascend to the tropopause over the Warm Pool, and then move both poleward and eastward away from the Warm Pool, possibly descending in the subtropical eastern Pacific. It can then join the trades, and repeat its westward and Equatorward journey through the boundary layer.

Bjerknes (1969) theorized that the cool, dry air of the trade winds is heated and moistened as it moves westward until it finally undergoes large-scale moist-adiabatic ascent over the Warm Pool. If there were no mass exchange with adjacent latitudes, a simple circulation would develop in which the flow is easterly at low levels and westerly at upper levels. When meridional mass exchange is considered, this simple picture has to be altered, because absolute angular momentum is exported to adjacent latitudes. Under steady-state conditions, the flux divergence of angular momentum at the equator must be balanced by an easterly surface wind stress. Thus surface easterlies on the equator are stronger than those imposed by the Walker circulation. The net result is that a thermally driven Walker cell is imposed on a background of

easterly flow, the intensity of which depends on the strength of the angular momentum flux divergence.

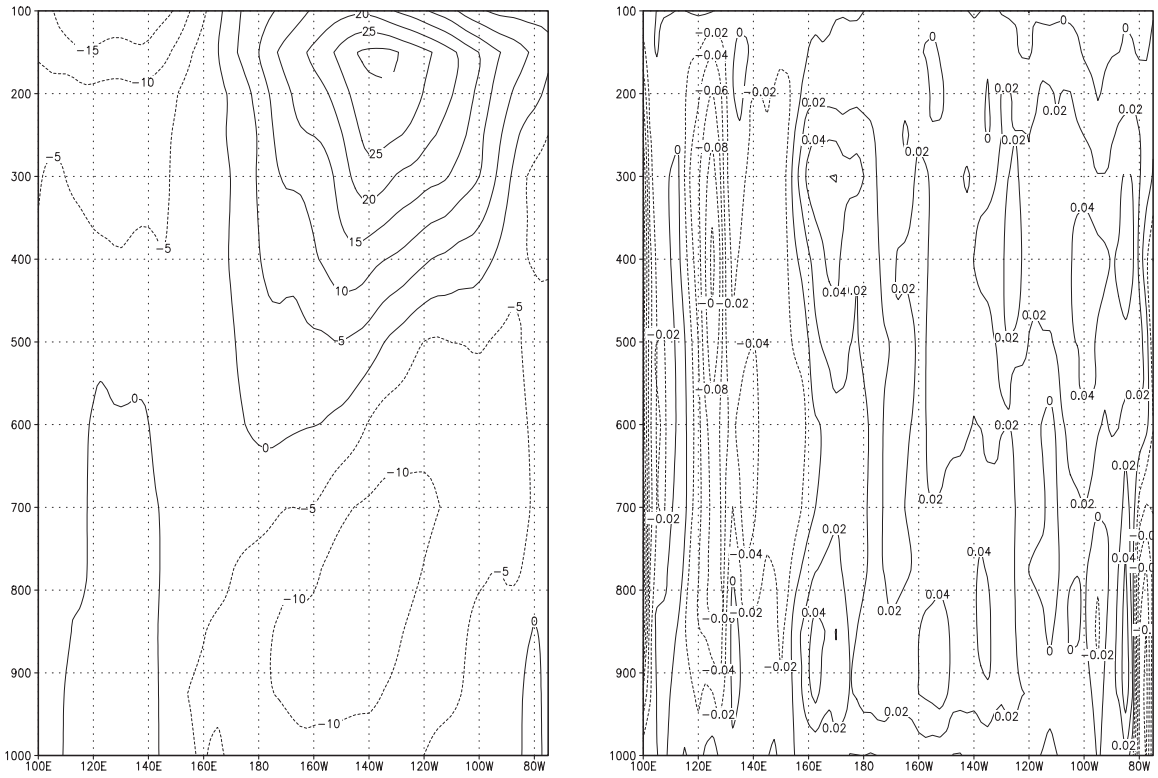

Figure 8.41: The longitude-height cross-sections of the zonal wind (left, in m s^{-1}) and vertical velocity (right, in Pa s^{-1}) along the Equator, for January, as analyzed by ECMWF.

Fig. 8.42 shows that the 1000-mb winds above the tropical Pacific (between 10° N and 10° S) have an easterly component in both solstitial seasons. For both seasons, easterly flow near the equator occurs west of about 90° W. In January, the easterly component is particularly strong above the central equatorial Pacific, and convergence is evident along the ITCZ near 8° N. In July, a notable characteristic is the strong cross-equatorial flow in the eastern Pacific. The zone of convergence at 1000 mb during the NH summer has moved north of 10° N over the eastern Pacific. At the latitude of the ITCZ, the easterly fetch originates to the east of Central America during both seasons. If we consider the Walker circulation to occur at near-equatorial latitudes, then easterly flow at 1000 mb cannot originate over the continents because the mountains of Peru act as a vertical barrier on the eastern boundary of the ocean. However, we note that a strong southerly component is evident during both seasons.

Lindzen and Nigam (1987) used a simple model to show that SST gradients are capable of forcing low-level winds and convergence in the tropics. They assumed that near the surface the Coriolis acceleration is balanced by the sum of the horizontal pressure-gradient force and wind stress; this is called an Ekman balance. Linearizing about a state of rest, they found a pressure field that qualitatively resembles the observations, although the wind speeds were

unrealistically strong. Neelin et al. (1998) show that the model used by Lindzen and Nigam (1987) is very similar to that of Gill (1980).

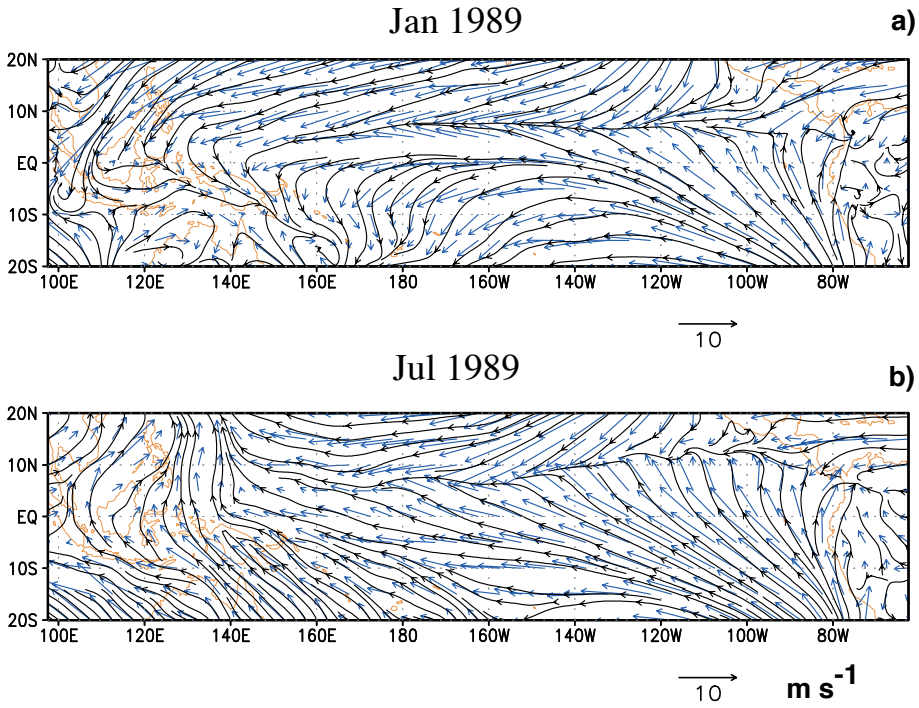

Figure 8.42: Streamlines and horizontal wind vectors for the tropical Pacific at 1000 mb for a) January 1989 and b) July 1989. The units are m s^{-1}.

Newell et al. (1996; hereafter N96) compared water-vapor data from the Upper Atmosphere Research Satellite (UARS) with upper-air wind data from the ECMWF reanalysis dataset to deduce horizontal and vertical motions in the tropical atmosphere. Their results indicate regions of strong ascending motion over the western Pacific Warm Pool and the South Pacific Convergence Zone. The main regions of sinking motion, which are located off South America and extend westward to the dateline just south of the equator, exhibit little seasonal movement. For comparison, Fig. 8.43 shows the vertical velocity fields at 300-mb from the ECMWF reanalysis dataset for the solstitial months. During January 1989, centers of ascending motion were located near 145° E at latitudes 5° N and 5° S. The SPCZ is clearly evident in the January 1989 data, with a large region of ascending motion that extends southeastward from 145° E to 160° W. A region of strong sinking motion straddles the equator and extends eastward from 160° E. During July 1989, the ascending region remains fixed at 145° E, but the NH and SH centers of ascending motion have merged on the equator. During the NH summer, the ITCZ is well developed at 5° N, and so the zone of sinking motion has slipped southward from its January position, particularly the zone over the central Pacific. The general pattern is one in which ascending motion dominates over the tropical western Pacific, while sinking motion occurs over the tropical central and eastern Pacific. Easterlies extend across the equatorial Pacific from South America to 170° W. West of 160° E, the low-level equatorial winds are very weak.

However, easterlies span the equatorial Pacific at 5° S and 5° N. Fig. 8.44 completes the picture, showing the upper branch of the Walker circulation. West of the dateline, the zonal winds over

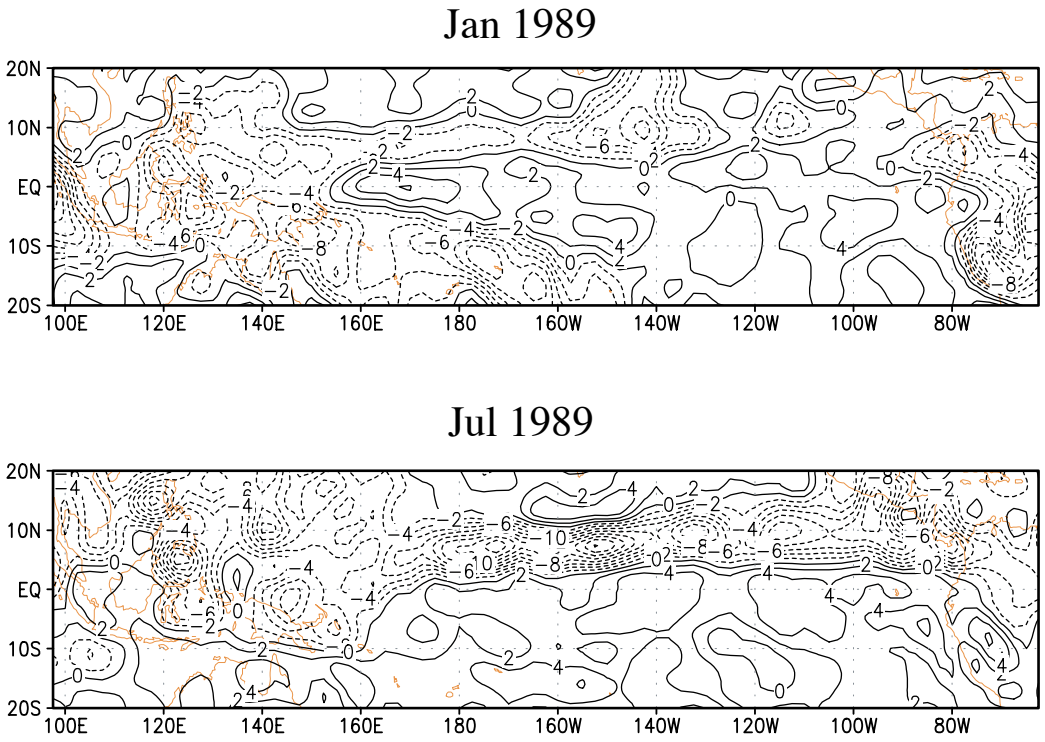

Figure 8.43: Contour plot of mean vertical velocity at 300 mb (units, 10^{-2} Pa s^{-1}) from the ECMWF reanalysis dataset for January and July, 1989. The contour interval is 2 x 10^{-2} Pa s^{-1}; negative contours are dashed. Data were obtained from NCAR.

the equator are easterly. Upper-level westerly flow occurs to the east of the rising motion. During July 1989, the upper-level flow above the equatorial Pacific ocean is entirely from the east. In the northern hemisphere (NH), weak westerly flow appears between 170° W and 140° W poleward of 15° N. In the southern hemisphere (SH), a westerly component of the wind exists south of 5° S to the east of the dateline. An interpretation is that the Walker circulation has migrated into the SH. A reexamination of Fig. 8.43b indicates that sinking motion is confined mainly to the SH, and occurs as far west as 165° W.

The Walker Circulation is an atmospheric phenomenon, but it is closely tied to east-west sea surface temperature gradients that are produced by atmospheric phenomena including aspects of the Walker Circulation itself. The Walker Circulation can thus be viewed as a phenomenon of the coupled atmosphere-ocean system. Fig. 8.45 shows that a sea-surface temperature (SST) maximum occurs over the tropical region centered on 120° E, and for this reason the region is known as the tropical Warm Pool. The "cold tongue" is a band of relatively cold waters along the equator that stretches from South America westward to near 160° E. Although a noticeable SST gradient exists along and across the cold tongue, the temperature variation is still much smaller than that which is generally observed in extratropical or polar regions of the globe. The tropical

climate is characterized by sea surface and horizontal air temperature gradients that are weak compared to the corresponding mid-latitude gradients. As explained by Charney (1963), for cloud-free regions of the tropics, pressure and temperature gradients must be small compared to those of midlatitudes.

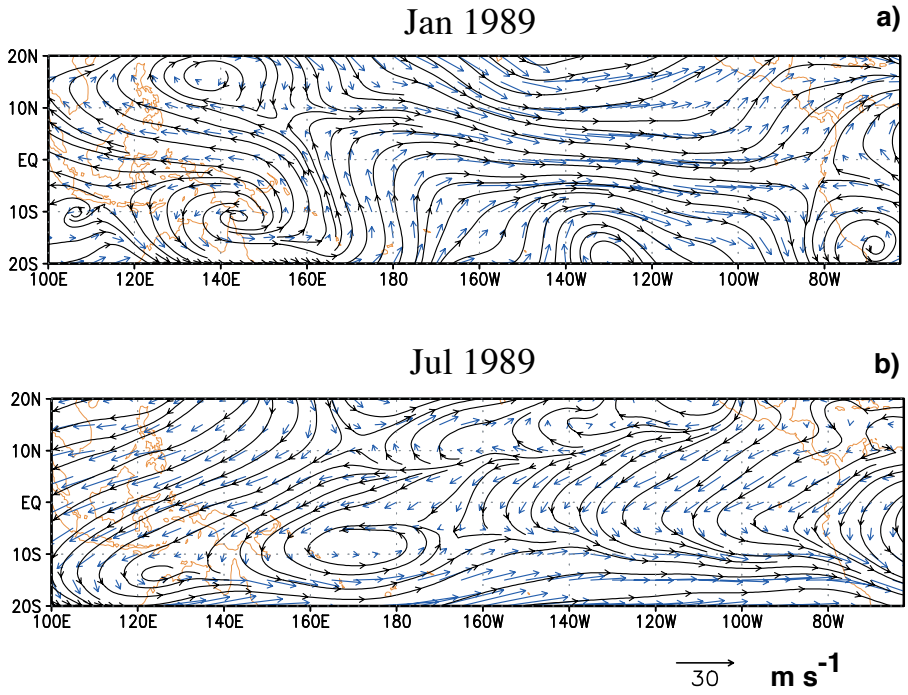

Figure 8.44: Streamlines and horizontal wind vectors for the tropical Pacific at 200 mb for a) January 1989, and b) July 1989.

The distribution of tropical convection is strongly related to both the local SST and the SST gradient. The tropical-Pacific Warm Pool is a region of intense deep convection. In Fig. 8.46, regions in which the outgoing longwave radiation (OLR) is less than 225 W m^{-2} can be identified as areas of frequent convection (Webster 1994). The OLR threshold corresponds to a

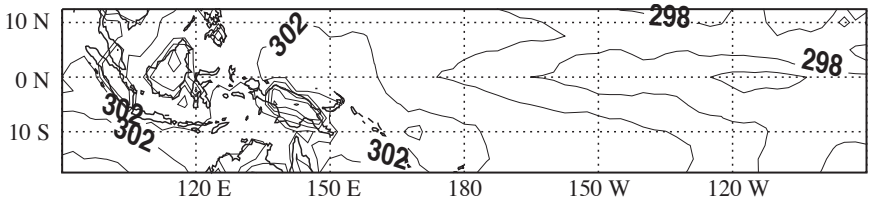

Figure 8.45: Tropical skin temperature for January 1989 from the ECMWF reanalysis dataset obtained from NCAR. Resolution for this dataset is 2.5º and the contour interval is 2 K.

monthly mean emission temperature of 250 K. Due to longwave trapping by optically thick anvil clouds, which are produced by deep convection, the OLR is reduced and threshold values of OLR can therefore be used as surrogates to infer the presence of convection. From the figure, we see that convection occurs throughout the Warm Pool, and in the South Pacific convergence zone

(SPCZ). On the other hand, the OLR is generally larger than 275 W m^{-2} across the equatorial cold tongue, indicating that convection is infrequent there.

The high, cold, and sometimes bright clouds of the Warm-Pool region limit the radiative cooling of the atmosphere over the Warm Pool, but they also limit the solar warming of the ocean. Ramanathan and Collins (1991) hypothesized that cirrus clouds act as a thermostat to regulate tropical SST. They used Earth Radiation Budget Experiment (ERBE) data to deduce the inter-relationships among shortwave and longwave cloud radiative forcings and radiative forcing

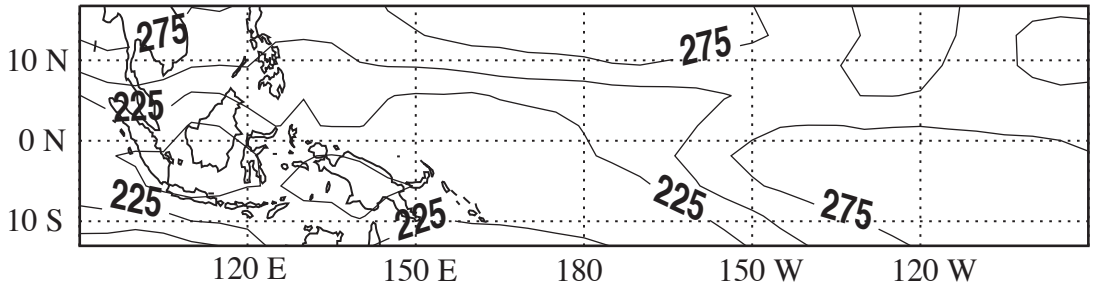

Figure 8.46: Tropical OLR for January averaged over 1985 to 1988. Daily means from the Earth Radiation Budget Experiment (ERBE) were averaged and interpolated onto a 5º x 4º (longitude-latitude) grid. The units are W m^{-2}.

of the clear atmosphere. They emphasized that the shortwave effects of clouds dominate over the longwave effects in regulating SST. According to their hypothesis, as SST increases, the cloud albedo increases. According to their idea, the atmosphere warms as a result of longwave cloud radiative effects, stronger latent-heat release by convection, and a stronger SST gradient over the tropical Pacific. This warming leads to an amplification of the large-scale flux convergence of moisture. The process continues until the reflectivity clouds increases sufficiently to cool the surface. A criticism of their study is that changes in the strength of the ocean and atmosphere circulations were not included. Nevertheless it undoubtedly true that the blocking of shortwave radiation by deep cloud systems tends to limit the sea-surface temperature in the Warm Pool.

As seen in Fig. 8.47, in the eastern tropical Pacific, stratus clouds in the boundary layer intercept sunlight and strongly reduce the heat flux into the ocean below (e.g. Hartmann et al. 1992). In this way, the atmosphere helps to maintain the cooler SSTs of the eastern Pacific. Stratus clouds form preferentially over cold water (Klein and Hartmann 1993), so a positive feedback is at work here (Ma et al. 1996). Latent heat exchange between the ocean and atmosphere is influenced by the surface relative humidity and the surface winds. For fixed relative humidity and SST, the evaporative cooling of the ocean increases as the surface wind stress increases. The winds also influence the SST distribution by generating cold-water upwelling in the eastern Pacific, and along the equator in the eastern and central Pacific. As discussed in Chapter 2, the equatorial cold tongue is due to upwelling driven by the trade winds.

The driving force behind the Walker circulation is the zonally varying heating that is balanced by zonally varying adiabatic heating/cooling due to sinking/rising motions. Over the

warm waters of the western Pacific, latent heat release due to intense convection and radiative warming of the atmospheric column are balanced by the adiabatic cooling associated with rising

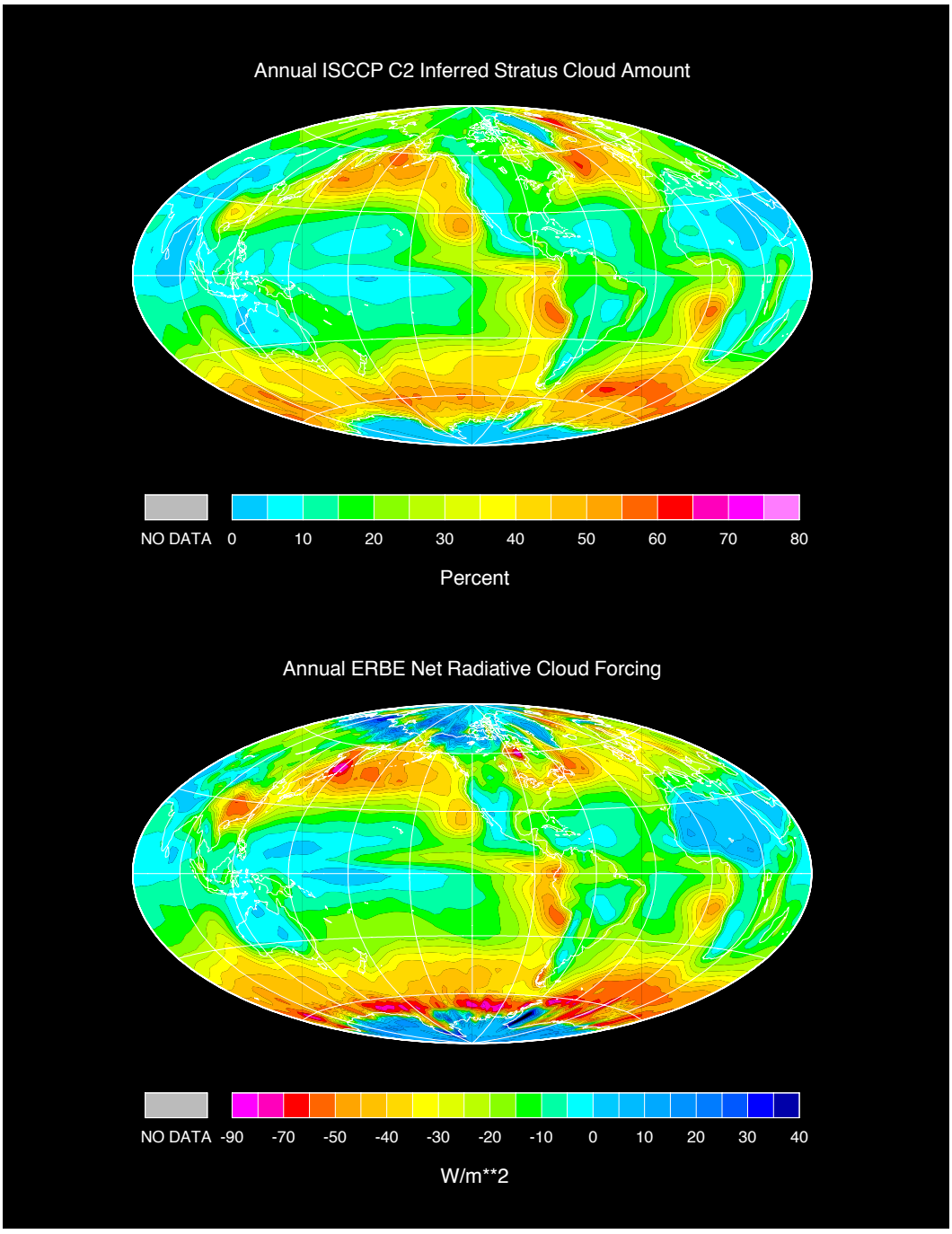

Figure 8.47: Observed annual-mean low-cloud amount (upper panel) and the net effects of clouds on the Earth's radiation budget (lower panel). Negative values in the lower panel indicate a cooling, i.e., shortwave reflection dominates longwave trapping. From Norris and Leovy (1994).

motion (Webster 1987). Over the eastern tropical Pacific, where the SST is relatively cold,

convection is infrequent, and so a balance between radiative cooling and subsidence exists, as discussed in Chapter 2.

Pierrehumbert (1995; hereafter P95) presented a two-box model of the Hadley/Walker circulation that has strongly influenced recent studies of the tropical climate. Fig. 8.48 presents a

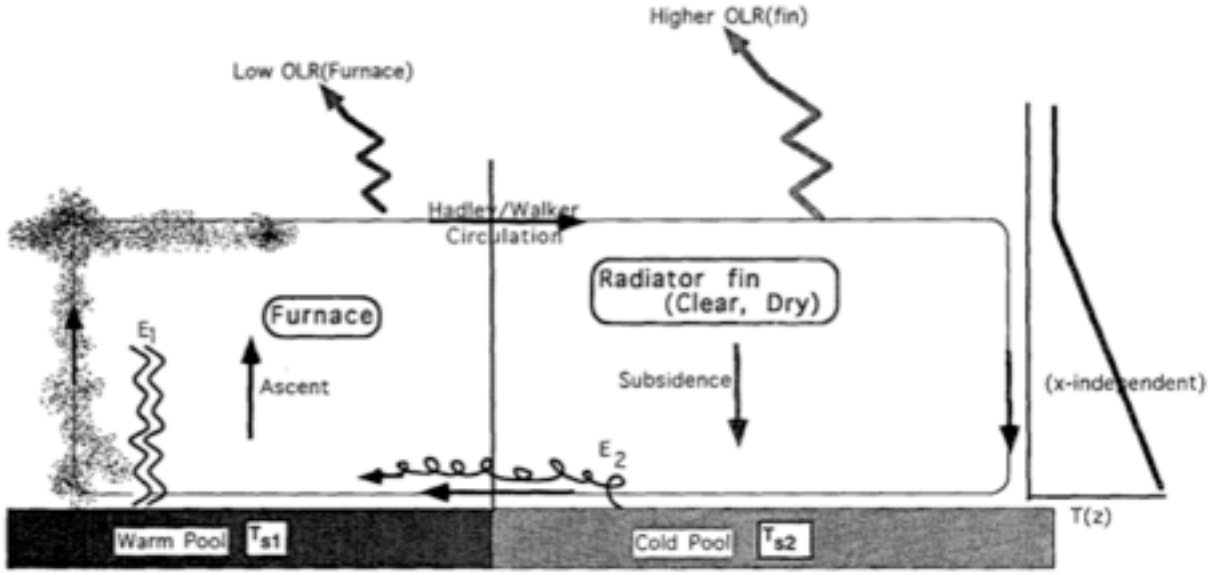

Figure 8.48: Pierrehumbert's (1995) schematic representation of the 'furnace/radiator-fin' model of tropical circulation. The symbols E and TS represent the evaporation rate and SST, respectively. The subscript 1 denotes the Warm Pool or "furnace," and the subscript 2 denotes the Cold Pool or "radiator fin."

schematic of his "furnace/radiator-fin" model. The model has separate energy budgets for its Cold-Pool and Warm-Pool regions. The SSTs of the Cold Pool and Warm Pool are assumed to be those that give energy balance for each box of the model atmosphere and for the Cold-Pool ocean. Surface energy balance for the Warm Pool was not explicitly included in the model. A vertically and horizontally uniform lapse rate was assumed, and the free-tropospheric temperature profile was assumed to be uniform across the tropics. The radiating temperature of the Cold-Pool free atmosphere was assumed to be the air temperature at $z=\frac{z_T}{2}$, where z_T is the height of the tropopause. The solution was obtained by first computing the net energy flux at the top of the Warm-Pool atmosphere for a given SST and relative humidity profile. The net radiative flux at the Warm-Pool TOA was assumed to be balanced by a horizontal energy transport to the Cold Pool. The Cold-Pool SST and radiating temperature were then computed under the constraint that the net diabatic cooling must balance the energy imported laterally from the Warm Pool.

The mass flux of the Hadley-Walker circulation was assumed to be that required to give a balance between adiabatic warming by dry subsidence and the net radiative cooling of the Cold-

Pool region. It can be shown that the horizontal heat transport by the Warm-Pool atmosphere is proportional to the diabatic cooling of the Cold-Pool atmosphere and to the ratio of Cold-Pool area and Warm-Pool area. This area ratio is a prescribed parameter of the model. P95 showed that for very small values of the Cold-Pool emissivity, the Warm-Pool SST increases without limit, because the Cold Pool cannot radiate enough energy to balance the energy absorbed in the Warm Pool. Because the Warm Pool controls the temperature profile, its equilibrium SST must decrease as the Cold-Pool radiating temperature decreases. As the Cold-Pool emissivity increases, the Warm-Pool cools off. The simulated Cold-Pool and Warm-Pool SSTs resemble the present-day climate for a range of conditions. The diagnosed mass flux is realistic. A weakness of the model is that it fails to account for cloud-radiative effects.

Miller (1997; hereafter M97) extended Pierrehumbert's model by studying the radiative effects of low clouds in the Cold-Pool region. Extending the basic concepts of P95, Miller constructed a three-box model, which includes energy- and moisture-balance equations for the boundary layer and free troposphere and a surface energy budget for each of three boxes: the updraft region, the Warm Pool, and the Cold Pool. Taking advantage of the small surface area covered by the updrafts, M97 simplified the model for the limit of vanishing updraft surface area. M97 demonstrated that in this limit, the boundary layer and tropopause of the Warm Pool region must be connected by a moist adiabat. Miller assumed that the lapse rate of the Warm Pool region is moist adiabatic. Following P95, atmospheric dynamics were implicitly included by assuming a uniform free-tropospheric temperature sounding across both the Warm and Cold Pool regions.

The main finding of M97 is that low clouds act as a thermostat for tropical SST. Without a realistic distribution of stratus clouds, the SST was too warm beneath the subsiding branch of the tropical circulation. Although low clouds reduce the surface-absorbed solar radiation locally, M97 also found that the temperature drop in the Warm Pool region was nearly as large as that in the Cold Pool region. In order to obtain a realistic Warm-Pool SST, the model needed an additional $-5\ \mathrm{W\ m^{-2}}$ forcing for the Cold Pool, and $-12\ \mathrm{W\ m^{-2}}$ forcing for the Warm Pool. This suggests that additional cloud types contribute to the surface forcing. In contrast, P95 did not need cloud radiative forcing to simulate realistic SSTs. It is possible that the simplified radiative transfer parameterization used by P95 compensated for the lack of cloud radiative effects.

The Madden-Julian Oscillation

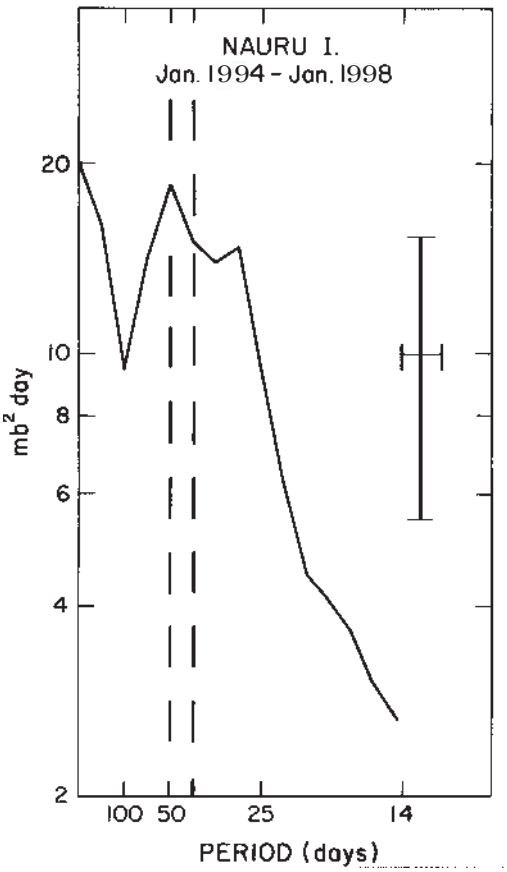

Figure 8.49: Variance spectrum for station pressures at Nauru Island, 0.4° S, 161.0° E. Ordinate (variance/ frequency) is logarithmic and abscissa (frequency) is linear. The 40-50-day period range is indicated by the dashed vertical lines. Prior 95% confidence limits and the bandwidth of the analysis (0.008 day-1) are indicated by the cross. From Madden and Julian (1994). Taken from Madden and Julian (1972 a).

The MJO (Madden-Julian Oscillation) can be defined as a broad region of humid air and vigorous precipitation that maintains itself as it drifts slowly eastward across the tropical Indian and Western Pacific Oceans (Madden and Julian, 1971, 1972). It reoccurs with a period in the range 30 - 60 days (Fig. 8.49). The oscillation appears most clearly over the Indian and western Pacific Oceans, and involves many meteorological variables, including the zonal wind, surface pressure, temperature, and humidity.

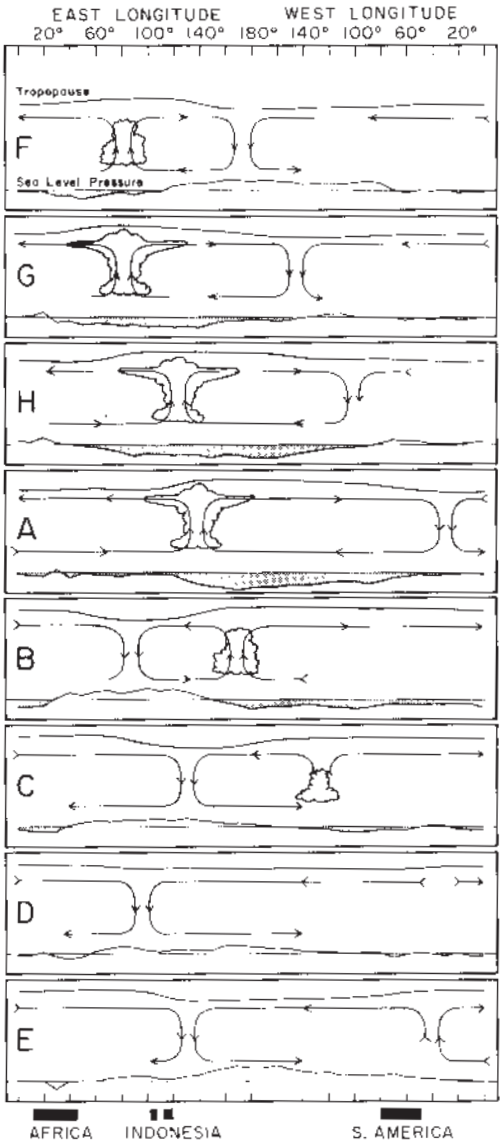

Figure 8.50: Schematic depiction of the time and space (zonal plan) variations of the disturbance associated with the 40-50-day oscillation. Dates are indicated symbolically by the letters at the left of each chart and correspond to dates associated with the oscillation in Canton's station pressure. The letter A refers to the time of low pressure at Canton and E is the time of high pressure there. The other letters represent intermediate times. The mean pressure disturbance is plotted at the bottom of each chart with negative anomalies shaded. The circulation cells are based on the mean zonal wind disturbance. Regions of enhanced large-scale convection are indicated schematically by the cumulus and cumulonimbus clouds. The relative tropopause height is indicated at the top of each chart. From Madden and Julian (1994), based on Madden and Julian (1972 a).

Observations have revealed corresponding oscillations of various measures of cumulus activity (Murakami et al., 1986), including precipitation (Hartmann and Gross, 1988) and outgoing longwave radiation (e.g., Weickmann and Khalsa, 1990). A schematic depiction of the

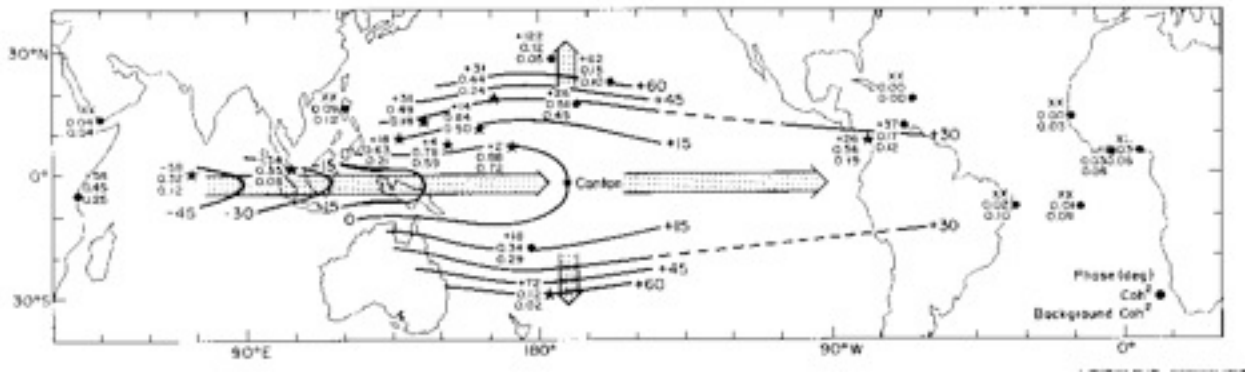

Figure 8.51: Mean phase angles (in degrees), coherence squares, and background coherence squares for approximately the 36-to-50-day period range of cross spectra between surface pressures at all stations and those at Canton. The plotting model is given in the lower right-hand corner. Positive phase angle means Canton times series leads. Stars indicate stations where coherence squares exceed a smooth background at the 95% level. Mean coherence squares at Shemya (52.8º N, 174.1º E) and Campbell Island(52.6º S, 169.2º E; not shown) are 0.08 and 0.02, respectively. Both are below their average background coherence squares. Values at Dar es Salaam (0.8º S, 39.3º E) are from a cross spectrum with Nauru. The arrows indicate propagation direction. From Madden and Julian (1994), adapted from Madden and Julian (1972).

MJO is provided in Fig. 8.50. The observed phase propagation is shown in Fig. 8.51. The MJO does not produce oscillations of convection over the Amazon basin or the Congo basin, even though an MJO signal is sometimes observed in the winds at those longitudes.

The MJO is a deep, equatorially-trapped, convectively-coupled disturbance that propagates eastward at about 5 m s^{-1} as it travels from the Indian Ocean eastward to the dateline, at which point decouples from the convection and speeds up to about 12 m s^{-1}. Global circumnavigation can sometimes be seen in upper-tropospheric velocity potential, but is harder to detect in other fields closer to the surface. In the oceanic regions of the eastern hemisphere, the disturbance is closely coupled to convective processes and is characterized by strong surface fluxes. During its convectively inactive phase, the MJO is characterized by weak surface easterlies and high downward radiative fluxes. Embedded within the broad convectively active region are fluctuations on smaller space and time scales (Nakazawa, 1988; Fig. 8.52). In the convectively active phase, strong surface westerlies and high surface latent heat fluxes are observed.

As would be expected from the theoretical work of Matsuno (1966) and Gill (1980), the precipitation maximum of the MJO is accompanied by low-level winds that trace out twin cyclones on the west side of the precipitation maximum, and a zonally broader patch of easterly winds on the east side. The zonal wind field thus converges at low levels near the precipitation maximum, and diverges aloft.

The MJO has been the subject of intense research because of its intraseasonal time scale, which suggests that intraseasonal weather anomalies may be predictable; and also because of their apparent relationships with the Indian summer monsoon (Yasunari, 1979; Krishnamurti and

Subrahmanyam, 1982), the likelihood of tropical Pacific storms (Gray, 1979), and the initiation of El Niño events (Lau and Chan, 1985). Since tropical convection can force Rossby waves that propagate into the extratropics, the MJO can also influence the weather in middle latitudes (e.g., Rueda, 1991).

Atmospheric global circulation models (GCMs) have difficulty in simulating the MJO. In addition, there is no consensus on the basic physical mechanisms that give rise to the MJO,

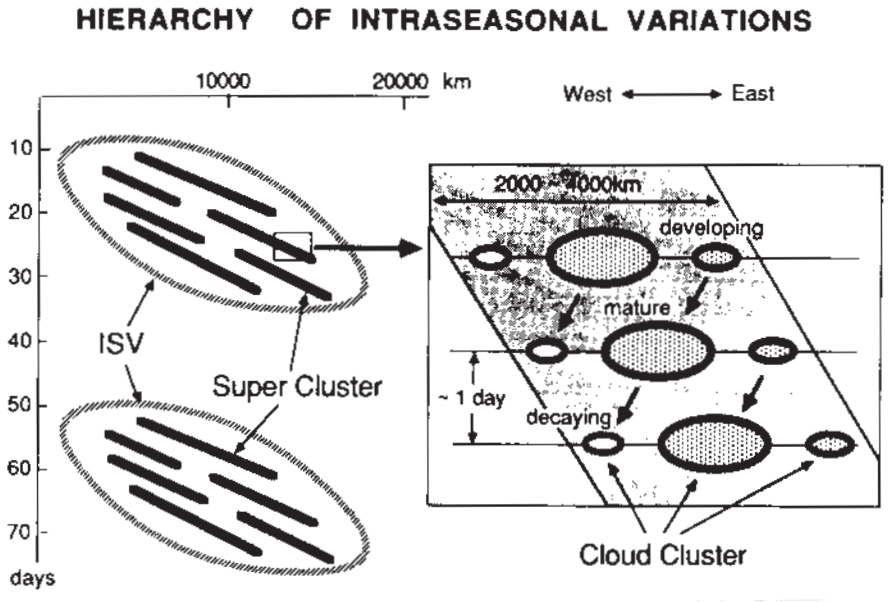

Figure 8.52: Schematic describing the details of the large-scale eastward-propagating cloud complexes (slanting ellipses marked ISV (intraseasonal variability) on the left-hand side). Slanting heavy lines represent super cloud clusters (SCC) within the larger complexes or ISV. The right-hand side illustrates the fine structure of the SCC with smaller westward-moving cloud clusters that develop, grow to maturity, and decay in a few days. From Nakazawa (1988).

although progress is being made. The problem of understanding the MJO can be separated into several linked parts. First of all, the steady motion generated by a moving heat source on an Equatorial beta plane is well described by the the model of Matsuno (1966) and Gill (1980), hereafter called the "MG model," and has been studied by by Hendon and Salby (1994) and Schubert and Masarick (2006), among others. The relevance of the MG model to the MJO has been recognized for decades (e.g., Chao, 1987). Given a realistic moving heat source, the dynamical core of any GCM should be able to simulate a wind field similar to that of the observed MJO.

On the other hand, the failure of Matsuno's (1966) theory of Equatorial waves to produce free modes resembling the MJO, despite the theory's success in predicting the other observed Equatorial waves (with the possible exception of Easterly Waves; Kiladis et al., 2009), implies that the MJO depends fundamentally on processes that were not included in Matsuno's model. It

is now believed that moist processes are essential to the MJO (e.g., Raymond, 2001; Grabowski and Moncrief, 2004; Bony and Emanuel, 2005, and the term "moisture mode," which was coined by Fuchs and Raymond (2007), is now widely used to describe the MJO (Sugiyama , 2009). The dry MG model cannot describe moisture modes.

Although variations of longwave radiative heating and surface evaporation may be important for the MJO (e.g., Raymond, 2001; Bony and Emanuel, 2005; Andersen and Kuang, 2012; Arnold et al., 2013), GCMs should be able to simulate them, at least qualitatively, so they are probably not the missing ingredients that prevent many conventional GCMs from simulating the MJO.

The eastward propagation of the MJO is favored by moistening of the atmosphere its east side, and drying on its west side. Moisture advection is a key process in the MJO (Maloney, 2009; Maloney et al., 2010; Andersen and Kuang, 2012; Pritchard and Bretherton, 2013). Both meridional and vertical advection favor strong drying to the west of the humid, strongly precipitating core of the MJO; under such conditions, positive water vapor anomalies must move eastward or be destroyed. In modeling terms, the drying is due to resolved-scale advection, rather than parameterized processes. The precipitation rate is observed to decrease as advection dries the air on the west side of the MJO. Perhaps surprisingly, not all models simulate this (Holloway et al., 2013). In an analysis of numerical experiments with an aquaplanet version of the SP-CAM, Andersen (2012) found that the eastward drift of the MJO speeds up if the subtropical reservoir of dry air is brought closer to the Equator.

WISHE (wind-induced surface heat exchange) theories of the MJO focus on the moistening on the east side (Emanuel, 1987; Neelin and Yu, 1994; Yu and Neelin 1994). Neelin et al. (1987) and Emanuel (1987) hypothesized a feedback in which the maintenance and eastward propagation of convection depend on the existence of a "basic state" easterly flow. The perturbation easterlies on the east side of the MJO reinforce the basic state easterlies, favoring strong evaporation on the east side. Westward zonal advection can carry the evaporatively moistened air towards the convective disturbance (Sobel and Maloney, 2012, 2013). The observed specific humidity fluctuations are strongest near the 700 hPa level, however (Sherwood, 1999; Holloway and Neelin, 2009). The MJO needs a process that lifts the water vapor from the surface and moistens the 700 hPa level. The *upward* transport of moisture on the east side of the MJO is due to processes that are parameterized in GCMs. Some GCMs that fail to simulate the MJO are unable to moisten the air in regions of strong precipitation (Thayer-Calder and Randall, 2009; Kim et al., 2009; Landu and Maloney, 2011; Mapes and Bacmeister, 2012; Hung et al., 2013; Kim et al., 2013).

As discussed by Thayer-Calder and Randall (2009), an essential requirement for successful simulation of the MJO is the ability to produce realistically deep layers of high relative humidity in regions of strong precipitation. Chikira (2013) presents an insightful analysis of the processes that lead to mid-tropospheric humidity changes in tropical convective systems. Recall from Chapter 6 that the dry static energy and water vapor are governed by

$$\rho\frac{\partial\overline{s}}{\partial t}=-\rho\overline{\mathbf{V}}\cdot\nabla\overline{s}-\widetilde{M}\frac{\partial\overline{s}}{\partial z}+\overline{Q_R}+\rho L\widetilde{C}+D\left(s_c-\overline{s}\right),$$

(106)

$$\rho\frac{\partial\overline{q_v}}{\partial t}=-\rho\overline{\mathbf{V}}\cdot\nabla\overline{q_v}-\widetilde{M}\frac{\partial\overline{q_v}}{\partial z}-\rho\widetilde{C}+D\left[\left(q_v\right)_c-\overline{q_v}\right],$$

(107)

where $\widetilde{M}$ is the environmental mass flux (positive upward), $\widetilde{C}$ is the net environmental condensation rate, and D is the rate at which mass is detrained from the convective updrafts.

Following Charney (1963), we assume that in the tropical free atmosphere the dry static energy is independent of time and horizontal position. This is used to justify the "weak temperature gradient approximation." Then (106) becomes

$$0\cong-\widetilde{M}\frac{\partial\overline{s}}{\partial z}+\overline{Q_R}+\rho L\widetilde{C}+D\left(s_c-\overline{s}\right),$$

(108)

or

$$\widetilde{M}\cong\left[\overline{Q_R}+\rho L\widetilde{C}+D\left(s_c-\overline{s}\right)\right]\left(\frac{\partial\overline{s}}{\partial z}\right)^{-1}.$$

(109)

According to (109), the environmental mass flux is whatever it takes to maintain a constant dry static energy. Assuming that $\frac{\partial\overline{s}}{\partial z}>0$, the condition for the environmental mass flux to be upward is $\overline{Q_R}+\rho L\widetilde{C}+D\left(s_c-\overline{s}\right)>0$. This is relevant because an upward environmental mass flux will moisten the environment.

We can use (109) to eliminate $\widetilde{M}$ in (107). The resulting equation does not contain an explicit vertical advection term. We obtain

$$\rho\frac{\partial\overline{q_v}}{\partial t}=-\rho\overline{\mathbf{V}}\cdot\nabla\overline{q_v}+\frac{\alpha}{L}\left[\overline{Q_R}+\rho L\widetilde{C}+D\left(s_c-\overline{s}\right)\right]-\rho\widetilde{C}+D\left[\left(q_v\right)_c-\overline{q_v}\right],$$

(110)

where

$$\alpha\equiv-L\left(\frac{\partial\overline{q_v}}{\partial z}\right)\left(\frac{\partial\overline{s}}{\partial z}\right)^{-1},$$

(111)

so that

$$1-\alpha=\left(\frac{\partial\overline{h}}{\partial z}\right)\left(\frac{\partial\overline{s}}{\partial z}\right)^{-1}.$$

(112)

From (111), we expect $\alpha>0$. From (112), we see that that throughout most of the tropics α should be greater than one in the lower troposphere, and close to zero in the upper troposphere. If the environment becomes saturated with a moist adiabatic lapse rate, then $\alpha=1$.

By combining terms, we can rewrite (110) as

$$\rho\frac{\partial\overline{q_v}}{\partial t}=-\rho\overline{\mathbf{V}}\cdot\nabla\overline{q_v}+\frac{\alpha}{L}\overline{Q_R}-(1-\alpha)\rho\widetilde{C}+D\left\{\left[\left(q_v\right)_c-\overline{q_v}\right]+\frac{\alpha}{L}\left(s_c-\overline{s}\right)\right\}.$$

(113)

If detrainment occurs at levels where $s_c\cong\overline{s}$, then (113) can be approximated by

$$\boxed{\rho\frac{\partial\overline{q_v}}{\partial t}\cong-\rho\overline{\mathbf{V}}\cdot\nabla\overline{q_v}+\frac{\alpha}{L}\overline{Q_R}-(1-\alpha)\rho\widetilde{C}+D\left[\left(q_v\right)_c-\overline{q_v}\right]}.$$

(114)

Inspection of (114) shows that:

- Radiative cooling dries, and radiative heating moistens. If water vapor and radiative heating are positively correlated, then radiation may be able to drive an instability, as suggested by Raymond (2001).
- For $\alpha>1$, which is expected in the lower troposphere, evaporation into the environment (i.e., $\widetilde{C}<0$) dries the lower troposphere. *This means that for $\alpha>1$ the evaporation of stratiform precipitation in the lower troposphere actually has a drying effect.*
- Condensation in the environment (i.e., $\widetilde{C}>0$) leads to a net moistening when $\alpha>1$. However, if the environment is saturated with a moist adiabatic lapse rate, then we expect $\alpha\cong1$, in which case environmental condensation has no effect on the water vapor.
- For $0<\alpha<1$, the evaporation of falling rain into the environment can drive an increase in moisture variance, consistent with the "stratiform instability" of Mapes (2000) and Kuang (2008). It appears that the best chance for instability comes with $0<\alpha<1$.

Summary

In this chapter, we explored the nature of the wave motions that fill the atmosphere, from both theoretical and observational perspectives. The basic theory of planetary waves, governing even the tropical waves that were rediscovered by Matsuno in 1966, was worked out by Laplace.

The waves of largest scale are by far the most energetic, and produce the lion's share of the fluxes of heat, moisture, and momentum that affect the zonally averaged flow.

Monsoons and the Walker Circulation can be regarded as thermally forced eddies.

Transient and stationary eddies are both important components of the general circulation of the atmosphere. Their effects on the mean flow are the subject of the next Chapter.

Problems

1. Show that for an isothermal atmosphere the static stability, S_p , increases strongly upward.

2. Prove that an isothermal atmosphere has only one equivalent depth for free oscillations, given by

$$\hat{h} = \gamma H \ , \tag{115}$$

where $\gamma \equiv c_p / c_v$, and $H = RT / g$.

3. Show that for a resting isentropic basic state with no heating and no gravitational forcing, perturbations satisfy

$$\frac{\partial z'}{\partial t} + H_0 \nabla_p^2 \chi = 0 \tag{116}$$

for all p, where $H_0 = R\overline{T}_0 / g$ and subscript zero denotes a surface value. Eq. (107) looks very much like the continuity equation for shallow water.

4. Show that the FOSC with the nondivergent approximation $(\chi = 0)$ must satisfy

$$\nabla_p \cdot \left(f \nabla_p \psi \right) = g \nabla_p^2 z' \ . \tag{117}$$

This implies that the oscillation is not in exact geostrophic balance.

5. Planetary waves are often observed to "bend westward" on their equatorward sides (see Fig. 8.53). Give an explanation for this based on Rossby wave dynamics.

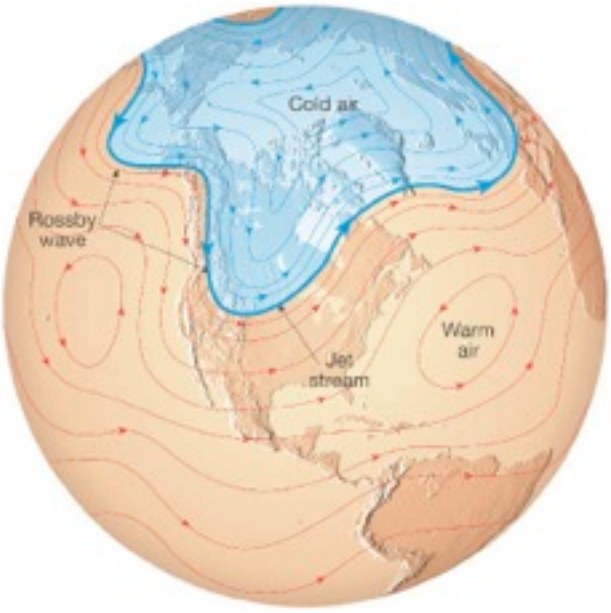

Figure 8.53: Schematic illustration of the structure of a Rossby wave. From http://www.daukas.com/Geoscience/MAtour/images/Rossby2.jpg.

6. Show that when there is a basic zonal current from the west, with a constant angular velocity $\dot{\lambda}$, the apparent "phase speed" relative to the Earth's surface of the free oscillation of the second class is given by

$$-\frac{\sigma}{s} = \dot{\lambda} - \frac{2\left(\Omega + \dot{\lambda}\right)}{n\left(n+1\right)}. \tag{118}$$

7. Laplace's tidal equation has a special solution when the period of oscillation is 1/2 day and the longitudinal wave number is zero.

a) Show that the solution is given by

$$\Theta_n = A\sin\left(\sqrt{\varepsilon_n}\mu\right) + B\cos\left(\sqrt{\varepsilon_n}\mu\right) \tag{119}$$

where

$$\varepsilon_n = \frac{4\Omega^2 a^2}{g h_n} \tag{120}$$

b) Show that

$$\varepsilon_n = \left(\frac{1}{2}n\pi\right)^2, n = 1,2,\ldots$$

$$A / B = \tan\left(\frac{n\pi}{2}\right)$$

(121)

is required to satisfy the boundary condition that $\mathbf{V}_n'$ remains finite at the poles. This solution was found by Solberg (1936).

10. Verify by substitution that

$$\hat{v}(y) = Ce^{-\frac{1}{2}y^2} H_n(y)$$

(122)

is a solution of

$$\frac{d^2\hat{v}}{dy^2} + \left(\sigma^2 - k^2 + \frac{k}{\sigma} - y^2\right)\hat{v} = 0 .$$

(123)

11. For the model of Charney and Eliassen, derive an expression for the "mountain torque" associated with a stationary, frictionally damped, resonant wave (i.e., $K^2 = K_S^2$).

QuickStudies Referenced

Laplace Tidal Equations

Spherical Harmonics

Shallow Water Equations

Hermite Polynomials