Technical Report
RAL-TR-2003-008

CLRC

The Physics of the Equatorial Atmosphere

L Gray K Hamilton

7th March 2003

COUNCIL FOR THE CENTRAL LABORATORY OF THE RESEARCH COUNCILS

Enquiries about copyright, reproduction and requests for additional copies of this report should be addressed to:

The Central Laboratory of the Research Councils
Library and Information Services
Rutherford Appleton Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QX
Tel: 01235 445384 Fax: 01235 446403
E-mail library@rl.ac.uk

ISSN 1358-6254

The Physics of the Equatorial Atmosphere

Lecture notes from a summer school held on this topic at the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy September 24 – October 5, 2001.

Edited by Lesley Gray and Kevin Hamilton

Contents

Preface

A School on "Physics of the Equatorial Atmosphere" was held at the Abdus Salam International Centre for Theoretical Physics from September 24 through October 5, 2001. The School was aimed primarily at advanced graduate students and young postdoctoral scientists with basic background in atmospheric science. The School featured 11 lecturers from Australia, Brazil, Germany, India, the UK, and the USA, and attracted 43 students from 27 countries including 12 students from Africa, 8 from South Asia and Indonesia, and 8 from Latin America.

The School lectures discussed a wide range of topics. The first week of lectures covered dynamics and transport in the low-latitude neutral atmosphere, with a focus on the stratosphere and mesosphere. The second week dealt with topics related to the equatorial ionosphere and magnetosphere.

This volume reproduces slightly revised versions of some of the lecture notes from the first week of the School. The overall goal of this volume is to provide an informal survey of key issues in the meteorology of the low-latitude middle atmosphere, one that significantly supplements available treatments of the subject in textbooks such as Holton (1975), Andrews et al. (1987) and Holton (1993), or in relevant review articles, notably Wallace (1973), Holton et al. (1995) and Baldwin et al. (2001).

Chapter 1

Introduction to the Dynamics of the Tropical Stratosphere

Kevin Hamilton
University of Hawaii, U.S.A.

1. Introduction

There are a number of features that distinguish the the dynamics of the tropical stratosphere and mesosphere from the dynamics elsewhere in the atmosphere and hence justify the focussed discussion of the low-latitude circulation that is presented in this School.

1. The small Coriolis parameter at low latitudes leads to a breakdown of the validity of the geostrophic approximation for the wind and invalidates the quasi-geostrophic theory that is so useful in explaining the large-scale circulation of the extratropical atmosphere. Another consequence of the small Coriolis parameter is that temperature observations (e.g. from satellite radiometers) are of limited use in inferring winds at low latitudes. Thus it is much harder in the tropics to diagnose the detailed dynamics from available observations, leading to more reliance on indirect theoretical and modelling approaches to study the general circulation.

2. The mean temperature structure in low latitudes is quite distinctive. Fig. 1 shows idealized temperature vs. height plots appropriate for the midlatitudes and for the equator. The equatorial region is notable for having higher temperatures in the lower troposphere and a colder/higher tropopause than is seen at higher latitudes. The height-latitude section of zonal-mean temperature in April determined from climatological observations is shown in Fig.2. The temperatures near the equatorial tropopause are extremely cold with observations ~190K.

3. The mean temperature structure combined with the the extremely nonlinear dependence of water vapor mixing ratio have important implications for the dynamics and chemistry of the atmosphere. In the lower and mid-troposphere the warm tropical temperatures and high water vapor mixing ratios lead to the possibility of very strong latent heat release. When combined with the small Coriolis

parameter this can lead to strong convectively-driven, strongly-divergent motions. At higher latitudes much more of the forcing of weather systems comes from conversion of potential energy associated with north-south temperature gradients.

Moist convection and precipitation have strong peaks near the equator. Fig. 3 shows the annual-mean precipitation estimated from observations. The low-latitude peak of rainfall is evident in this picture, but at a particular time the narrow meridional concentration of convection is even more pronounced. Over the oceans the convection is concentrated in the Intertropical Convergence Zone (ITCZ), which moves north and south following the sun through the year. As we will see, the time-variation of tropical convection excites vertically-propagating waves that are crucial to the dynamics of the tropical stratosphere.

4. The very cold temperatures in the tropical stratosphere are accompanied by extremely dry conditions. Fig. 4 shows the observed zonal-mean water vapor mixing ratios in the upper troposphere and lower stratosphere. The strong gradients near the tropopause are evident. Also noteworthy are the very small values throughout the stratosphere, particularly above the height of the equatorial tropopause, what has come to be called the "overworld". The implications of this dryness for the atmosphere (and indeed human life) is profound. With little water, the chemistry of the stratosphere is very different from the troposphere - there are no clouds (except in winter at very high latitudes), no precipitation, and much lower concentration of OH radical which catalyzes much of the homogeneous chemistry in the lower atmosphere.

Explaining the extremely dry conditions in the overworld seems possible only if it assumed that all the air entering the overworld enters through the equatorial tropopause. Fig. 5 shows a schematic of the large-scale circulation of the troposphere. There are many issues now being researched relating to how the

air actually gets into the stratosphere and exactly what determines how dry it is. However, it seems clear that any changes in the equatorial tropopause temperature will result in changes in the stratospheric water vapor, with important implications for stratospheric ozone chemistry. Interestingly, there is now considerable evidence of a rising trend in stratospheric water vapor. Fig. 6 shows the trend in annual mean water vapor concentration during 1981-2000 determined from balloon observations at Boulder (40N). So far there seems to be no clear connection of this trend with tropopause temperatures, but concerns about stratospheric chemistry response to tropospheric climate change represents one of the key practical reasons for studying the detailed dynamics of the equatorial middle atmosphere.

5. Perhaps the most distinctive features of the circulation in the neutral tropical middle atmosphere are the large-amplitude, long-period oscillations seen in the zonally-averaged flow. In particular, the winds and temperatures of the equatorial stratosphere undergo a very strong quasi-biennial oscillation (QBO) while the region from near the stratopause to the lowermost thermosphere displays a prominent semiannual oscillation (SAO). Fig. 7 shows a height-time section of monthly-mean zonal wind near 8N during 1962-1966 put together from balloonsonde measurements at Balboa (9N) to 30 km, and rocketsonde measurements from 30-56 km. This figure is strikingly different from the result one would obtain from higher latitude stations. Notably the annual (12-month) harmonic would dominate any higher latitude plot, but it is almost absent in this near-equatorial section. Instead, in the lower stratosphere the variation in wind seems to be interannual, while in the upper stratosphere there is a variation with period of 6 months.

In the lower stratosphere, say 16-35 km, the observations show a systematic variation from westerly (*eastward*) to easterly (*westward*) roughly every other year. In a later lecture we will look at observations of this phenomenon in much longer

data records. We will see that it has a variable period that averages about 28 months, leading to the term "quasi-biennial oscillation" (QBO) to describe it.

The QBO is a fascinating feature of the circulation, and is the closest thing to an internally-generated periodic phenomenon in the general circulation, to be contrasted with the genuinely periodic (and astronomically forced) diurnal and annual cycles. The QBO is much more regular than other widely-studied quasi-periodic phenomena in the atmosphere-ocean system, such as the Southern Oscillation or North Atlantic Oscillation.

The existence of the QBO demonstrates another way in which the tropical middle atmosphere is unique. The QBO is so regular that the winds in the tropical stratosphere are predictable months, and even years, in advance. The normal expectation based on tropospheric and midlatitude experience is that purely atmospheric phenomena have a limit of practical predictability of 2-3 weeks at most. The enhanced predictability of the tropical stratospheric wind is associated with the ability of the momentum in the zonally-averaged flow in this region to act as an effective long-term "memory" for the atmospheric circulation. In the next section of these notes we will try to explain the physics behind this phenomenon.

6. The equatorial atmosphere is also characterized by unique electrodynamic phenomena. The terrestrial geomagnetic equator is shown in Fig. 8 along with the geomagnetic latitude lines defined relative to this equator. The ionosphere displays fascinating phenomena that seem largely confined to low geomagnetic latitudes just as the neutral atmosphere has phenomena restricted to near the geographical equator. One early observation was that solar diurnal variations in the magnetic field tend to be enhanced near the geomagnetic equator. This has been attributed to the effects of a narrow ionospheric current confined near the geomagnetic equator and termed the "equatorial electrojet". See Forbes (1981) for a review. Another

prominent and widely-studied electrodynamic effect characteristic of low-latitudes is the instability leading to the dramatic "spread-F" phenomena. These and other related phenomena will be discussed in the second week of our School.

2. Basic Dynamical Considerations - The "Radiative Spring"

While not perfectly monochromatic, the equatorial QBO is the most spectacular example of a quasi-periodic oscillation in the atmosphere that is not astronomically forced (such as the annual and diurnal cycles). The QBO has a time-scale of deterministic predictability of months or even years - in strong contrast to meteorologists' usual picture of predictability in the atmosphere. An attempt will be made to explain this surprising feature within the simplest possible contexts. We will first consider the mean flow response to eddy forcing and show why the circulation in the equatorial region is likely to behave quite differently from that in higher latitudes. Then in a later lecture we will treat the very simple case of two internal gravity waves interacting with the mean-flow, to illustrate how long-period oscillatory equatorial mean-flow variations may be produced.

Governing Equations and Notation

We will assume a basic knowledge of the equations (momentum, continuity, thermodynamic, equation of state) governing fluid motion on a rotating sphere. Derivations can be found in standard textbooks such as Holton (1992). To study the large-scale circulation it is often useful to make the "primitive equation" approximations, replacing the vertical momentum equation by hydrostatic balance and considering only the Coriolis effects associated with the component of the earth's rotation vector parallel to the local vertical (termed the "Coriolis parameter", $f=2\Omega\cos(\theta)$).

We will adopt "standard" meteorological notation where u, v, w represent the zonal, meridional and vertical components of the winds, with westerly (eastward), southerly (northward), and upward velocities defined as positive. Location is de-

fined by coordinates θ, the latitude, ϕ, longitude (east as positive) and z, the geometrical height above the earth's surface. Fig. 10 shows a schematic of this coordinate system.

In many instances it it convenient to approximate the full spherical geometry by planar geometry, indeed several different "flavours" of this kind of approximation are useful. In a simple instance we will be considering models of the equatorial flow that ignore rotation and replace the zonal direction by Cartesian direction, x, with an assumption of periodic boundary conditions in this direction. Another approach, the so-called "f-plane" retains the Coriolis force, but with a constant value of the Coriolis parameter, f. Finally, the equatorial region is often treated with an "equatorial β-plane", in which the Coriolis parameter is taken to depend linearly on the meridional distance from the equator, $f=\beta y$.

The decay of mean atmospheric density with height has important implications for many aspects of atmospheric dynamics. However, it also introduces complications in the mathematical treatment of the equations. In some instances it is useful to make one further assumption, familiar to oceanographers, namely the Boussinesq approximation. In the Boussinesq approximation the atmospheric density is treated as constant except in the hydrostatic equation (where varying density affects the horizontal gradients of pressure which drive the circulation). For quantitative results in the atmosphere, the Boussinesq approximation is almost always inadequate, but - if the system studied is hydrostatic - some of the mathematical simplification of the Boussinesq approximation can be retained in the compressible equations by transforming the vertical coordinate from geometrical height to something proportional to the logarithm of pressure. We will discuss these "log-pressure" coordinates later.

Zonal Mean and Eddies

For many purposes it is useful to divide the atmospheric circulation into its zonal-mean component and all deviations from the zonal-mean (which we refer to as the "eddy" component). Thus, for example, the zonal wind can be written as

$$u = \overline{u} + u'$$

where the overbar refers to an average around a latitude circle (or equivalently over the full x-domain in a Cartesian geometry model).

The separation between the zonal-mean and eddy components can be applied to the governing equations. Consider, for example, the zonal momentum equation in the very simple 2D (x-z) Boussinesq case:

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - w\frac{\partial u}{\partial z} - \frac{1}{\rho_o}\frac{\partial P}{\partial x}$$

where P is the pressure and ρ_o is the (constant) mean density. Substituting the zonal-mean, eddy decomposition and then averaging over x yields an equation for the rate of change of the zonal-mean zonal wind:

$$\frac{\partial \overline{u}}{\partial t} = -\overline{w}\frac{\partial \overline{u}}{\partial z} - \overline{w'\frac{\partial u'}{\partial z}}$$

the Boussinesq version of the continuity equation is

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0$$

When averaged over x the continuity equation implies that

$$\overline{w} = 0$$

and

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0$$

This can then be used to rewrite the zonal-mean momentum equation in the familiar form

$$\frac{\partial \overline{u}}{\partial t} = -\frac{\partial \overline{u'w'}}{\partial z}$$

Here the acceleration of the mean flow is forced by the effect of the eddies which can be regarded as a convergence of a "Reynolds stress" or "vertical eddy flux of zonal momentum". It emphasizes the role of eddies as "catalysts" transferring mean momentum from one level to another. Similar analysis can be performed in a full three-dimensional case and for the thermodynamic equation as well. The result is a set of equations for the zonal-mean variables that are "forced" by the eddy fluxes of momentum and heat.

Linearized Zonal-Mean Model

We will discuss here what is perhaps the simplest useful model for the mean circulation in the middle atmosphere. In particular, we consider the steady-state response of the mean flow to a specified eddy forcing. The model has the realistic feature that diabatic effects act to restore radiative equilibrium. For simplicity we will model this radiative effect with a Newtonian cooling towards a state with

no meridional temperature gradient (it is easy to generalize to the case where the "radiative equilibrium" state has a meridional gradient). Also, for mathematical tractability, we will consider only the Boussinesq equations on an f-plane (periodic boundary conditions in the x-direction will represent the zonal periodicity of the real atmosphere). The base state static stability is assumed constant and some non-linear advection terms are ignored. With all these assumptions the governing equations for the zonally-averaged circulation reduce to:

$$-f\bar{v} = \bar{F}_{eddy}$$

$$f\bar{u} = -\frac{1}{\rho_0}\frac{\partial \bar{P}'}{\partial y}$$

$$\frac{\partial \bar{P}'}{\partial z} = -g\bar{\rho}'$$

$$\bar{w}\bar{\rho}_z = -\alpha\bar{\rho}'$$

$$\frac{\partial \bar{v}}{\partial y} + \frac{\partial \bar{w}}{\partial z} = 0$$

where

$$\bar{\rho} = \rho_0 + z\bar{\rho}_z + \bar{\rho}'$$

and

$$\bar{P} = \bar{P}_0 + \bar{P}'$$

Note that we have ignored the eddy terms in the thermodynamic equation for simplicity. [For those familiar with the Transformed Eulerian Mean (TEM)

formulation (e.g., Andrews et al., 1987) another approach is to take these same equations and replace $\overline{w}$ and $\overline{v}$ with the TEM velocities $\overline{w^*}$ and $\overline{v^*}$ - this accounts for the effects of the eddy heat transports in a consistent way.]

The thermodynamic equation can be rewritten as

$$N^2 \bar{w} = \frac{\alpha g}{\rho_0} \bar{\rho}'$$

where

$$N^2 = -\frac{g \bar{\rho}_z}{\rho_0}$$

Here, f is the Coriolis parameter (a constant), u is the zonally-averaged zonal velocity, $\bar{v}$ and $\bar{w}$ are the meridional and vertical components of the mean meridional circulation, N is the buoyancy frequency, $\bar{P}$ is the zonal-mean pressure and $\bar{\rho}$ is the zonal-mean density. The density is the sum of a background state and the perturbation $\tilde{\rho}$. The background state density is the sum of the constant Boussinesq mean density, $\bar{\rho}_0$, and a linear decrease with height. The background pressure $\bar{P}_0$ is assumed to be in hydrostatic balance with the background density. The body force per unit mass associated with the eddy Reynolds stress convergence [or Eliassen-Palm flux divergence in the TEM case] is $\bar{F}_{eddy}$, and the Newtonian cooling has a relaxation coefficient α.

These linear, constant coefficient equations are easily solved if we write the solution in terms of Fourier harmonics. Consider one harmonic:

$$\bar{F}_{eddy}, \bar{P}', \bar{\rho}', \bar{u}, \bar{v}, \bar{w} \propto e^{i(ky+mz)}$$

Substitution into the equations of motion gives

$$\bar{v} = -\frac{\bar{F}_{eddy}}{f} \tag{1}$$

$$f\bar{u} = -\frac{ik\bar{P}}{\rho_0} \tag{2}$$

$$im\bar{P} = -g\bar{\rho}' \tag{3}$$

$$\frac{\rho_0 N^2 \bar{w}}{\alpha} = g\bar{\rho}' \tag{4}$$

$$ik\bar{v} + im\bar{w} = 0 \tag{5}$$

Equations (3) and (4) can be combined to give

$$\frac{\rho_0 N^2 \bar{w}}{\alpha} = -im\bar{P}$$

or

$$im\bar{w} = \frac{\alpha}{\rho_0 N^2} m^2 \bar{P}$$

Substituting this into (5) and using (1) to eliminate $\bar{w}$ and $\bar{v}$ gives

$$-ik\frac{\bar{F}_{eddy}}{f} + \frac{\alpha m^2}{\rho_0 N^2}\bar{P} = 0$$

or

$$\bar{P} = \frac{ik\rho_0 N^2}{f\alpha m^2}\bar{F}_{eddy}$$

From equation (2) then

$$\bar{u} = -\frac{ik}{f\rho_0}\bar{P} = \frac{k^2 N^2}{f^2 m^2 \alpha}\bar{F}_{eddy} \tag{6}$$

The response of $\bar{u}$ to the forcing is proportional to $1/f^2$, which implies that, for given spatial scales of the forcing, the constraint of the "radiative spring" on the mean flow is very much less effective near the equator than at higher latitudes. One factor of $1/f$ arises because the temperature perturbation associated with a given wind perturbation scales as f, and there is an additional factor of $1/f$ because the induced meridional circulation affects the zonal momentum equation through the $f\bar{v}$ term.

Since m^2 in equation (6) is equivalent to a second derivative in z on the zonal velocity, this acts as a vertical diffusion, i.e.

$$\frac{\partial^2 \bar{u}}{\partial z^2} = -\frac{k^2 N^2}{f^2 \alpha} \bar{F}_{eddy}$$

or in general, with a non-resting background state,

$$\frac{\partial^2 (\bar{u} - \bar{u}_{rad})}{\partial z^2} = -\frac{k^2 N^2}{f^2 \alpha} \bar{F}_{eddy}$$

where $\bar{u}_{rad}$ is the zonal mean wind that is in thermal wind balance with the radiative equilibrium temperature structure. The radiative spring acts as an "effective vertical diffusion" diffusing away any deviations from the radiative equilibrium state.

This process has a characteristic time-scale $\bar{u}/F_{eddy}$ or $\frac{k^2 N^2}{f^2 m^2 \alpha}$. At sufficiently small f this will become very long, and it is unlikely that the steady-state solution has much relevance. Near the equator the dominant balance is presumably between $\partial \bar{u}/\partial t$ and $\bar{F}_{eddy}$. Thus at low latitudes the mean flow can respond to eddy driving with sustained accelerations. The momentum in the zonal-mean flow at any time has responded to the accumulated effect of the eddy driving and serves as a "memory" for the atmospheric circulation.

This treatment of the mean flow effects in the tropical middle atmosphere has been highly idealized, of course. Some aspects involved in more realistic models are discussed by Dunkerton (1991). For example, Dunkerton notes that in a compressible atmosphere the effective diffusion does not act symmetrically in the upward and downward directions as it does in the simple Boussinesq case considered here. The work of Haynes et al. (1991) is also relevant. This is a more general discussion of the response of the zonal-mean circulation to imposed forces, including such aspects as the nonlinear advection of the relative mean momentum by the mean meridional circulation, and the spherical geometry. Their results also demonstrate

the difficulty in finding a steady-state mean flow response at low latitudes that would be relevant to the real atmosphere.

3. Phenomenology of Tropical Circulation

3.1 Historical Introduction

The first scientific knowledge of the winds in the tropical stratosphere was obtained from observations of the motion of the aerosol cloud produced by the eruption of Mt. Krakatoa (modern-day Indonesia) in August 1883. The optical phenomena caused by the aerosol were remarkable enough that their first appearance was widely noted. The UK Royal Society collected observations from over 30 locations in the tropics and plotted the motion of the edge of the aerosol cloud (see Fig. 10). The regular westward motion is evident, and these observations imply a mean easterly wind velocity between about 31 and 34 m-s^{-1}.

The wind in the tropical lower stratosphere was first measured with pilot balloons in 1908 by von Berson at two locations in equatorial East Africa. Over the next three decades these observations were followed by sporadic measurements at a number of tropical locations (see Hamilton, 1998, for a review of these early observations). The results sometimes indicated easterly winds and sometimes westerly winds, a state of affairs reconciled at the time by assuming that there was a narrow ribbon of westerlies (the "Berson westerlies") embedded in the prevailing easterly current revealed by the Krakatoa observations (e.g., Palmer, 1954).

Regular balloon observations of the lower stratospheric winds in the tropics began at a number of stations in the early 1950's. By the end of the decade it was obvious that both the easterly and westerly regimes at any height covered the entire equatorial region, but that easterlies and westerlies alternated with a roughly biennial period (Veryard and Ebdon, 1961; Reed et al., 1961). Initially it was thought that the period of the oscillation might be exactly two years, but as measurements accumulated it soon became clear that the period of oscillation

was somewhat irregular and averaged over 2 years. By the mid-1960s the term "quasi-biennial oscillation" (QBO) had been coined to denote this puzzling aspect of the stratospheric circulation.

3.2 Modern Observations of the QBO

Height-Time Dependence at Low latitudes

Much of what we know about the QBO has come from analysis of operational balloon-borne radiosonde observations, although such observations are generally restricted to a ceiling near 10 hPa (~30 km) and the number of near-equatorial stations is relatively limited. Fig. 11 shows an 8-year time series of monthly-mean 30 mb zonal wind computed simply by averaging twice-daily balloon observations at Singapore (1.4N). This illustrates many of the key features of the QBO. Note that the time series is clearly dominated by an alternation between easterly and westerly wind regimes roughly every other year. The extremes in the prevailing winds vary somewhat from cycle to cycle, but the peak easterly (in the monthly mean) usually exceeds 30 m-s^{-1}, while the westerly extreme is usually between 10 and 20 m-s^{-1}. The time series has a rough square-wave character with rapid transitions (~2-4 months) between periods of fairly constant prevailing easterlies or westerlies.

A hand-drawn height-time contour plot of the monthly-mean zonal wind from 100 hPa (~17 km) and 10 hPa (~30 km) for four years at Canton Island (2.8N) is shown in Fig. 12. This shows that the wind reversals appear first at high levels and then descend. At any level the transition between easterly and westerly regimes is rapid, so that the transitions are also associated with strong vertical shear.

The height-time evolution of the monthly-mean wind near the equator for over four decades is shown in Fig. 13 (an updated version of the figure published in

Naujokat, 1986) Unfortunately there has not been a single, continuously-operating, near-equatorial station taking appropriate observations over this period. This results shown in Fig. 13 were spliced together from records at Canton Is. (2.8N), Gan (0.7S) and Singapore (1.4N). These near-equatorial data suggest that the mean period of the oscillation is close to 28 months and varies between about 20 months and 36 months. Much of the variability in period is associated with the changes in the length of the easterly phase at and above about 30 hPa and in the westerly phase at and below about 50 hPa. The maximum amplitude occurs near 30 mb and the amplitude drops off to small, but apparently still detectable, values near the tropopause (~17 km). The dropoff in QBO amplitude above 30 mb is very gradual and the oscillation is still very strong at the 10 mb level. At almost all levels and all times the easterly-to-westerly transitions are more rapid than *vice versa*, and the associated westerly shear zones are considerably more intense than the easterly shear zones.

There is only a very limited network of radiosonde stations near the equator regularly reporting stratospheric winds. An assumption implicit in most observational studies is that the prevailing winds near the equator are essentially zonally-symmetric. Belmont and Dartt (1968) tried to check this with available radiosonde data and concluded that the QBO was indeed very nearly zonally-symmetric up to 50 hPa (above this level they felt they had inadequate data to verify this). Recently the advent of Doppler-radiometer observations of horizontal winds by the High Resolution Doppler Imager (HRDI) instrument on the Upper Atmosphere Research Satellite (e.g., Ortland et al., 1996) has provided another opportunity to examine this issue. Ortland (1997) finds that there may be some modest (~10 $m\text{-}s^{-1}$) zonal asymmetries in the monthly-mean wind at the equator in the westerly phase of the QBO, particularly at and above 10 mb. This issue will be discussed

further below, but, to first order, the assumption of a zonally-symmetric QBO is appropriate.

The vertical structure of the equatorial QBO above the operational balloon ceiling was initially studied by rocket-borne sondes. Fig. 14 shows the zonal wind observations from actual individual rocket soundings for 25 months at Ascension Island (8S). Results at 28 km show an obvious QBO signal, but at higher altitudes the QBO amplitude clearly drops off while the semiannual (6-month) variation dominates by 48 km. Fig. 15 shows a long height-time series of monthly-mean zonal wind near the equator constructed by using the balloon observations at Canton Is., Gan and Singapore below 31 km, and an average of rocketsonde observations at Kwajalein (9N) and Ascension Is. (8S) above 31 km. The top panel shows the deseasonalized values, while the bottom panel shows a band-passed version including only periods between 9 and 48 months. The filtering in both cases eliminates the mean semiannual variation, allowing the QBO to stand out. The drop off of the QBO amplitude in the upper stratosphere is evident.

The advent of the HRDI satellite observations has allowed the structure of the wind QBO at still higher levels to be investigated. The top panel of Fig. 16 shows the height-time section of equatorial zonal-mean winds estimated from the HRDI observations from Jan. 1992-July 1997 and from 10-100 km (with a gap near 45 km). The bottom panel shows the same section, but with the annual and semiannual harmonics removed, emphasizing the interannual variations. The QBO in the lower-middle stratosphere evident. Also apparent is a less smooth interannual variation peaking around 85 km. The vertical wiggly lines emphasize the apparent connection of the variations near 85 km with those in the lower stratosphere. Whenever there are easterly (westerly) anomalies through a thick layer in the lower stratosphere, strong westerly (easterly) anomalies appear near

the mesopause. The high level variations are often referred to as the "mesopause QBO (MQBO)".

Meridional Modulation

While the QBO in zonal-wind has little longitudinal structure, it does have a very strong dependence on latitude. Fig. 17, taken from Dunkerton and Delisi (1985), shows the evolution of 30 hPa zonal winds during 1964-1983 determined for various latitudes by averaging balloon observations at clusters of stations located near the indicated latitude. So, for example, the results shown for the equator are actually averages over Singapore (1.4N), Gan (0.7S) and Nairobi (1.3N), and those labelled "19N" are averages over Hilo (19.7N), Wake (19.3N) and Grand Cayman (19.3N). The strongly near-equatorial character of the QBO is evident from even a cursory inspection of these data. Near the equator the data show almost nothing but the QBO, by 19 or 20 degrees latitude the QBO is hardly apparent, and is swamped by the seasonal cycle.

Fig. 18 shows a determination of the amplitude (solid contours) and phase (dashed contours) of the QBO in zonal wind as a function of height and latitude in the tropical stratosphere. It was derived by R.J. Reed who performed a least-squares fit of a single 26-month harmonic to about 8 years of balloonsonde zonal wind data at a number of low latitude stations (see Wallace, 1973). Unlike purely periodic phenomena the determination of amplitude and phase for the QBO is not straightforward. Other studies have followed Reed in fitting a fixed-period harmonic to the data. Such a procedure will not provide stable results - amplitudes will drop the longer the record analyzed, so Reed's application to a relatively brief record may be essentially optimal. Fig. 18 shows that the amplitude has a peak centered squarely on the equator and a roughly Gaussian dropoff in latitude with an e-folding width of between 13 and 15 degrees of latitude. The phase lines

are remarkably regular indicating a steady downward propagation of about 2 km-month^{-1} and very little phase variation in latitude.

The single harmonic approach obscures some of the characteristic details of the meridional evolution of the zonal wind through the QBO cycle. In particular, the initial westerly accelerations away from the easterly extreme tend to be very narrowly-confined near the equator, and then the "westerly tongue" expands so that the extreme westerly jet is of about the same width as the easterly jet. This is illustrated in Fig. 19 (from Hamilton, 1987) that shows the zonal wind at 30 hPa month-by-month from October 1979 through February 1980.

Reed's analysis shows the amplitude of the QBO dropping to near zero at 25 degrees latitude. Later on we will discuss the influence of the QBO at higher latitudes.

QBO in Temperature

A QBO in temperature has also been clearly observed. The bottom curve in Fig. 20 shows the temperature anomalies (deviations from long-term climatology) averaged over the 30-50 hPa layer based on Singapore balloon observations. A QBO with peak-to-peak amplitude of around 8°C is evident. The top curve shows the vertical shear in the zonal wind over the 30-50 hPa layer, again from the Singapore balloon observations. The strong positive correlation between the two time series is clear, and can be explained in terms of the thermal wind balance discussed earlier. In general, the usefulness of the geostrophic approximation breaks down at low latitudes, of course, but in fact the *zonal-mean* component of the circulation should be close to geostrophic balance.

Period Modulation of the QBO

When a long record of QBO winds is examined, some order appears in the

variability from cycle-to-cycle. The QBO amplitude appears to be fairly constant, but the period length varies significantly with time. This variability has some very characteristic features. Below about 30 hPa the easterly phases are of nearly constant length, but the westerly phases vary by a factor of two. Fig. 21 shows a determination of the length of the westerly phases at 45 hPa through 4 decades. As noted by Salby and Callaghan (2000) the period seems to vary systematically with a quasi-decadal time scale. Salby and Callaghan speculated that the 11-year solar cycle may be somehow controlling the QBO period. Later on we will consider other possible explanations for the variability of QBO period.

Above about 30 hPa, the situation reverses: the westerly phases are of roughly comparable length, but the easterly phases are quite variable.

Another interesting aspect of the QBO is the subtle effect of the annual cycle. As noted earlier, the long-term mean period of the QBO is incommensurate with the annual period or any obvious harmonic or subharmonic. However, the annual cycle does have some statistical connection with the observed QBO-related wind transitions. Fig. 22 is a histogram of the month-of-year in which the 50 hPa westerly-to-easterly and easterly-to-westerly zero crossings were observed. There is a strong concentration of such transitions in the May-July period of the year. Again this subtle connection with the annual cycle is something that needs to be explained dynamically.

3.3 Observations of the Semiannual Oscillation

The first knowledge of the wind field above the ceiling for balloon observations was obtained with rocket soundings. Fig. 14 shows all the zonal wind observations from rocket soundings during a two year period at Ascension Is. (8°S). In the middle stratosphere the QBO appears quite clearly (at least to $\sim$ 40 km) However,

at upper stratospheric levels the QBO is dominated by a shorter period oscillation. We now know that this is a semiannual oscillation (SAO). Unlike the QBO, the SAO is very clearly phase-locked to the calendar. Near the stratopause the easterly extremes are reached in January and July and the westerly extremes around April and October. Fig. 23 shows a climatological annual march of the zonal wind deduced from rocket observations at a number of sites. At the equator the semiannual variation clearly dominates in a thick layer around the stratopause. The vertical structure of the westerly accelerations displays a downward propagation that is similar to the QBO wind reversals. The easterly accelerations are more uniform in height. The stratopause SAO has been seen in rocket observations of zonal winds, rocket observations of temperatures (e.g., Garcia et al., 1997; Dunkerton and Delisi, 1997), satellite radiometer measurements of temperature (Hitchman and Leovy, 1986; Delisi and Dunkerton, 1988) and the recent HRDI Doppler radiometer measurements of winds (Burrage et al., 1996).

Fig. 24 shows the height-latitude structure of the amplitude of the semiannual component of the zonal wind variation determined from rocketsonde observations at several stations. Like the QBO, the SAO appears to have a similar equatorially-trapped structure. This is shown as well in Fig. 25 which displays the height-latitude structure of the amplitude of the semiannual component of zonal-mean temperature determined from satellite radiometer observations.

Using the limited number of rocket soundings available at very high altitudes, Hirota (1978) and Hamilton (1982a) were able to show that the amplitude of the SAO at low latitudes drops in the lower mesosphere, but rises again to a maximum near the mesopause. They also found that the SAO at the mesopause is nearly 180° out of phase with the stratopause SAO. Fig. 26 shows Hirota's hand-contoured plot of the zonal wind measured at Ascension Is. each month from 1970-1972. In

each month data up to the usual meteorological rocketsonde ceiling (~60 km) are available, and in about 18 months there are some data from special rockets that go up to 90 km. The long-term mean and annual cycle at each height have been removed. This figure shows the stratopause SAO centered around 50 km, and also a second maximum in SAO amplitude around 80-85 km. This "mesopause SAO" is nearly 180 degrees out-of-phase with the stratopause SAO.

These early rocket observations of the mesopause SAO have been confirmed recently with the wind measurements from the HRDI instrument. This can be seen in the raw HRDI timeseries in the top panel of Fig. 16. Fig. 27 (from Garcia et al., 1997) shows the equatorial zonally-averaged zonal wind as a function of height and time of year composited from the HRDI measurements in the height range 65-110 km. The prominent SAO at the mesopause seen in these data has amplitude comparable to that observed near the stratopause.

3.4 Observations of High-Frequency Waves in the Tropical Atmosphere

The discussion thus far has focussed on analysis of monthly-mean observations. In addition to the slow variations that are revealed in the monthly-means, the tropical stratosphere and mesosphere display significant day-to-day and hour-to-hour variations. The top panel in Fig 28 shows the height-time section of zonal wind at Balboa (9N) for 6 months plotted using daily balloon data, while the bottom panel shows the same thing, but for temperature observations. The overall descending QBO phases are evident in each of these sections. However, just as prominent are oscillations of periods appearing to be ~2-20 days. These variations also display a predominant downward phase progression. As we will see, these higher frequency waves are thought to be essentially internal gravity waves, modified by the effects

of the earth's rotation. In this interpretation, the downward phase propagation corresponds to upward wave energy (or, more exactly, wave activity) propagation, suggesting a dominant tropospheric source for the waves.

Similar waves are observed higher up as well. Fig. 29 (from Hirota, 1978) shows the difference in u and T between two rocket soundings at Ascension Is. separated by two days. There are wiggles at various scales apparent, but most prominent is a variation with vertical wavelength of about 20 km which looks very coherent between the zonal wind and temperature.

The wave activity in the low-latitude stratosphere is observed to be significantly modulated by the QBO. The contours in Fig. 30 shows the time power spectra of u and T variations calculated from twice-daily balloon observations over the 20-25 km layer at Singapore plotted each month for several years. Also shown as the heavy line is the time series of monthly-mean zonal wind. The spectra show maxima around periods of 10 days, but these peaks become much more prominent near the easterly extreme of the QBO. As we shall see, the coupling of the high frequency waves with the mean flow is thought to be crucial for the generation of the QBO.

Balloon and rocket observations are usually taken at most twice-pre-day. Other kinds of observations such as radars and lidars show that significant wave variations in the middle atmosphere occur at even higher frequencies.

There is considerable data concerning the height-time variations of middle atmospheric wind and temperature at single stations. Rarer is useful information on the horizontal structure of the wave variations, and even rarer are observations that reveal the time-horizontal structure of middle atmospheric variations. At fairly low frequency (say many days to weeks) there is some indication of fairly coherent, long

zonal scale structures. Some of the early studies of waves in the tropical stratosphere focussed on planetary-scale waves that can be examined by using the daily balloon observations from widely separated radiosonde stations (see Wallace, 1973, for a review of this early work). It is now fairly clear that such coherent, planetary-scale waves are only a part of the full spectrum of waves in the middle atmosphere. Aircraft observations suggest that there is a rather shallow continuous spectrum of horizontal variations out to scales as small as 10 km. Fig. 31 shows mean horizontal power spectra in the upper troposphere calculated from observations from thousands of instrumented commercial aircraft. One key unresolved question that will arise in various contexts in these lectures is the relative importance of relatively small-scale gravity waves and the lower-frequency planetary-scale waves in the general circulation of the tropical middle atmosphere. Unfortunately, the currently available observations do not provide definitive information about this issue.

4. Simple Model of Wave-Driving of the Mean Flow

4.1 Introduction

The considerations outlined in Section 2 above explain the propensity of the zonal-mean tropical middle atmospheric circulation to undergo slow variations. The basic explanation for why the variations can take the form of the quasi-regular QBO wind reversals was provided by Lindzen and Holton (1968) and Holton and Lindzen (1972). They showed that the eddy driving of the mean flow from a combination of eastward and westward travelling, vertically-propagating, internal gravity waves or equatorial planetary waves could produce a long-period mean flow oscillation with many of the observed characteristics of the QBO. This section will discuss simple models that demonstrate how waves interacting with the mean flow may give rise to the long-period QBO phenomenon. later lectures will discuss related issues such as the details of the wave dynamics in the equatorial region (in particular the role of planetary rotation), and the tropospheric forcing mechanisms for these waves.

Plumb (1977) presented a slightly simplified version of the Holton-Lindzen model which simply considers the mean flow accelerations induced by two internal gravity waves. We will describe the Plumb model in some detail here. The first step will be a brief discussion of the properties of linear internal gravity waves.

4.2 Two-Dimensional Internal Gravity Waves

We consider the case of purely 2-D, non-rotating flow in a domain with periodic boundary conditions in the x-direction. For simplicity we will treat the problem within the Boussinesq and hydrostatic approximations. Consider then the equa-

tions linearized about a time-independent, zonal-mean state characterized by mean flow $\bar{u}(z)$, and mean static stability $N(z)$.

$$\frac{\partial u'}{\partial t} + \bar{u}\frac{\partial u'}{\partial x} + w'\frac{\partial \bar{u}}{\partial z} = -\frac{1}{\rho_0}\frac{\partial P'}{\partial x} - \alpha u' \tag{7}$$

$$\frac{\partial \rho'}{\partial t} + \bar{u}\frac{\partial \rho'}{\partial x} - \frac{\rho_0 N^2}{g}w' = -\alpha \rho' \tag{8}$$

$$\frac{\partial P'}{\partial z} = -g\rho' \tag{9}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0 \tag{10}$$

Where we have allowed for the possibility of wave dissipation by inclusion of the linear relaxations in the momentum and thermodynamic equations. Now assume travelling wave solutions

$$u' = U(z)e^{i(kx-\omega t)}$$

$$w' = W(z)e^{i(kx-\omega t)}$$

etc.

Substitution into equations (7)-(10), and assuming that $\alpha = 0$, gives

$$\frac{\partial^2 W}{\partial z^2} + \left(\frac{N^2 k^2}{(\omega - k\bar{u})^2}\right)W + \left(\frac{k}{(\omega - k\bar{u})}\right)\frac{\partial^2 \bar{u}}{\partial^2 z}W = 0$$

If we now assume that the background state does not vary with height so that $N^2(z)$ and $\bar{u}(z)$ are constant, then

$$\frac{\partial^2 W}{\partial z^2} + \left(\frac{k}{\omega - k\bar{u}}\right)^2 N^2 W = 0 \tag{11}$$

The solution to this equation is of the form $W \propto e^{imz}$ and the resulting dispersion relation is

$$m^2 = \frac{N^2 k^2}{(\omega - \bar{u}k)^2} = \frac{N^2}{\hat{c}^2}$$

where $\hat{c}$ is the intrinsic (Doppler-shifted) horizontal phase speed for the wave. Waves with high intrinsic horizontal phase speed have small m, or long vertical wavelength. Also, near critical levels, where the horizontal phase speed matches the background flow, the vertical wavelength should become very small.

The vertical group velocity is given by

$$c_{g_z} = \frac{\partial \omega}{\partial m} = \frac{\partial(\omega - \bar{u}k)}{\partial m}$$

Using as our convention that the intrinsic frequency of the waves is positive, the appropriate root must be taken:

$$\omega - \bar{u}k = \pm \frac{Nk}{m}$$

where the sign of k and m determine the sense of phase propagation. This gives a vertical group velocity

$$c_{g_z} = \mp \frac{Nk}{m^2} = \mp \frac{(\omega - \bar{u}k)^2}{Nk}$$

This implies that for $m < 0$ (i.e. for waves with downward phase propagation), $c_{g_z} > 0$ and so there is upward group propagation. Also, higher intrinsic horizontal phase speed waves have larger vertical group velocity. As $\omega \rightarrow \bar{u}k$ the vertical group velocity $c_g \rightarrow 0$.

One can also easily show for $c_{g_z} > 0$, that $\overline{u'w'} > 0$, if $k > 0$, and that $\overline{u'w'} < 0$, if $k < 0$. That is, for upward-propagating wave energy with eastward (westward) intrinsic phase speed the eddy momentum flux is positive (negative).

For the general case when the mean state varies with height, the full equation for $W(z)$ is needed. However, if we assume that the mean state varies sufficiently slowly with height relative to the wave phase, i.e.

$$\frac{k}{|\omega - \bar{u}k|}\frac{\partial^2 \bar{u}}{\partial z^2} \ll m^2$$

then the dominant balance is just that described by equation (11). Again if $\bar{u}$ and N^2 vary only slowly with z over a vertical scale m^{-1}, then we can use the WKB solution

$$W(z) = W(z_0)\Big(\frac{kN}{\omega - \bar{u}k}\Big)^{1/2} exp\Big(\int_{z_0}^{z} -\frac{ikN}{\omega - \bar{u}k}dz\Big) \tag{12}$$

where we have considered a wave with upward group velocity. In this case, one can go back to the governing equations and - assuming a slowly-varying mean state - show that $\overline{w'w'}(z) = constant$, which is consistent with the Eliassen-Palm theorem (i.e. that for steady, unforced and undissipated linear gravity waves the divergence of the Reynolds stress should be zero; see Eliassen and Palm, 1961).

Including the dissipation terms $-\alpha u'$ and $-\alpha \rho'$, in the above derivation simply replaces ω by $\omega + i\alpha$. In this case, the integral in equation (12) becomes

$$\int_{z_0}^{z} -\frac{kN}{\omega + i\alpha - \bar{u}k}dz$$

$$= \int_{z_0}^{z} \frac{ikN}{(\omega - \bar{u}k)(1 + \frac{i\alpha}{\omega - \bar{u}k})}dz$$

Assuming weak dissipation (i.e. that $\alpha \ll (\omega - \bar{u}k)$) and using a binomial expansion, the integral becomes

$$\approx \int_{z_0}^{z}\Big(-\frac{ikN}{(\omega - \bar{u}k)} + \frac{i^2\alpha kN}{(\omega - \bar{u}k)^2}\Big)dz$$

$$= \int_{z_0}^{z} -\frac{ikN}{(\omega - \bar{u}k)}dz - \int_{z_0}^{z}\frac{\alpha kN}{(\omega - \bar{u}k)^2}dz$$

so that the solution has both an oscillating character and an overall amplitude modulation with an exponential decay with height. In this case, the Reynolds stress is no longer constant but

$$\overline{u'w'}(z) \sim exp\Big(\int_{z_0}^{z} -\frac{2\alpha k N}{(\omega - \bar{u}k)^2} dz\Big) = exp\Big(\int_{z_0}^{z} -\frac{2\alpha}{c_{g_z}} dz\Big)$$

Therefore, the slower the vertical group velocity, the longer the damping can act on the waves, the more rapid the decay of the Reynolds stress, and the stronger the local mean flow driving. Note that in general α can be a function of z. If the damping term is omitted from the momentum equation but retained in the thermodynamic equation, it is easy to show that

$$\overline{u'w'}(z) = \overline{u'w'}(z_0) exp\Big(\int_{z_0}^{z} -\frac{\alpha k N}{(\omega - \bar{u}k)^2} dz\Big). \tag{13}$$

This is the case considered by Plumb (1977) in his original paper, and it is possibly a reasonable treatment of the stratosphere in the sense that large-scale motions may be regarded as rather inviscid, while the effects of radiative transfer will definitely act to damp the temperature perturbations associated with wave motions. The effectiveness of radiative transfer in damping waves is known to depend on the vertical wavelength, with short wavelengths more strongly damped (e.g. Fels, 1982). In the limit of very long vertical wavelength perturbations the timescale for the radiative damping can be estimated to range from $\sim$ 100 days in the lowermost stratosphere to $\sim$ 5 days near the stratopause. Fig. 32 shows a typical profile used for the thermal damping rate in the stratosphere. The dissipation rates given in this figure are for long vertical wavelengths and thus are something of an underestimate of those appropriate for the waves (with wavelengths roughly 3-15 km) thought to be important in driving the mean flow accelerations in the tropical stratosphere (e.g. Wallace, 1973).

4.3 One-Wave Plumb model

Consider a system with a zonal-mean flow $\bar{u}(z,t)$ and one internal gravity wave, with wavenumber k and frequency ω, forced at the lower boundary. Assume that there is a constant effective vertical diffusion K acting on the mean flow. Then the zonal-mean zonal momentum equation is

$$\frac{\partial \bar{u}}{\partial t} = -\frac{\partial}{\partial z}(\overline{u'w'}) + K\frac{\partial^2 \bar{u}}{\partial z^2}$$

We follow the standard treatment of wave-mean flow interaction problems in assuming that the eddy fluxes can be adequately computed from linear wave theory. In addition we assume that the vertical wavelength of the wave is sufficiently short that the WKB scaling discussed earlier applies, and that the mean flow changes so slowly that the steady wave solution can be used. In this case we saw that the wave momentum flux is given by equation (13). In that equation z_0 is height of the lower boundary, which can be regarded as corresponding to the tropopause. All the wave parameters, including the momentum flux at z_0, must be specified, along with $N(z)$ and $\alpha(z)$. Using boundary conditions on the mean flow such that $\bar{u}(z_0) = 0$ and $\frac{\partial \bar{u}}{\partial z} = 0$ at the top boundary, one can integrate numerically very easily, calculating the wave flux, then using the flux convergence as forcing in the mean flow equation for some brief timestep, then recalculating the wave fluxes with the updated mean flow etc.

Fig. 33 shows the mean flow evolution in an integration of this model. The imposed gravity wave has a phase speed of +30 m-s^{-1} and a period of about 15 days. With an initial zonal wind $\bar{u}(t_0) = 0$, and with radiative damping that increases with height, the initial acceleration of the zonal wind is strongest higher up. As the wind is accelerated, however, the intrinsic phase speed of the wave decreases, leading to increased dissipation at lower levels. Eventually the wave is

almost choked off from reaching the upper levels and the accelerations there drop almost to zero. The net effect is to produce a mean flow jet with maximum that descends with time, very much like the wind regimes in the QBO. After 80 days the model is approaching a steady-state with the maximum mean flow comparable to the wave phase speed. At this point the wave is very strongly absorbed in the lowest few km of the domain and the eddy forcing is counteracted by the mean flow diffusion of momentum down into the lower boundary.

The model described here is easily generalized to include the effects of compressibility (e.g., Plumb, 1977). The WKB solution for the Reynolds stress associated with the vertically propagating gravity wave becomes

$$\rho_0(z)\overline{u'w'}(z) = \rho_0(z_0)\overline{u'w'}(z_0)exp\Big(\int_{z_0}^{z} -\frac{\alpha k N}{(\omega - \bar{u}k)^2}dz\Big).$$

where $\rho_0(z)$ is the mean density. The mean flow equation is generalized to

$$\frac{\partial \bar{u}}{\partial t} = -\frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0\overline{u'w'}) + \frac{K}{\rho_0}\frac{\partial}{\partial z}(\rho_0\frac{\partial \bar{u}}{\partial z})$$

The reduction of mean density with height leads to more effective wave driving of the mean flow accelerations at higher levels. In the compressible model one can obtain a very pronounced downward progression of the jet maximum even with a wave dissipation rate, α, that is constant with height.

4.4 Two-Wave Plumb Model

The example above showed that the interaction of a single dissipating internal wave with the mean flow can produce monotonic mean flow accelerations that are

self-limiting. We now show that adding a second wave with phase speed in the opposite direction allows the mean flow to oscillate.

Consider a system with two waves such that $k_1 = -k_2$ and $\omega_1 = \omega_2$. The governing equation for the mean flow is now

$$\frac{\partial \bar{u}}{\partial t} = -\frac{\partial}{\partial z}(\overline{u'w'}) + K\frac{\partial^2 \bar{u}}{\partial z^2}$$

where the total Reynolds stress is

$$\overline{u'w'}(z) = \overline{u'w'}_1(z_0)exp\Big(\int_{z_0}^{z} -\frac{\alpha k_1 N}{(\omega_1 - \bar{u}k_1)^2}dz\Big) + \overline{u'w'}_2(z_0)exp\Big(\int_{z_0}^{z} -\frac{\alpha k_2 N}{(\omega_2 - \bar{u}k_2)^2}dz\Big)$$

For simplicity we suppose the waves have equal amplitudes at the lower boundary, so that

$$\overline{u'w'}_1(z_0) = -\overline{u'w'}_2(z_0)$$

Fig. 34 shows results from an integration of a 2-wave Plumb model using waves with phase speeds of +25 m/s and -25 m/s. The result shown is for a compressible version and other details can be found in Hamilton (1982b). The mean flow evolution displays an oscillation between positive and negative winds with a period of slightly over 1000 days, downward propagation of the wind reversals, and the development of intense shear zones. The basic resemblance to the observed QBO is striking and represents a nice demonstration that the downward propagation of the jets and shear zones can result purely from interactions with upward propagating waves, i.e. the basic forcing for the stratospheric QBO may be regarded as residing in the troposphere where the waves are excited.

The basic reason for the oscillatory behavior in Fig. 34 is easy to understand. Once the flow develops with strong winds in, say, the +x direction within the lower part of the domain, the propagation to upper levels of the wave with negative

phase speed is enhanced. When this wave is disspated at these upper levels a negative jet is formed. This is a kind of "shadowing effect" in which the effects of a spectrum of vertically-propagating waves tends to produce a mean flow anomaly aloft of opposite sign to that at lower levels. The negative jet produced aloft can descend just as in the 1-wave case. Once the negative jet has descended low enough the effects of the mean flow viscosity destroy the positive jet below, allowing the negative mean flow region to descend right to the lower boundary. The process can then continue with a positive jet produced aloft, and so on. This process is illustrated schematically in Fig. 35.

The behaviour of the simple 2-wave Plumb model described here has been reproduced in a laboratory experiment by Plumb and McEwan (1978). They put salt-stratified water into an annulus with a flexible membrane at the bottom. The membrane was then forced up and down in order to excite a standing oscillation at the bottom. Since the standing oscillation can be regarded as the sum of two waves with equal but opposite phase velocities, the laboratory experiment is almost a direct analogue of the 2-wave Plumb model (with the annulus geometry supplying the periodic boundary conditions). The results were quite dramatic, as the waves produced a much longer period oscillation in the annulus-averaged circulation with downward-propagating flow reversals.

4.5 Further Development of Simple QBO Models

Perhaps the most fundamental problem impeding our understanding of the dynamics of the QBO is the inability to diagnose accurately from observations the actual contributions of different parts of the wave spectrum to the mean flow accelerations in the tropics. A direct calculation of the vertical Reynolds stress requires knowledge of the vertical wind, which cannot be directly measured on a

regular basis. Even the planetary-scale variations of the horizontal wind in the tropical stratosphere are hard to determine, given the sparse radiosonde network and the difficulty in obtaining reliable estimates of the wind from satellite temperature observations. This lack of a firm observational quantitative understanding affected the development of QBO theory from the beginning. In fact, the original model of Lindzen and Holton (1968) supposed that the forcing of the mean flow was due to gravity waves with a continuous spectrum of horizontal phase speeds, while the same authors four years later (Holton and Lindzen, 1972) showed that similar results could be obtained by including only one eastward-propagating and one westward-propagating large-scale equatorial planetary wave. Even today we cannot be certain of the relative contributions of planetary-scale waves and smaller-scale gravity waves.

The 1972 Holton and Lindzen paper presents a theory very similar to that of Plumb (1977), except that the expressions for vertical group velocity of the equatorial Kelvin and Rossby-gravity waves are used for the eastward and the westward-propagating waves, respectively (see Lindzen, 1971). The choice of wave parameters in the model was roughly based on available observations of planetary-scale equatorial waves (see Wallace, 1973, for a review). The results were quite encouraging, since with "reasonable" parameters the model was found to simulate an equatorial QBO with roughly the correct period, amplitude and vertical structure.

Much of the theoretical work on the QBO over the last 25 years has involved generalizations of the Holton-Lindzen (1972) model. Dunkerton (1981) considered the effects of wave transience in the model (i.e. he relaxed the assumption of steadiness in the calculation of the wave fluxes). Saravanan (1990) generalized the Holton-Lindzen model to include a large number of waves. Dunkerton (1983) added

an additional easterly forcing to the Holton-Lindzen model which was meant to account for the effects of mean flow forcing from quasi-stationary planetary waves forced by topography in the extratropics.

A major limitation of the Holton-Lindzen and Plumb QBO models is their restriction to one spatial dimension, i.e. they solve for the height and time dependence of a mean flow meant to represent the flow averaged over some latitude band around the equator. Plumb and Bell (1982) constructed a numerical model of the QBO forced by two waves, but including a treatment of the waves and mean flow in both height and latitude. Also included in the Plumb and Bell work was the effect of the vertical advection by the mean meridional circulation produced by mean flow radiative effects. They noted that the presence of westerly shear on the equator is associated with a warm anomaly at the equator and hence anomalous diabatic cooling, leading to mean sinking (or at least anomalously weak rising motion). This aspect of the QBO had been qualitatively discussed earlier by Reed (1965), and Reed's schematic diagram of this effect is reproduced here as Fig. 36. Plumb and Bell noted that the vertical advection associated with this component of the mean meridional circulation should act to intensify the accelerations as westerly jets descend, since the downward advection of the westerly momentum from above adds to the local westerly wave-driving. Conversely, the mean advection effect should act to weaken the easterly accelerations. Plumb and Bell proposed this as the mechanism to explain the observed asymmetry between the strengths of the easterly and westerly shear zones in the QBO. The effects of the QBO in modulating the mean meridional circulation will be the focus of a series of lectures later in this School.

Recently the possibility that a broad spectrum of gravity waves might be responsible for driving the QBO has been revived (e.g., Dunkerton, 1997). In one

interesting development, Alexander and Holton (1997) conducted high-resolution limited-area explicit numerical simulations of a tropical squall line and found that the momentum fluxes associated with the gravity waves forced by such storms could well be significant for the dynamics of the QBO.

References

Alexander, M.J. and J.R. Holton, 1997: A model study of zonal forcing in the equatorial stratosphere by convectively induced gravity waves. *J. Atmos. Sci., 54,* 408-419.

Andrews, D.G., J.R. Holton and C.B. Leovy, 1987: *Middle Atmosphere Dynamics*, Academic Press.

Baldwin, M., L. Gray, T. Dunkerton, K. Hamilton, P. Haynes, W. Randel, J. Holton, M. Alexander, I Hirota, T. Horinouchi, D. Jones, J. Kinnersley, C. Marquardt, K. Sato and M. Takahashi, 2001: The quasi-biennial oscillation. *Rev. Geophys., 39,* 179-229.

Belmont, A.D., and D.G. Dartt: 1968: Variation with longitude of the quasi-biennial oscillation. *Mon. Weath. Rev., 96,* 767-777.

Belmont, A.D., D.G. Dartt and G.D. Nastrom, 1974: Periodic variations in stratospheric zonal wind from 20 to 65 km at 80°N to 70°S. *Quart. J. Roy. Meteorol. Soc., 100,* 203-211.

Burrage, M.D. and coauthors, 1996: Long term variability in the equatorial mesosphere and lower thermosphere zonal winds. *J. Geophys. Res., 101,* 12847-12854.

Delisi, D. and T.J. Dunkerton, 1988: Equatorial semiannual oscillation in zonally averaged temperature observed by Nimbus-7 SAMS and LIMS. *J. Geophys. Res., 93,* 3899-3904.

Dunkerton, T.J., 1981: Wave transience in a compressible atmosphere, Pt. 2, Transient equatorial waves in the quasi-biennial oscillation. *J. Atmos. Sci., 38,* 298-307.

Dunkerton, T.J., 1983: Laterally-propagating Rossby waves in the easterly acceleration phase of the quasi-biennial oscillation. *Atmosphere-Ocean, 21,* 55-68.

Dunkerton, T.J., 1991: Nonlinear propagation of zonal winds in an atmosphere with Newtonian cooling. *J. Atmos. Sci., 48,* 236-263.

Dunkerton, T.J., 1997: The role of gravity waves in the quasi-biennial oscillations. *J. Geophys. Res., 102,* 26053-26076.

Dunkerton, T.J. and D. Delisi, 1985: Climatology of the equatorial lower stratosphere. *J. Atmos. Sci., 42,* 1199-1208.

Dunkerton, T.J. and D. Delisi, 1997: Interaction of the quasibiennial oscillation and the stratopause semiannual oscillation. *J. Geophys. Res., 102,* 26107-26116.

Eliassen, A. and E. Palm, 1961: On the transfer of energy in stationary mountain waves. *Geofys. Publ., 22,* 1-23.

Fels, S.B., 1982: A parameterization of scale-dependent radiative damping rates in the middle atmosphere. *J. Atmos. Sci., 39,* 1141-1152.

Forbes, J., 1981: The equatorial electrojet. *Rev. Geophys., 19,* 469-504.

Garcia, R.R., T.J. Dunkerton, R.S. Liebermann and R.A. Vincent, 1997: Climatology of the semiannual oscillation of the tropical middle atmosphere. *J. Geophys. Res., 102,* 26019-26032.

Hamilton, K., 1982a: Rocketsonde observations of the mesospheric semiannual oscillation at Kwajalein *Atmosphere-Ocean, 20,* 281-286.

Hamilton, K., 1982b: A note on the interaction between a Thermally Forced Standing Internal Gravity Wave and the Mean Flow: Implications for the Theory of the Quasi-biennial Oscillation. *J. Atmos. Sci., 39,* 1881-1886.

Hamilton, K., 1984: Mean wind evolution through the quasi-biennial cycle in the tropical lower stratosphere. *J. Atmos. Sci., 41,* 2113-2125.

Hamilton, K., 1985: The initial westerly acceleration phase of the stratospheric quasi-biennial oscillation as revealed in FGGE analyses. *Atmosphere-Ocean, 23,* 188-192.

Hamilton, K., 1987: A review of observations of the quasi-biennial and semi-annual oscillations of the tropical middle atmosphere. *Transport Processes in the Middle Atmosphere,* D. Riedel Publishing Co., pp. 19-29.

Hamilton, K., 1998: Observations of tropical stratospheric winds before World War II. *Bull. Amer. Meteorol. Soc., 79,* 1367-1371.

Haynes, P.H., C.J. Marks, M.E. McIntyre, T.G. Shepherd and K. Shine, 1991: On the downward control of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci., 48,* 651-678.

Hirota, I., 1978: Equatorial Kelvin waves in the upper stratosphere in relation to the semiannual oscillation of the zonal wind. *J. Atmos. Sci., 35,* 714-722.

Hitchman, M.H. and C.B. Leovy, 1986: Evolution of the zonal mean state in the equatorial middle atmosphere during October 1978-May 1979. *J. Atmos. Sci., 43,* 3159-3176.

Holton, J.R. and R.S. Lindzen, 1972: An updated theory for the quasi-biennial cycle of the tropical stratosphere. *J. Atmos. Sci., 29,* 1076-1080.

Lindzen, R.S., 1971: Equatorial planetary waves in shear. Part I. *J. Atmos. Sci., 28,* 1452-1463.

Lindzen, R.S. and J.R. Holton, 1968: A theory of the quasi-biennial oscillation. *J. Atmos. Sci., 25,* 1095-1107.

Naujokat, B., 1986: An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *J. Atmos. Sci., 43,* 1873-1877.

Ortland, D.A. and coauthors, 1996: Measurements of stratospheric winds by the high resolution Doppler imager. *J. Geophys. Res., 101,* 10351-10363.

Palmer, C.E., 1954: The general circulation between 200 mb and 10 mb over the equatorial Pacific. *Weather, 9,* 3541-3549.

Plumb, R.A., 1977: The interaction of two internal waves with the mean flow: implications for the theory of the quasi-biennial oscillation. *J. Atmos. Sci., 34,* 1847-1858.

Plumb, R.A., 1984: The quasi-biennial oscillation. *Dynamics of the Middle Atmosphere* (J.R. Holton and T. Matsuno, eds.), Terra Scientific Publishing, 217-251.

Plumb, R.A. and R.C. Bell, 1982: Model of the quasi-biennial oscillation on an equatorial beta-plane. *Quart. J. Roy. Meteorol. Soc., 108,* 335-352.

Plumb, R.A. and A.D. McEwan, 1978: The instability of a forced standing wave in a viscous stratified fluid: a laboratory analogue of the quasi-biennial oscillation. *J. Atmos. Sci., 35,* 1827- 1839.

Reed, R.J., 1965: The present status of the 26-month oscillation, *Bull. Amer. Meteorol. Soc., 46,* 374-386.

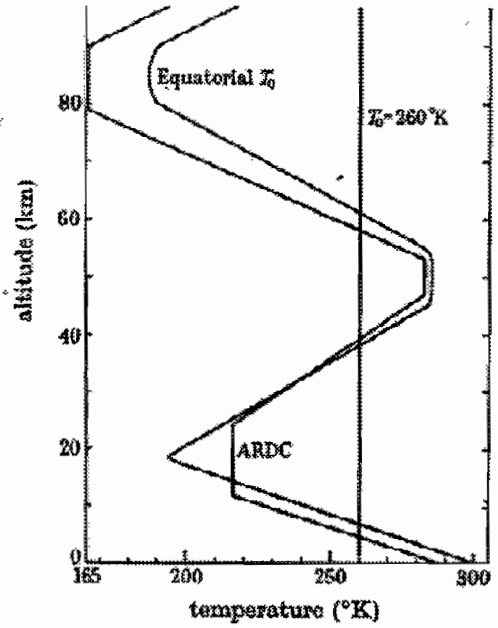
Reed, R.J., W.J. Campbell, L.A. Rasmussen and D.G. Rogers, 1961: Evidence of a downward-propagating annual wind reversal in the equatorial stratosphere. *J.Geophys. Res, 66,* 813-818.

Salby, M., and P. Callaghan, 2000: Connection between the Solar Cycle and the QBO: The missing link. *J. Clim., 13,* 2652-2662.

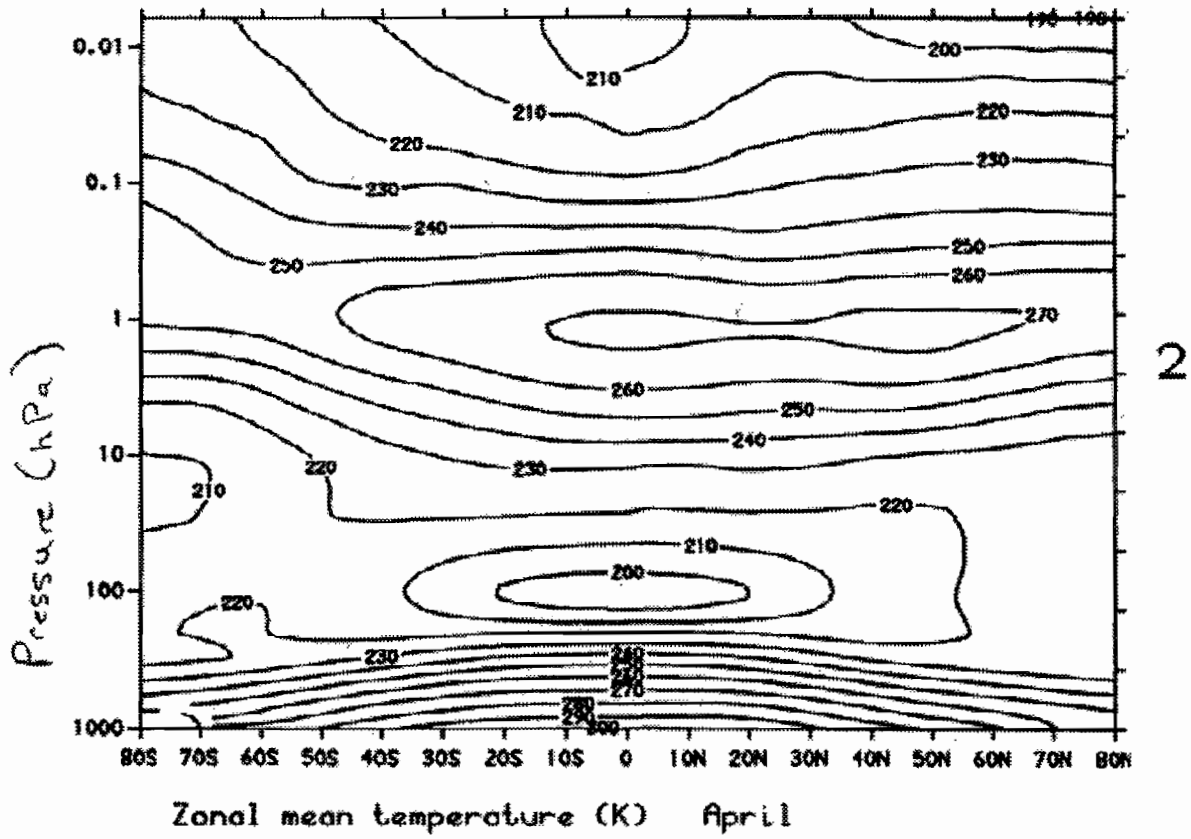
Saravanan, R., 1990: A multiwave model of the quasi-biennial oscillation. *J. Atmos. Sci., 47,* 2465-2474.

Veryard, R.G. and R.A. Ebdon, 1961: Fluctuations in tropical stratospheric winds. *Meteorological Magazine, 90,* 127-143.

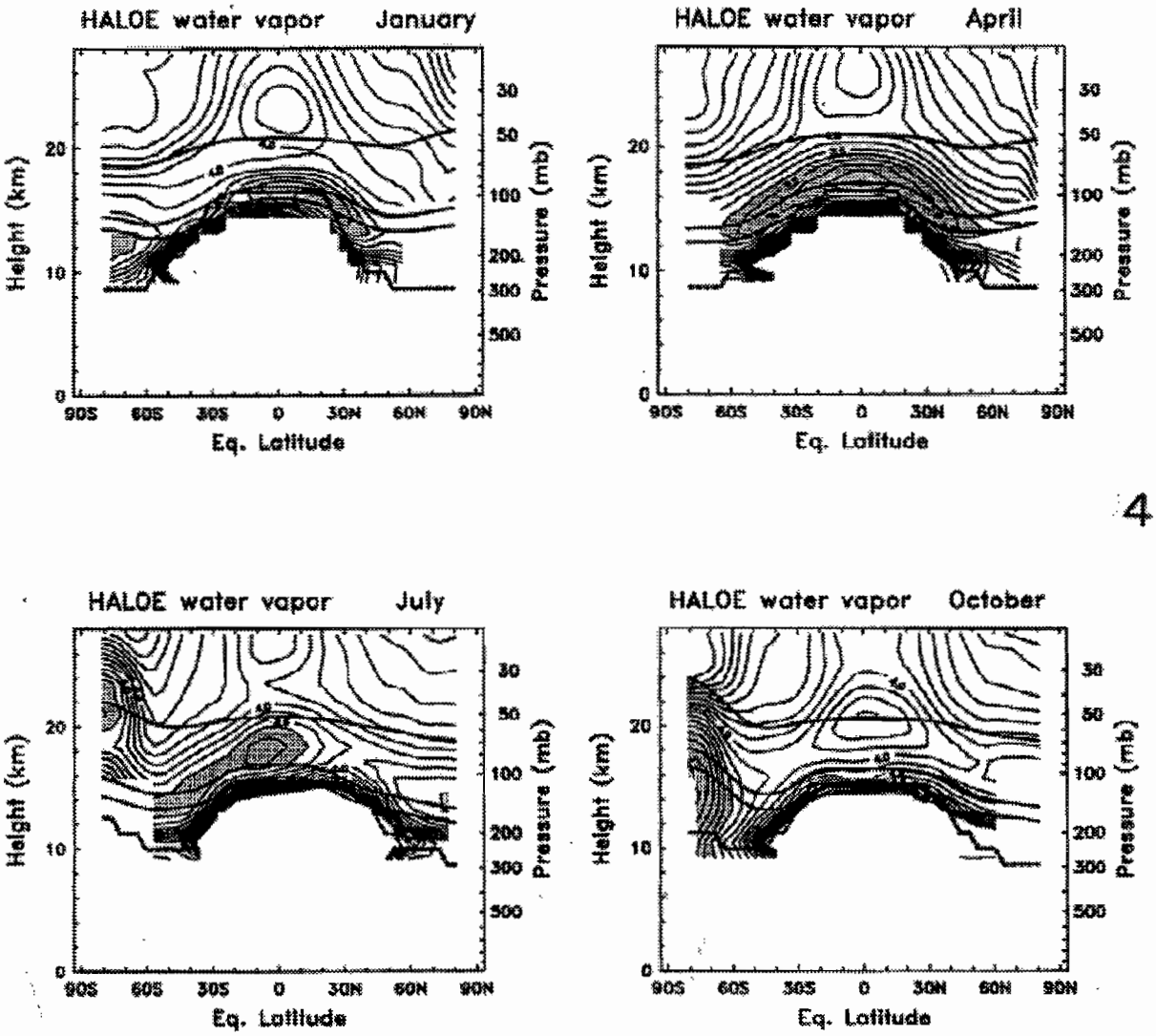
Wallace, J.M., 1973: The general circulation of the tropical lower stratosphere. *Rev. Geophys. Space Phys., 11,* 191-222.


Equatorial T_0
T_0=260°K
ARDC
altitude (km)
temperature (°K)
0
20
40
60
80
165
200
250
300

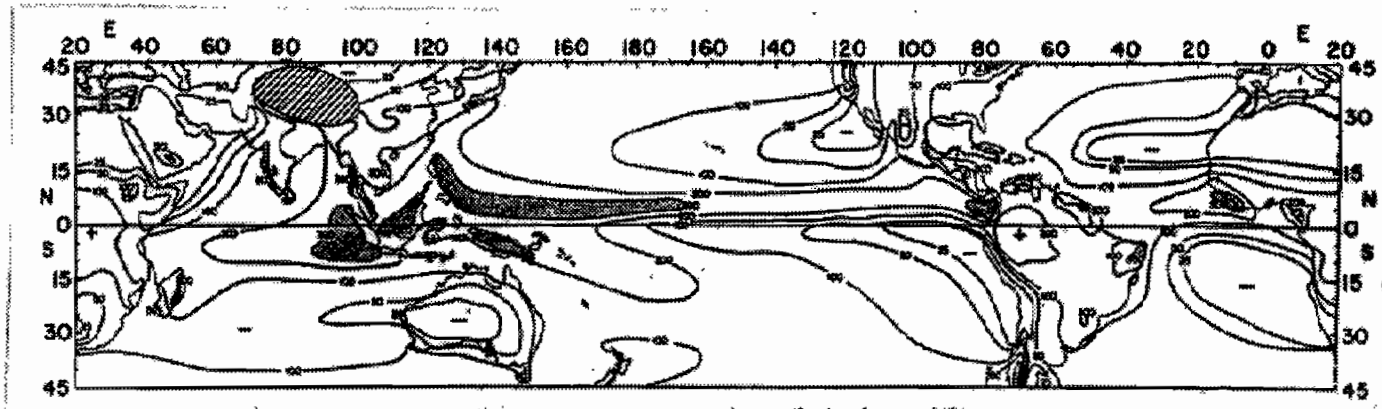
1


Pressure (hPa)
0.01
0.1
1
10
100
1000
80S 70S 60S 50S 40S 30S 20S 10S 0 10N 20N 30N 40N 50N 60N 70N 80N
Zonal mean temperature (K) April

2

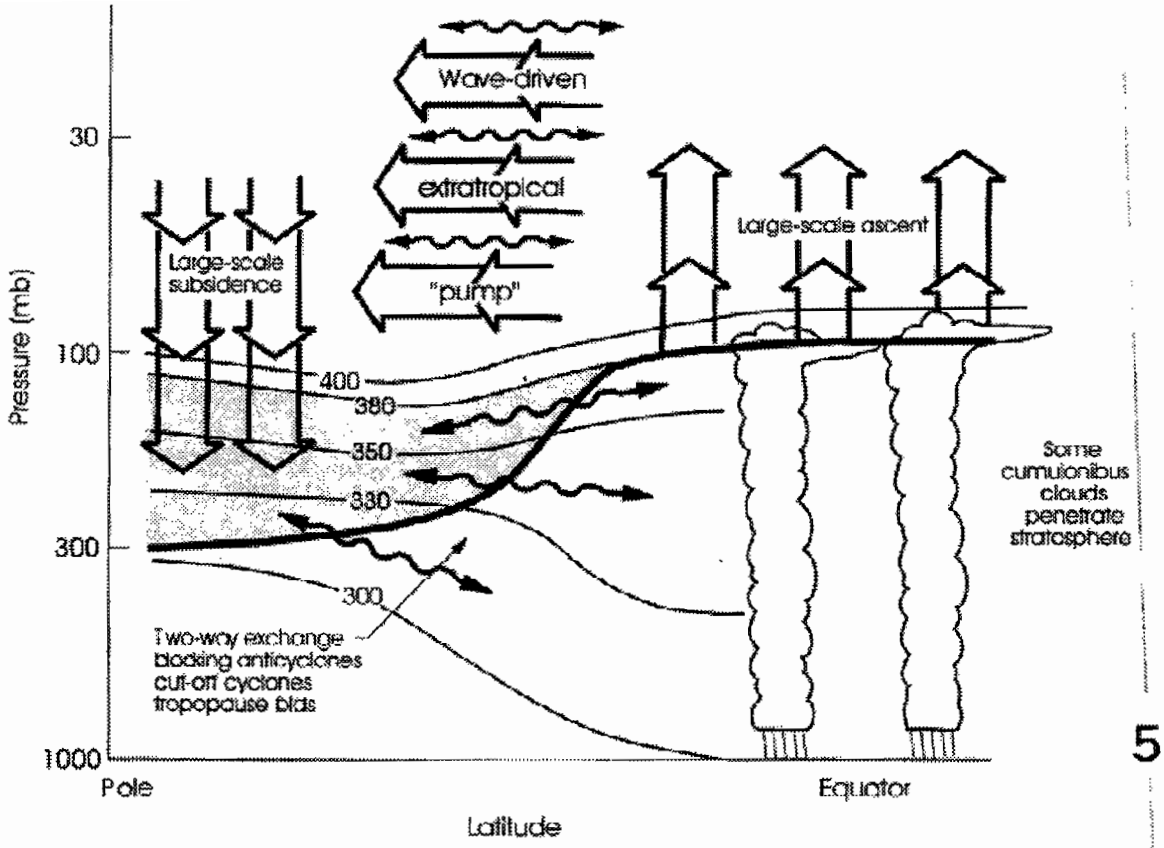


Meridional cross-sections of water vapour mixing ratio between the tropopause and ~28 km derived from HALOE measurements. Plots are shown for January, April, July and October, illustrating the seasonal cycle. For each month the dashed lines represent the tropopause, and the solid lines show the 380 K, 400 K, and 500 K isentropes.



(3)

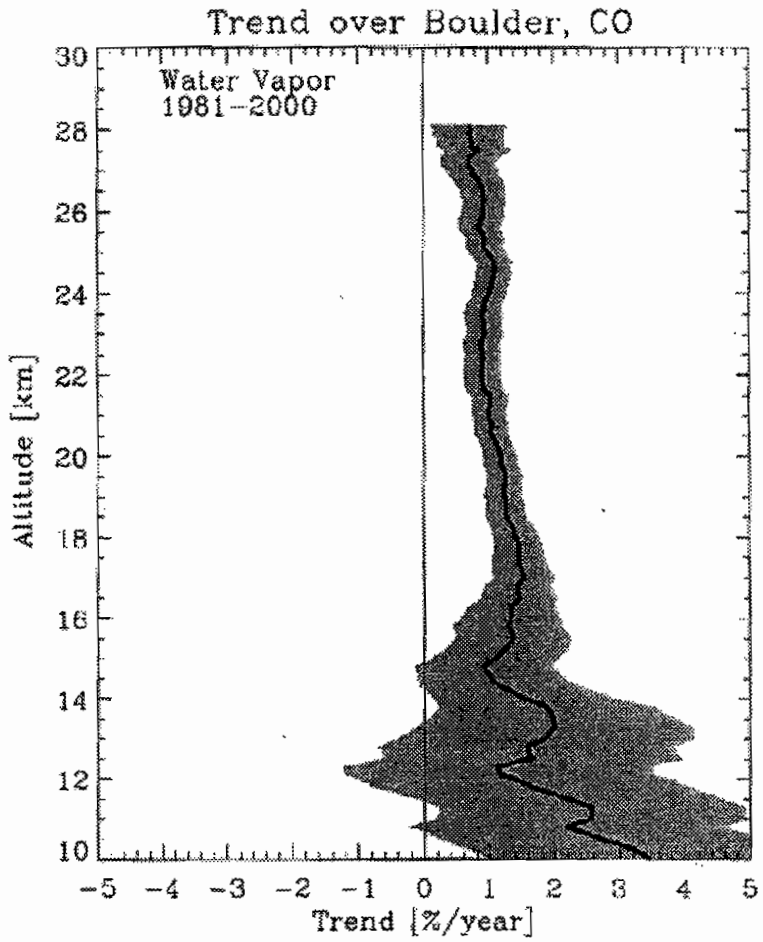
: mean annual precipitation (cm) (...). Shaded: regions with more than 300 cm/year.



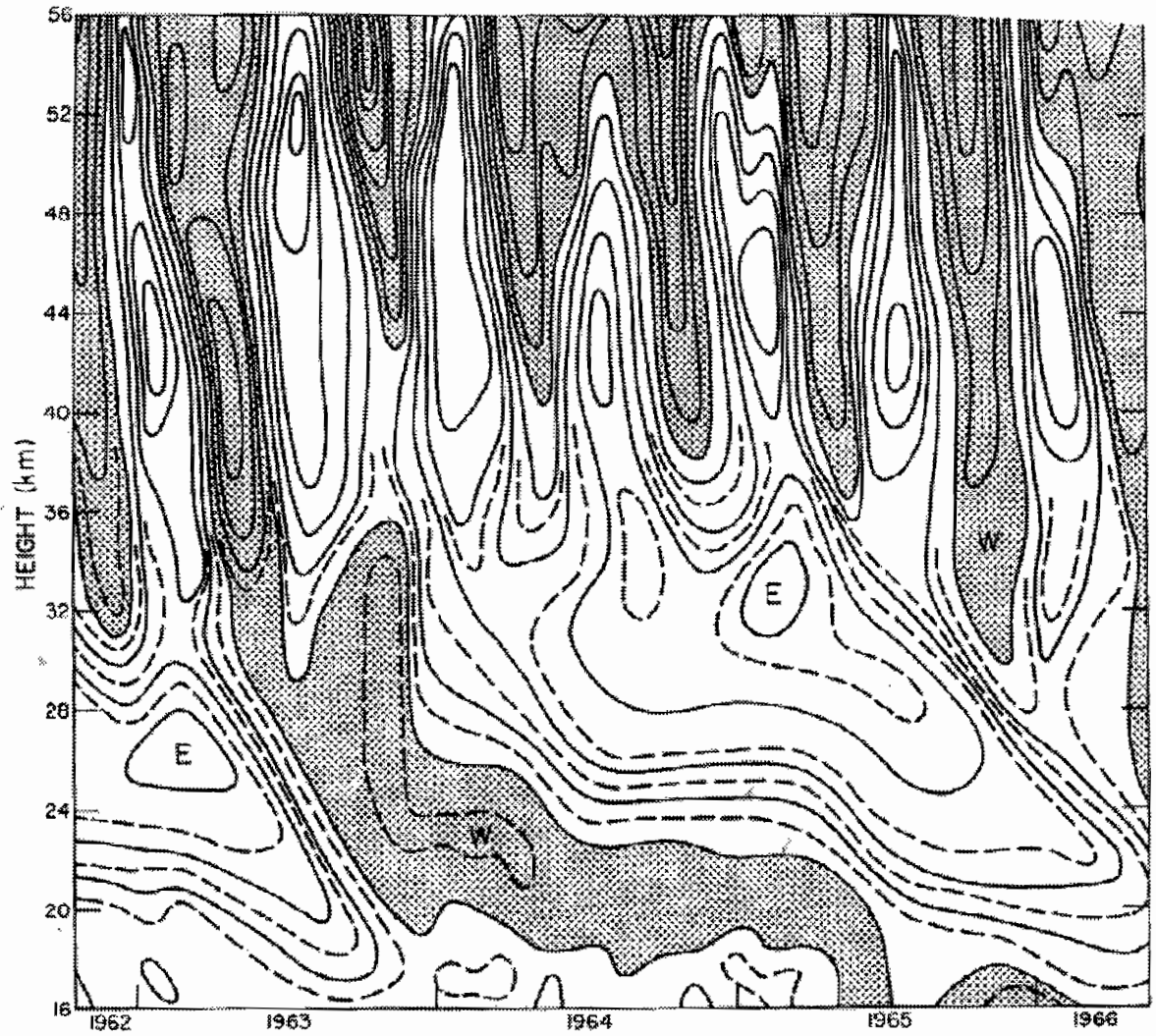
5

Dynamical Aspects of stratosphere-troposphere exchange.

The tropopause is shown by the thick line. Thin lines are isentropic or constant potential temperature surfaces labelled in Kelvins. Heavily shaded region is the "lowermost stratosphere" where isentropic surfaces span the tropopause and isentropic exchange by tropopause folding occurs. The region above the 380K surface is the "overworld", in which isentropes lie entirely in the stratosphere. Light shading in the overworld denotes wave-induced forcing (the extratropical "pump"). The wiggly double headed arrows denote meridional transport by eddy motions, which include tropical upper-tropospheric troughs and their cut-off cyclones as well as their midlatitude counterparts including folds. Not all eddy transports are shown; and the wiggly arrows are not meant to imply any two way symmetry. The broad arrows show transport by the global-scale circulation which is driven by the extratropical pump. This global scale circulation is the primary contribution to exchange across isentropic surfaces (e.g., the ~380 K surface) that are entirely in the overworld. (Holton et al., Reviews of Geophysics, 1995).

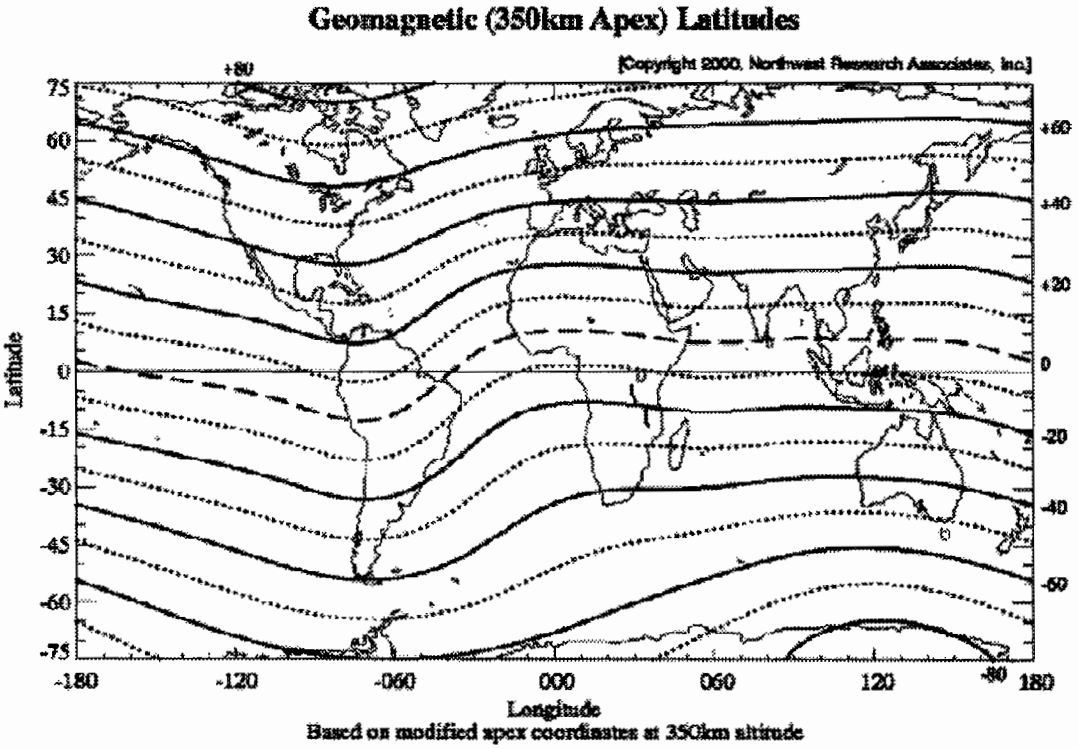


. The linear trend of water vapour mixing ratio (percent per year) and 95% confidence interval (shaded area) as a function of altitude over Boulder. Significant increases of about 1% per year are found at all altitudes above 16 km.



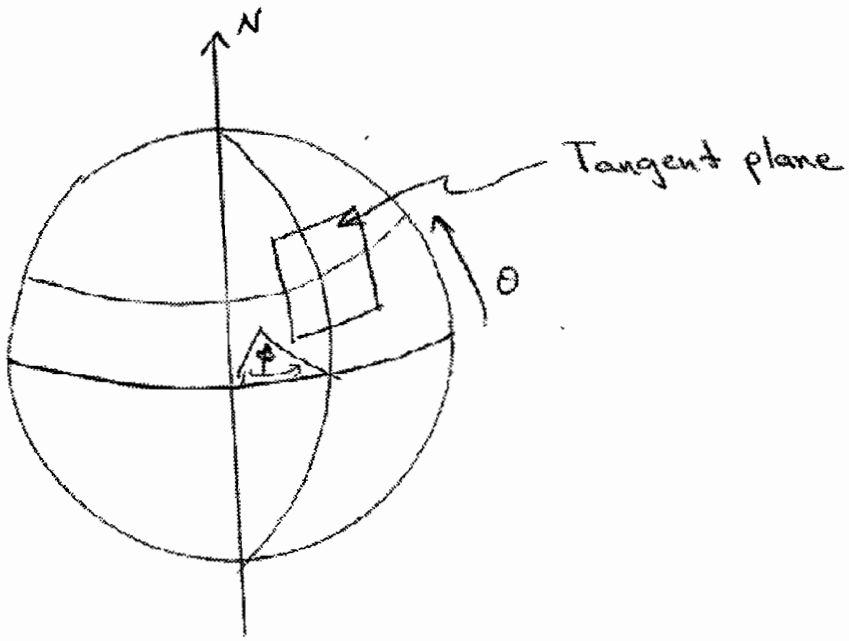
Time–height section of zonal wind in the vicinity of the equator based on monthly averaged data. Solid lines have been drawn at intervals of 10 m s^{-1}. Shaded regions denote flow from west to east, and the letters W denote maxima of such *westerly* flow. E's denote maxima of easterly (east to west) flow. At the higher levels there is a prominent semiannual oscillation while below 35 km there is a longer-term variability associated with the quasi-biennial oscillation. [From Wallace, *Rev. Geophys. and Space Phys.*, **11**, 196 (1973), copyrighted by American Geophysical Union.]

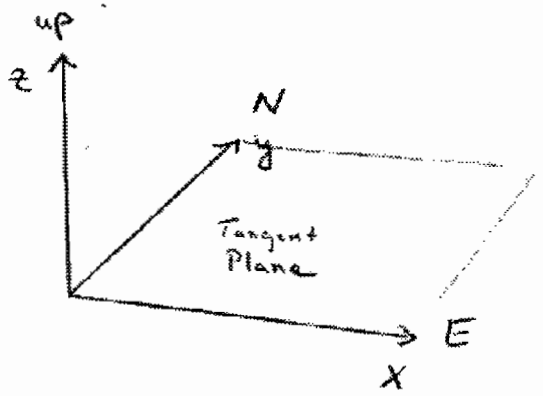
Geomagnetic Latitude



This maps shows lines of constant geomagnetic latitude based on apex geomagnetic coordinates referenced to an altitude of 350km. The geomagnetic equator is indicated by a dashed line. The coordinates were calculated using the IGRF-1985 geomagnetic field model updated to epoch 1990.0.

Definitions

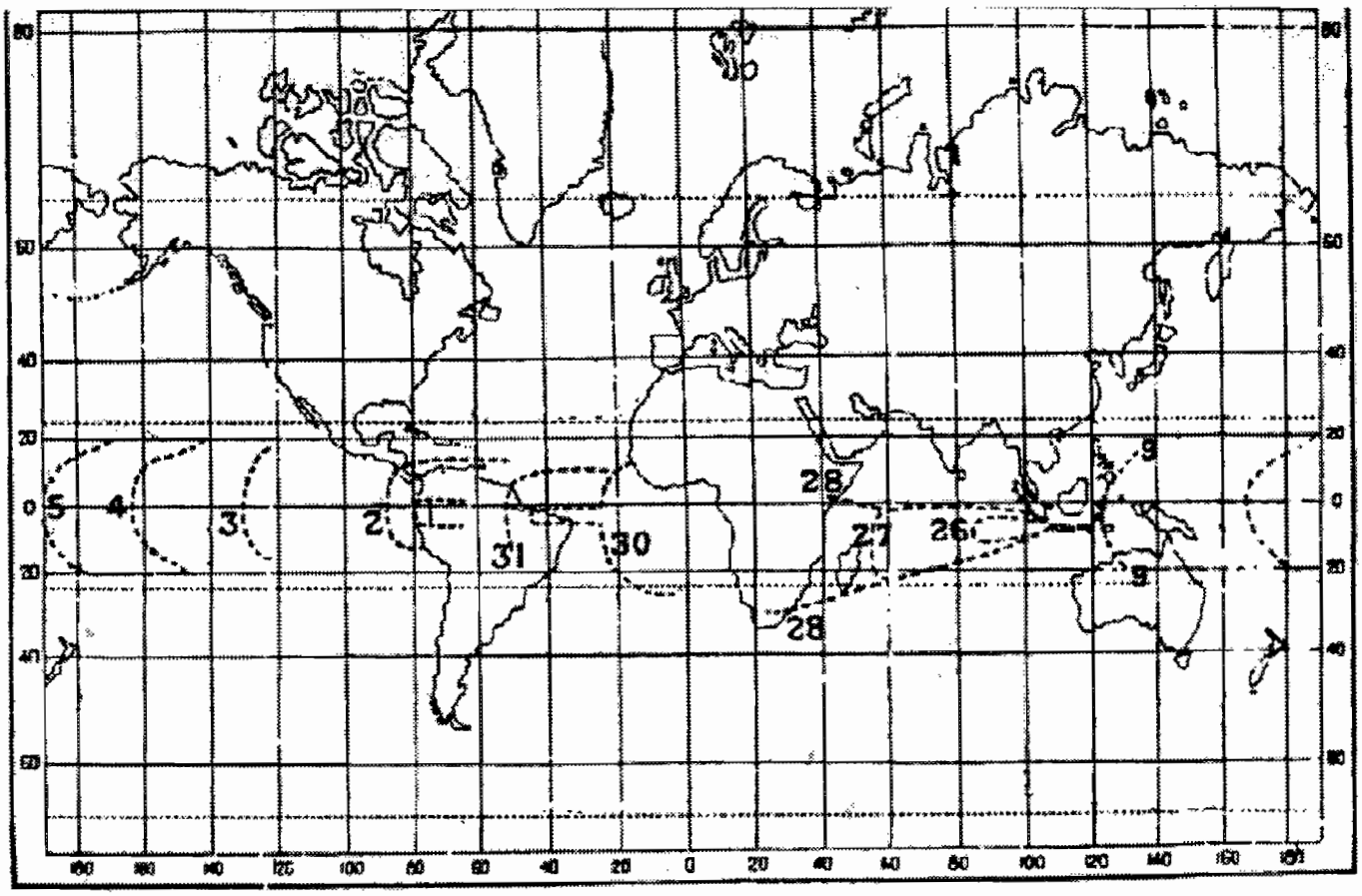




$u \equiv \frac{dx}{dt} \equiv$ "westerly wind"

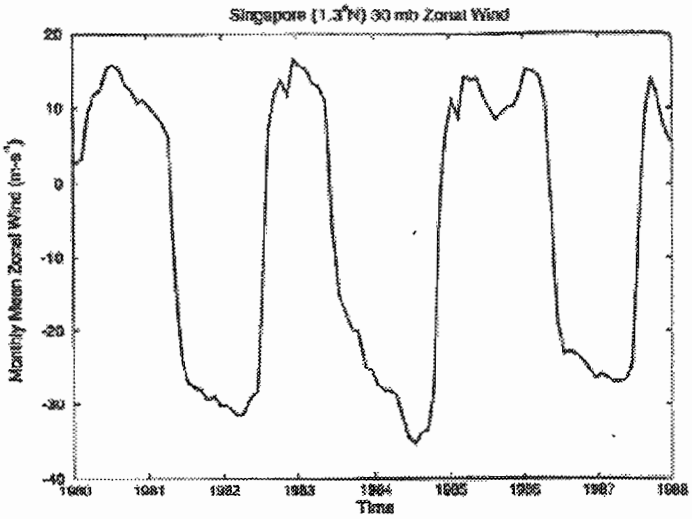
$v \equiv \frac{dy}{dt} \equiv$ "southerly wind"

$w \equiv \frac{dz}{dt} \equiv$ "vertical wind"



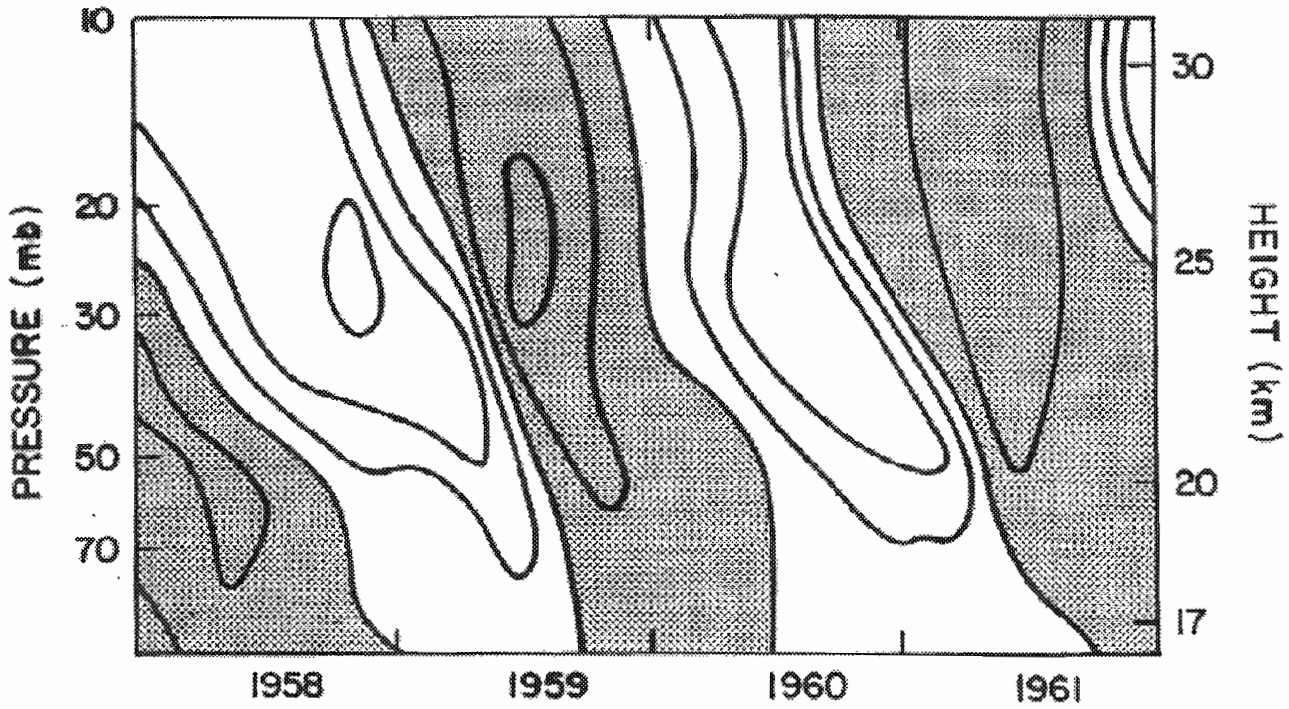
The spread of the optical phenomena observed after the eruption of Mt. Krakatoa on 26 August 1883. The dotted lines give the western boundary of the region where the phenomena had been observed on successive days, 26 August, 27 August ... 9 September. Reproduced from Russell (1888).

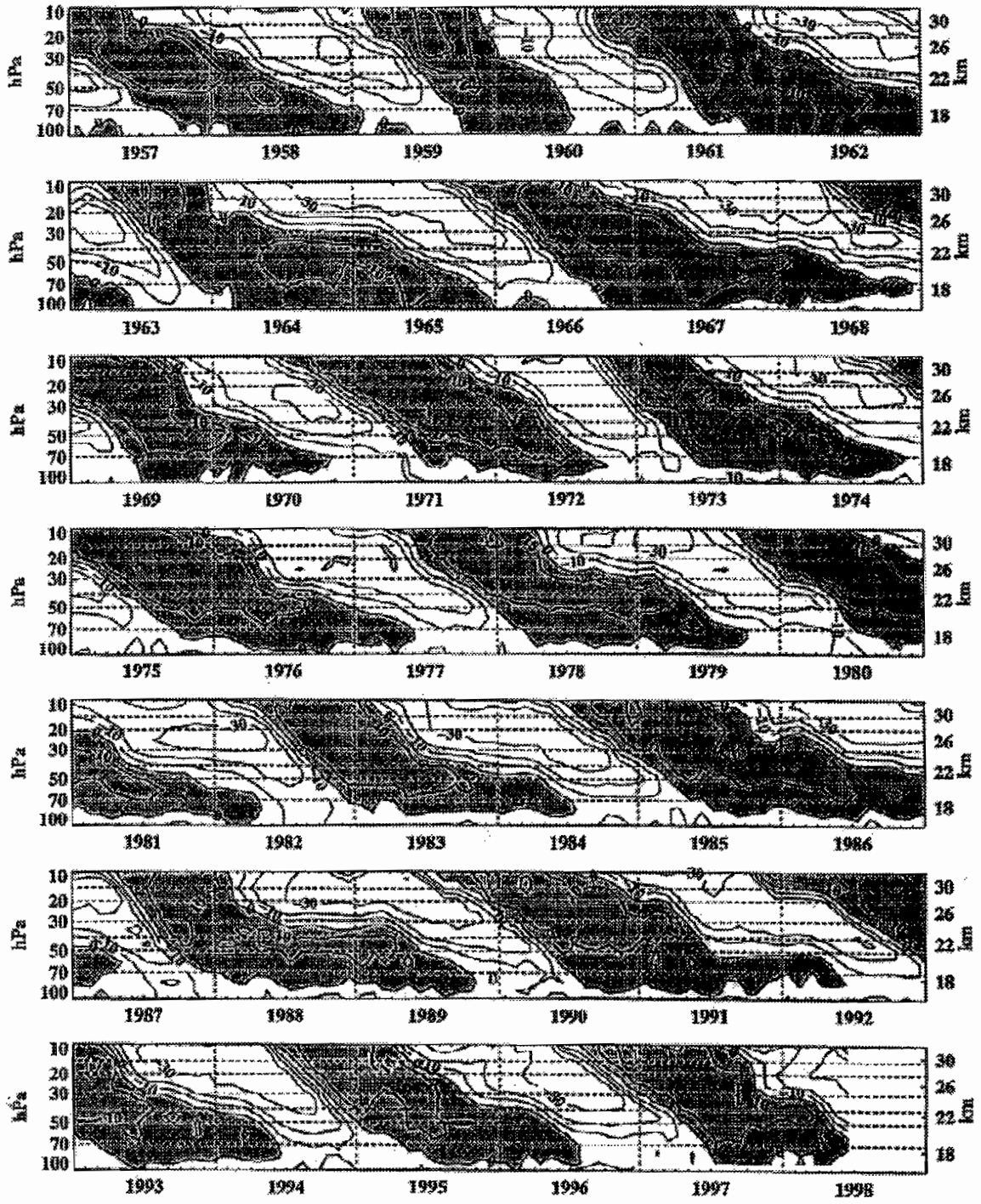
11



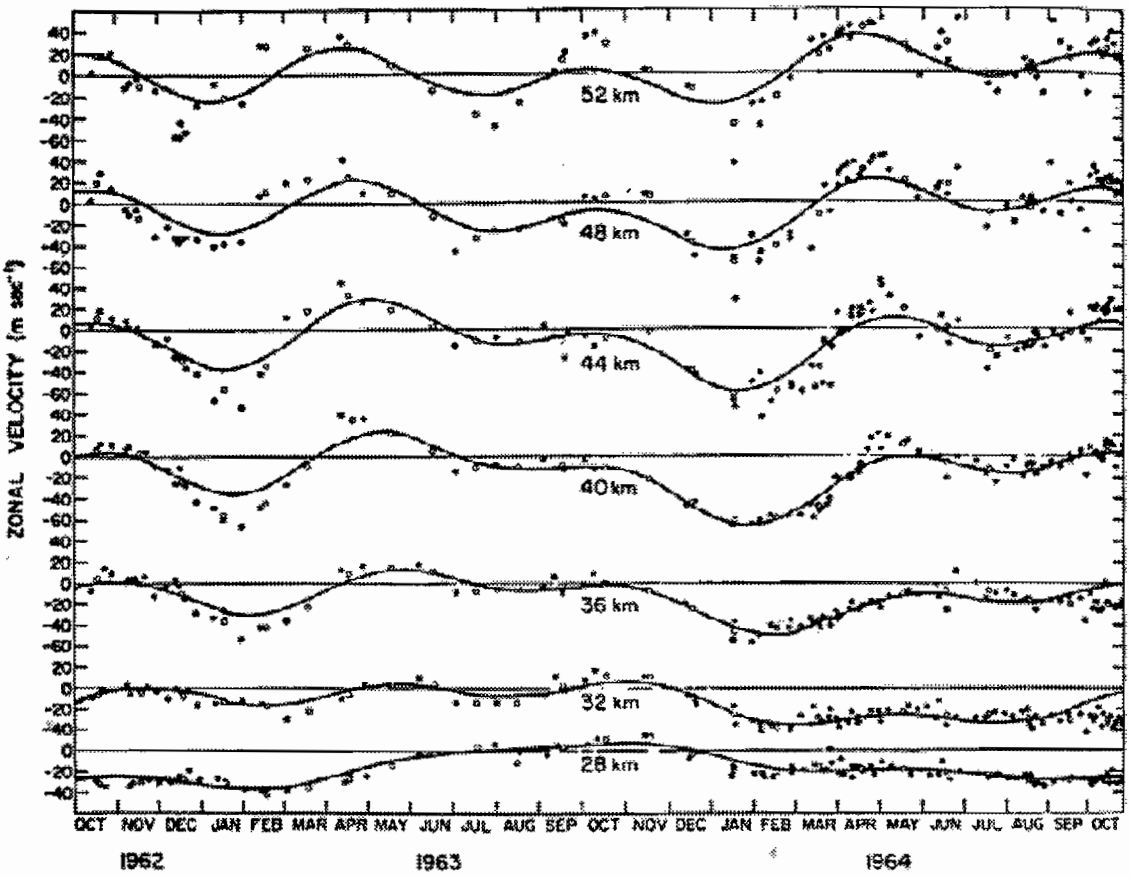
Time series of the monthly-mean zonal wind measured by balloons at Singapore during the period 1980–88.

12

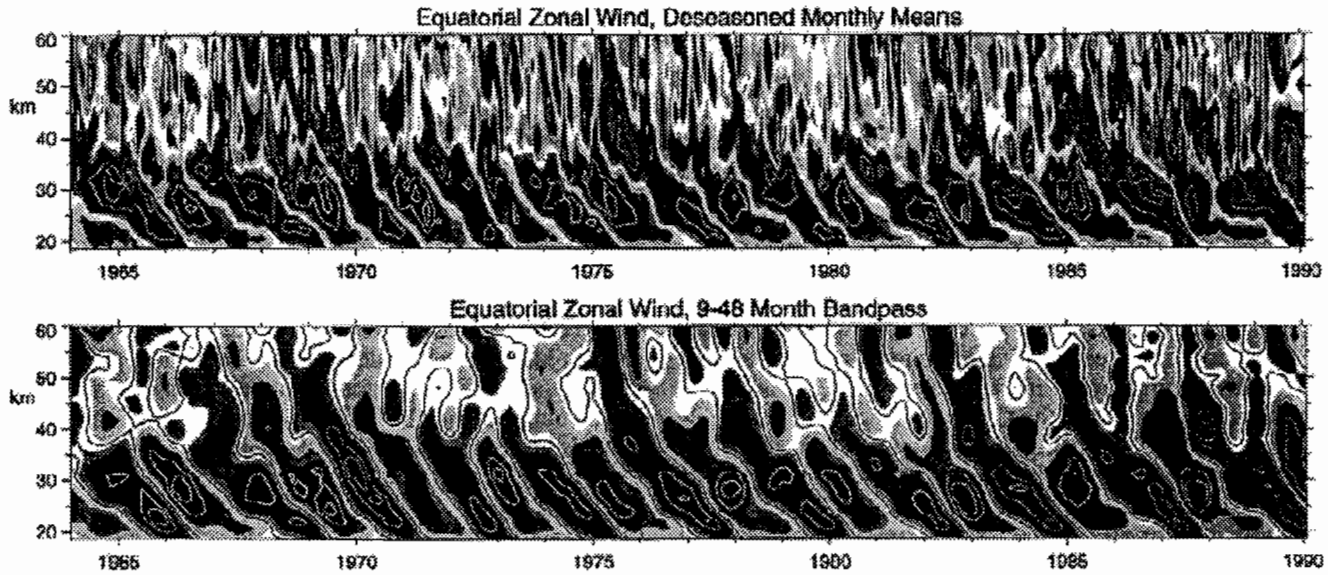




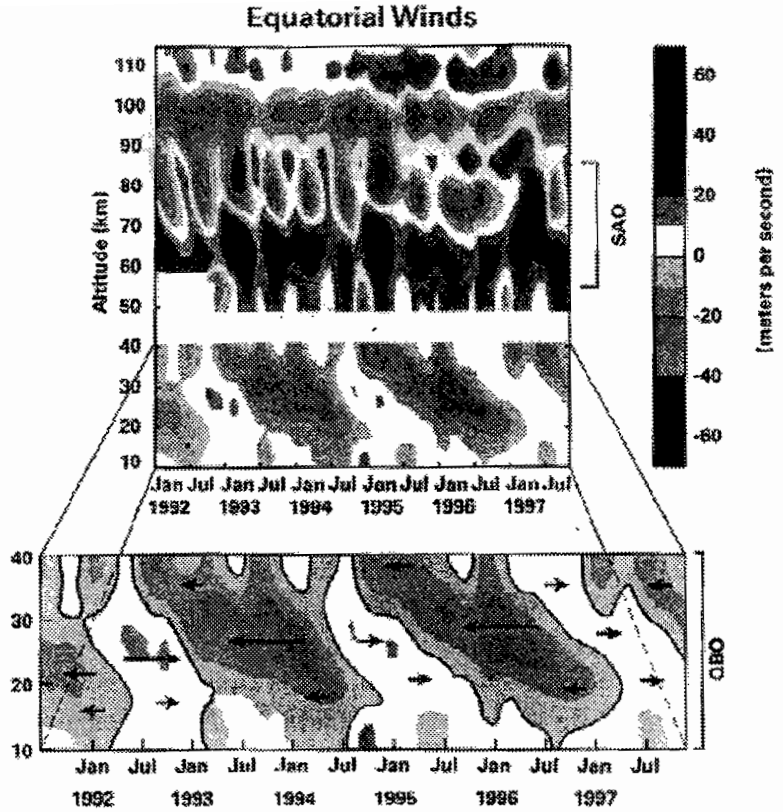
Time-height section of the monthly-mean wind at stations near the equator. Results represent observations from Canton Island (2.8°S, 171.7°W) during 1957–1967, Gan (0.7°S, 73.1°E) during 1967–1975, and Singapore (1.4°N, 103.9°E) during 1976–1998. Westerly winds are shaded and the contour interval is 10 m s^{-1}. Figure provided by B. Naujokat.

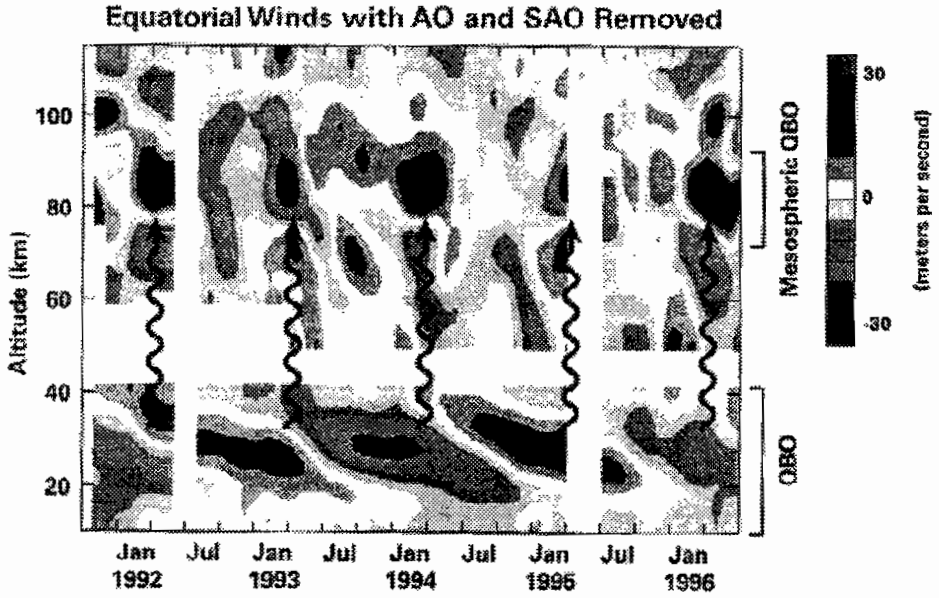


Zonal wind measurements taken at Ascension island (7.9°S) during the period October 1962 through October 1964. The solid circles show individual measurements and the open circles are monthly means for months when there was more than one measurement available. Reproduced

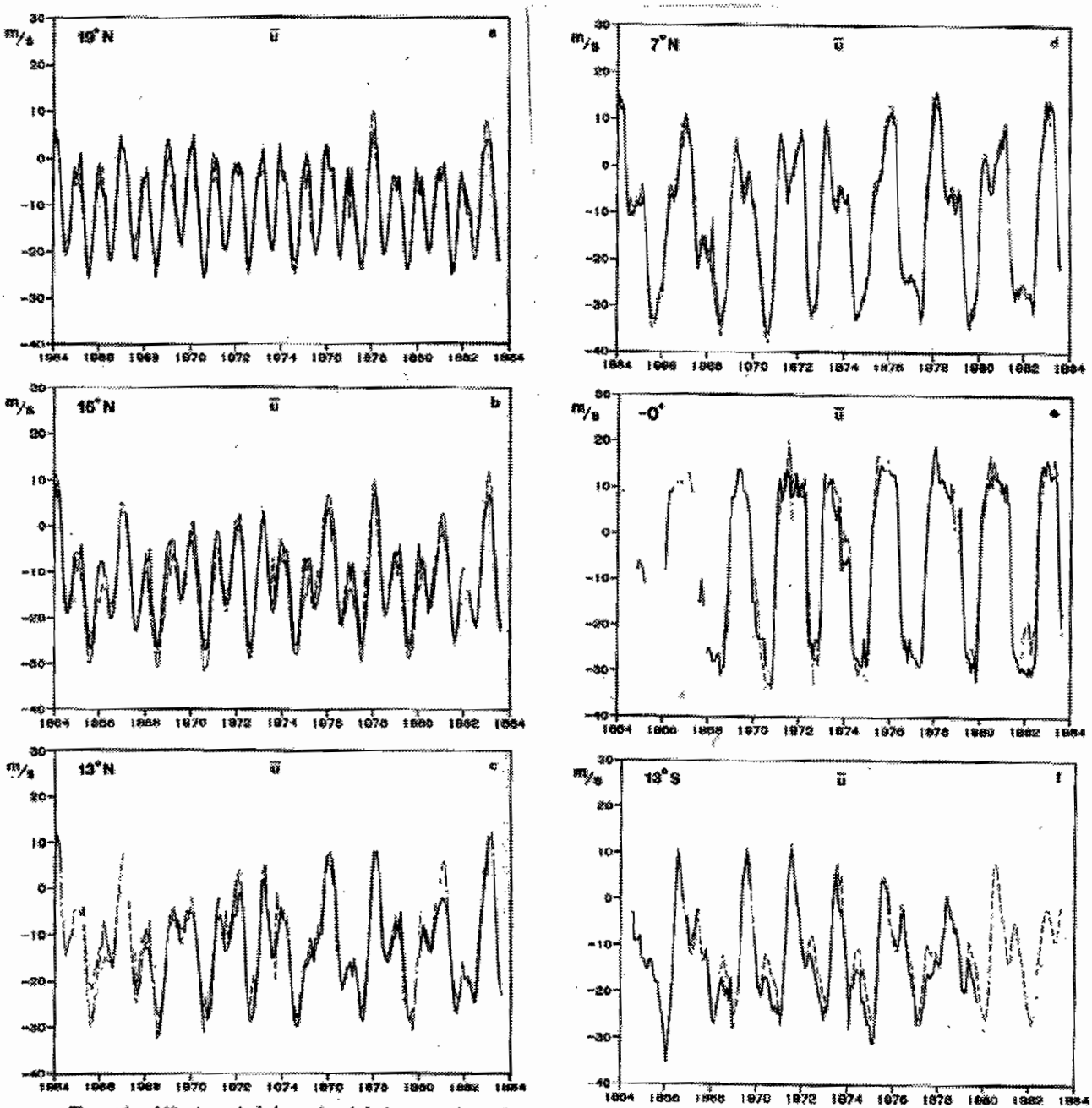


(top) Time-height section of the monthly-mean zonal wind component (m s^{-1}), with the seasonal cycle removed, for 1964–1990. Below 31 km, equatorial radiosonde data are used from Canton Island (2.8°N, January 1964 to August 1967), Gan/Maledive Islands (0.7°S, September 1967 to December 1975), and Singapore (1.4°N, January 1976 to February 1990). Above 31 km, rocketsonde data from Kwajalein (8.7°N) and Ascension Island (8.0°S) are shown. The contour interval is 6 m s^{-1}, with the band between −3 and +3 unshaded. Red represents positive (westerly) winds. After *Gray et al.* [2001]. In the bottom panel the data are band-pass filtered to retain periods between 9 and 48 months.

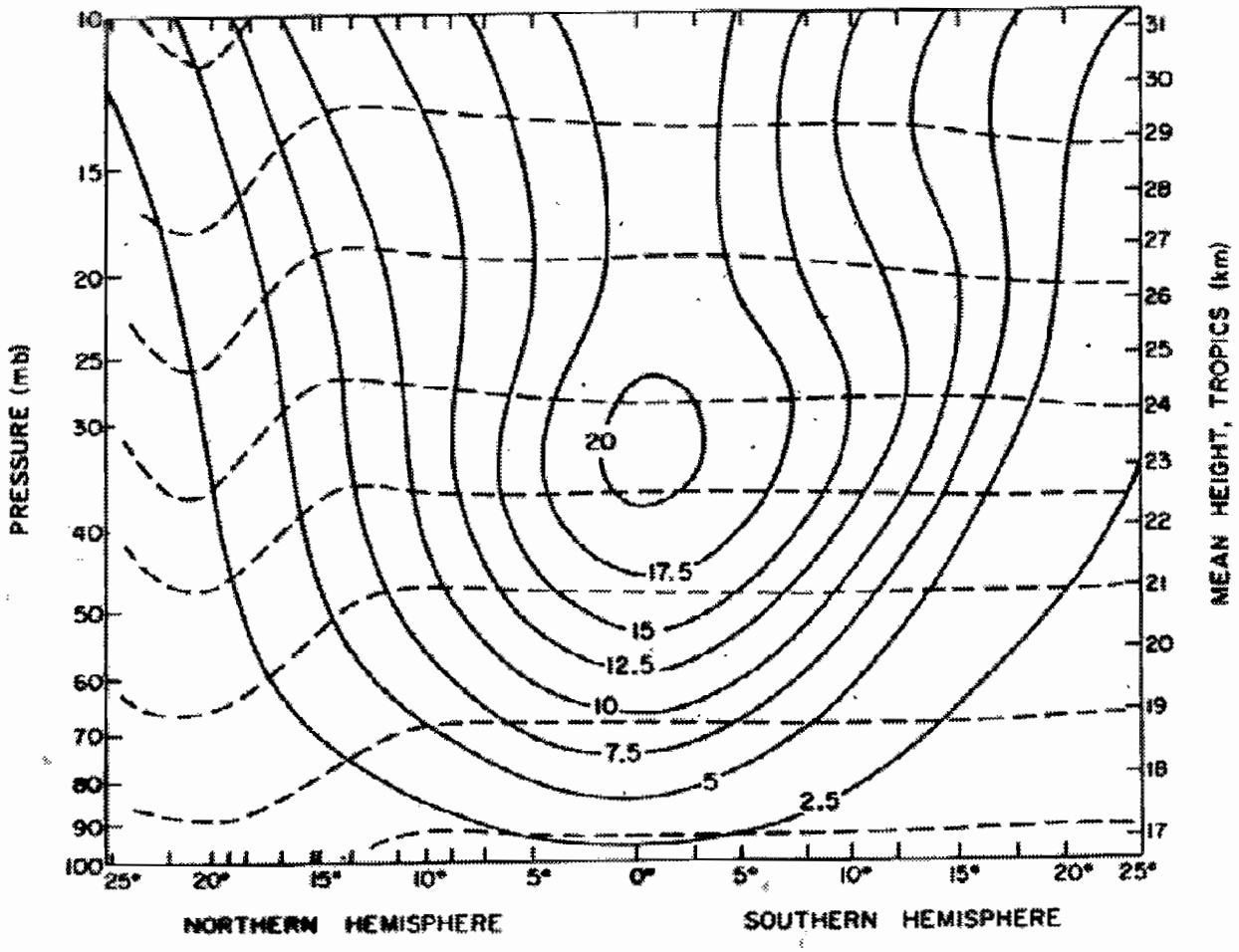


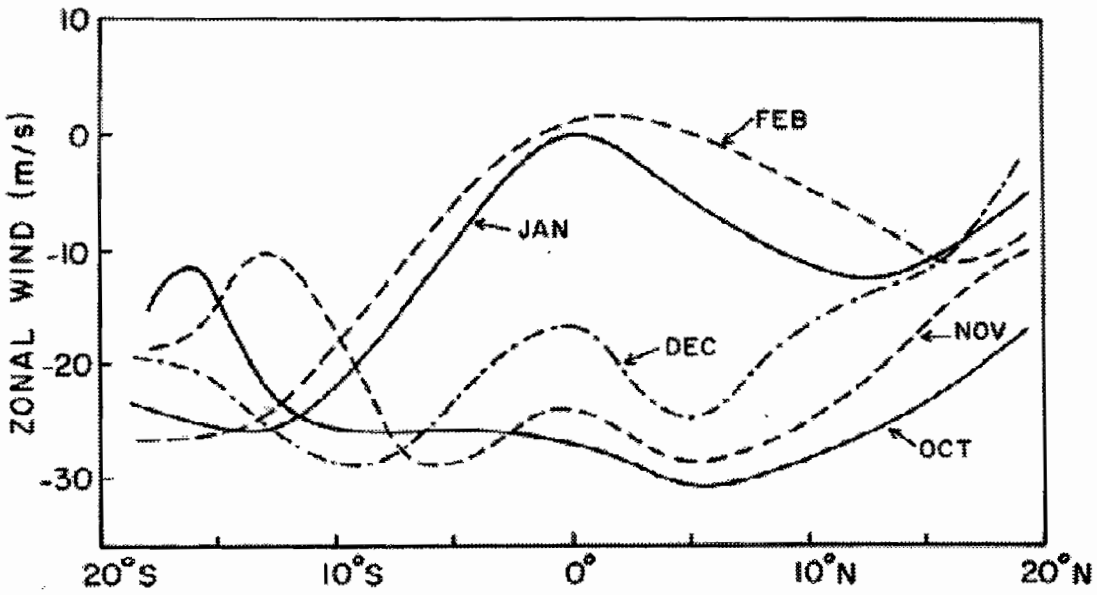


(top) High Resolution Doppler Imager (HRDI) measurements of the zonally averaged zonal wind in the tropical stratosphere and mesosphere from 1992 to 1998. The lower panel shows the QBO from 20 to 40 km as descending easterly (green to blue) and westerly (red to yellow) winds. In the lower mesosphere (60–80 km) the wind structure is dominated by the semiannual oscillation (SAO). (bottom) Removing the SAO and the annual oscillation (upper panel) shows that the influence of the QBO extends into the mesosphere (80 km). Mesospheric wind changes coincide with the change of the QBO winds near 30 km. This coupling between the mesosphere and the stratosphere is believed to be caused by small-scale upward propagating gravity waves, indicated by wavy arrows. From the UARS brochure, modified from the original provided by M. Burrage and D. Ortland. Courtesy M. Schoeberl.



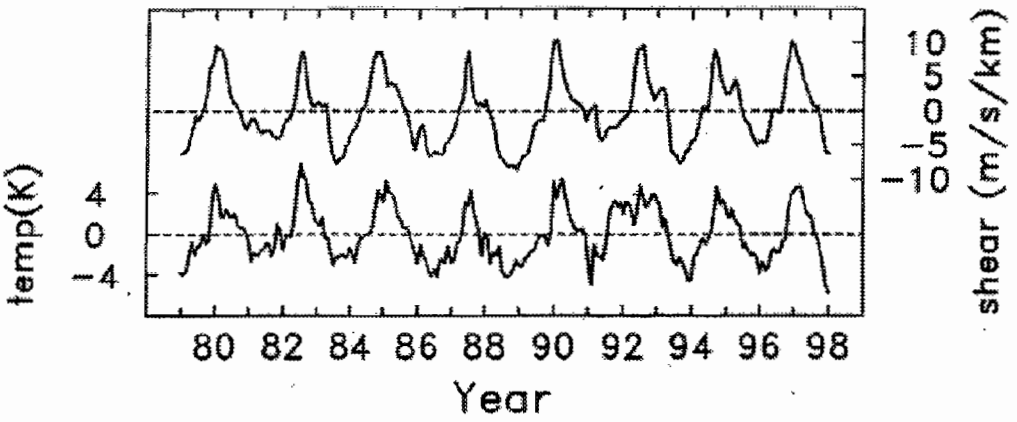
Time series of 30 mb zonal wind at various latitudes, comparing stations of similar latitude but different longitude. Positive values denote westerlies. (Refer to Table 1 for exact station locations.)


PRESSURE (mb)
MEAN HEIGHT, TROPICS (km)
10
15
20
25
30
40
50
60
70
80
90
100
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
20
17.5
15
12.5
10
7.5
5
2.5
25°
20°
15°
10°
5°
0°
5°
10°
15°
20°
25°
NORTHERN HEMISPHERE
SOUTHERN HEMISPHERE



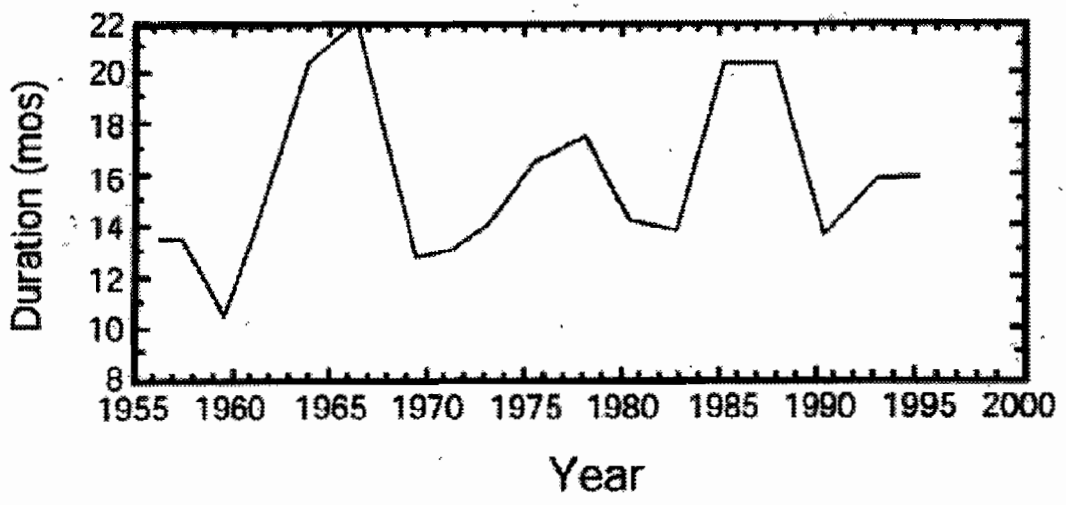
Profiles of the mean zonal wind at 30 mb (deduced from balloon observations) for each month from October 1979 through February 1980.

20

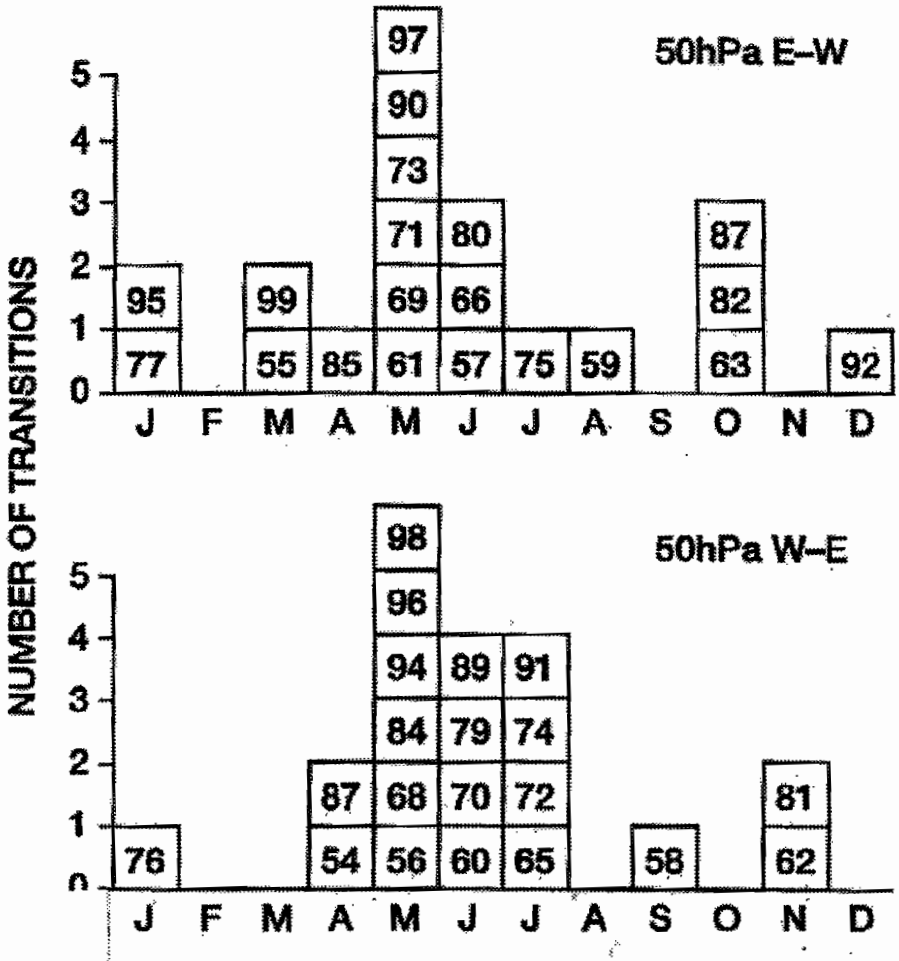


Equatorial temperature anomalies associated with the QBO in the 30- to 50-hPa layer (bottom) vertical wind shear (top curve).

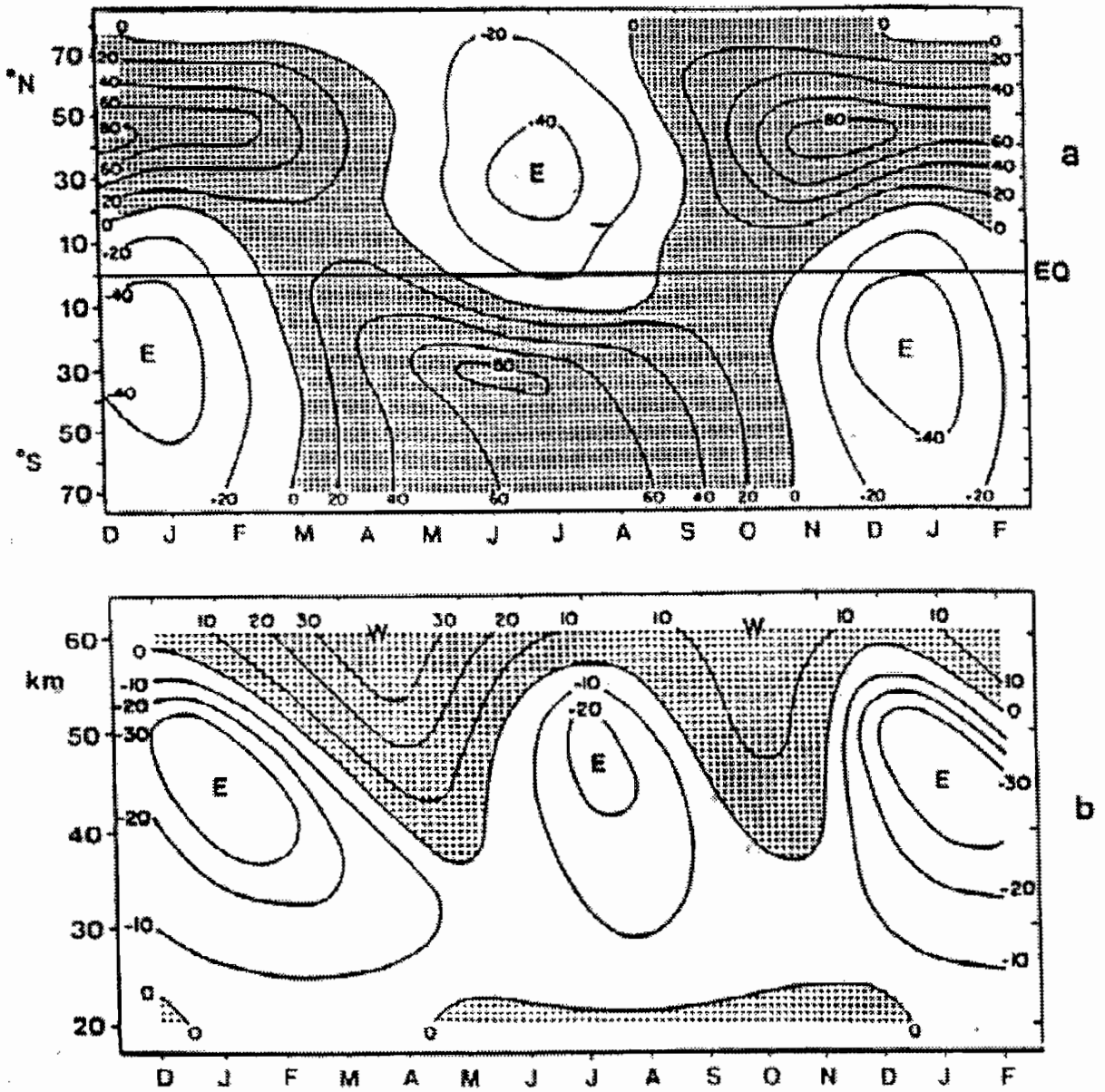
21



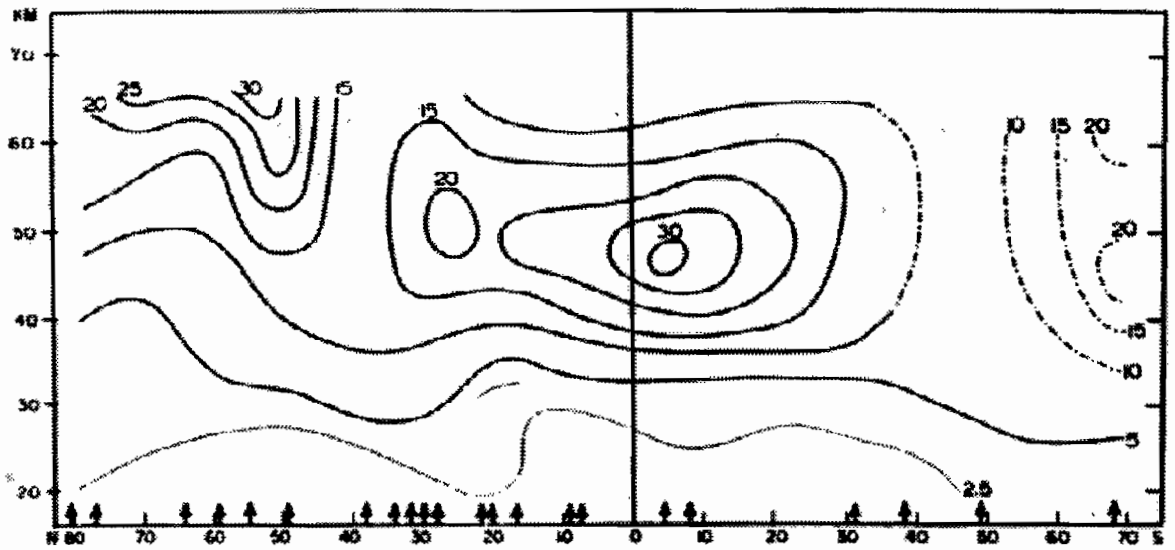
Duration of equatorial westerlies



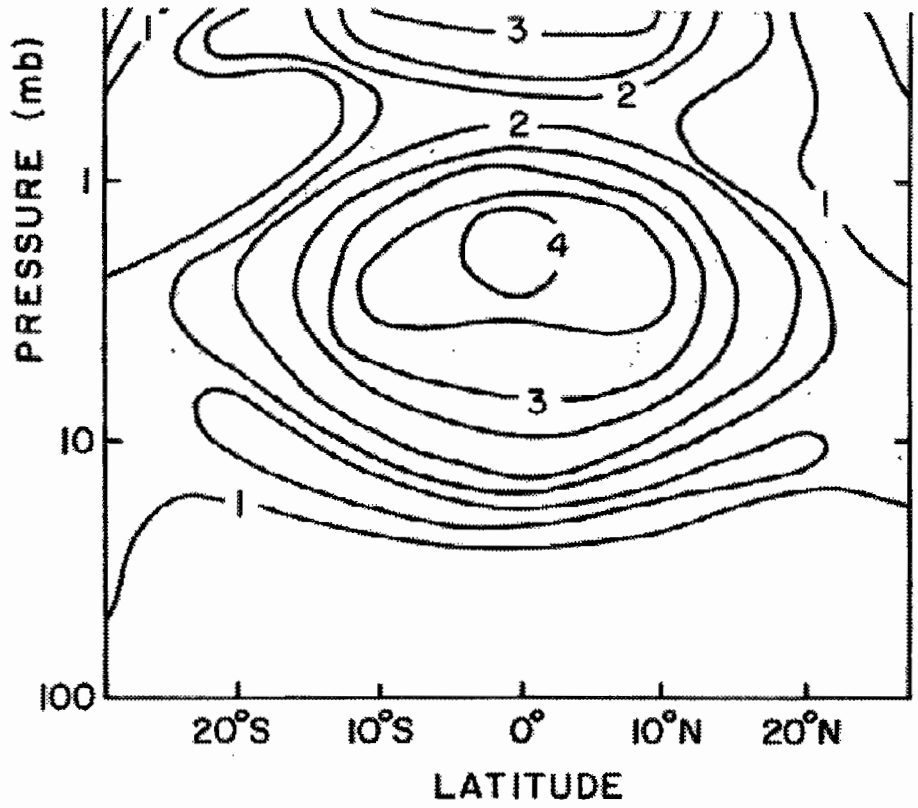
Histograms of the number of transitions (zero crossings) at 50 hPa grouped by month. Individual years are listed in the boxes. Easterly to westerly transitions are displayed in the top panel, while westerly to easterly transitions are shown in the bottom panel. After *Pawson et al.* [1993].



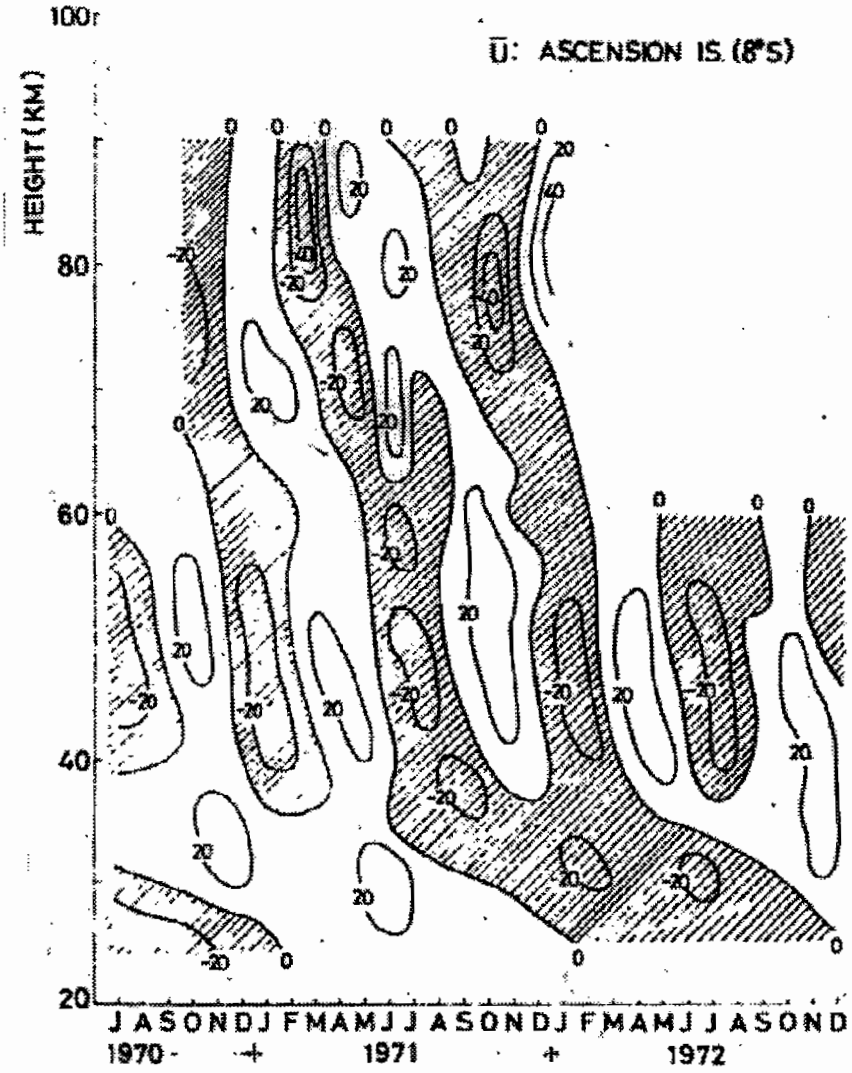
(top) Latitude-time section of the climatological annual march of zonal-mean zonal winds at 50-km height determined from rocketsonde observations at several stations. Contour interval is 20 m s^{-1} and regions of westerlies are shaded. (bottom) Altitude-time section of the annual march of equatorial zonal wind determined from interpolation of observations at Kwajalein (8.7°N) and Ascension Island (7.9°S). Contour interval is 10 m s^{-1} and regions of westerlies are shaded. Reproduced from Delisi and Dunkerton (1988b) and based on an earlier figure from Belmont et al. (1975).



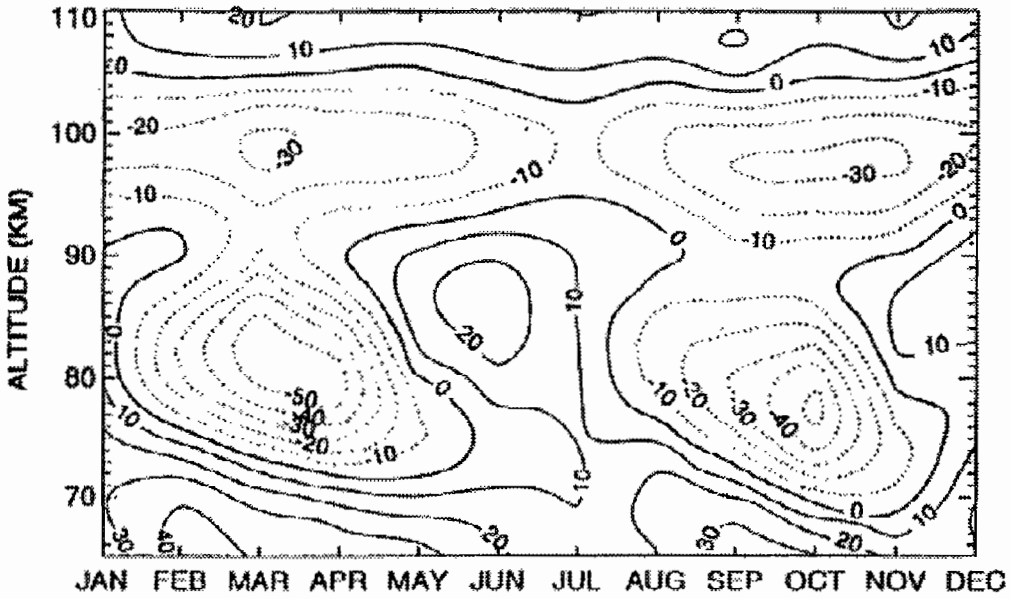
(Top) The amplitude of the quasi-biennial oscillation of the zonal wind as a function of height and latitude determined by Belmont et al. (1974) from rawindsonde and meteorological rocket data at 24 stations. The amplitude was computed by fitting a 29 month harmonic to the data. Contour labels are in m/sec. The arrows at the bottom show the station locations. (Bottom) The amplitude of the semiannual oscillation of the zonal wind. Also reproduced from Belmont et al. (1974).



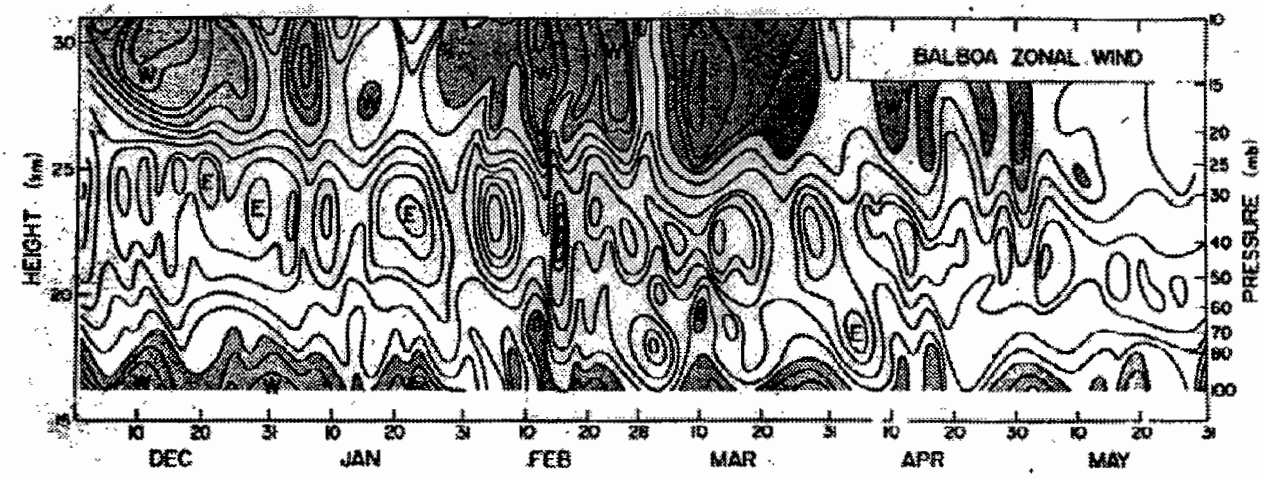
Amplitude of the semiannual harmonic of the zonally-averaged temperature



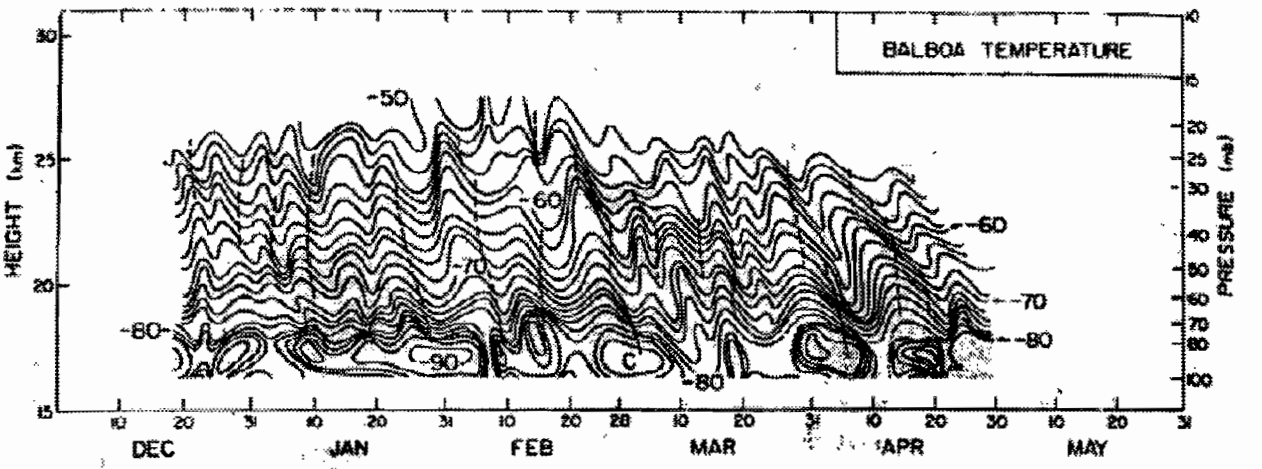
Vertical time-section of the monthly mean zonal wind at Ascension Island (8°S, 14°W) for the period July 1970–December 1972. The mean and annual component have been removed. Units are m s^{-1} and the easterlies are shaded.



Annual march of the zonally-averaged equatorial zonal wind measured by the HRDI Doppler radiometer. The contour intervals is 10 m s^{-1} and dashed contours denote easterly winds. Reproduced from Garcia et al. (1997).

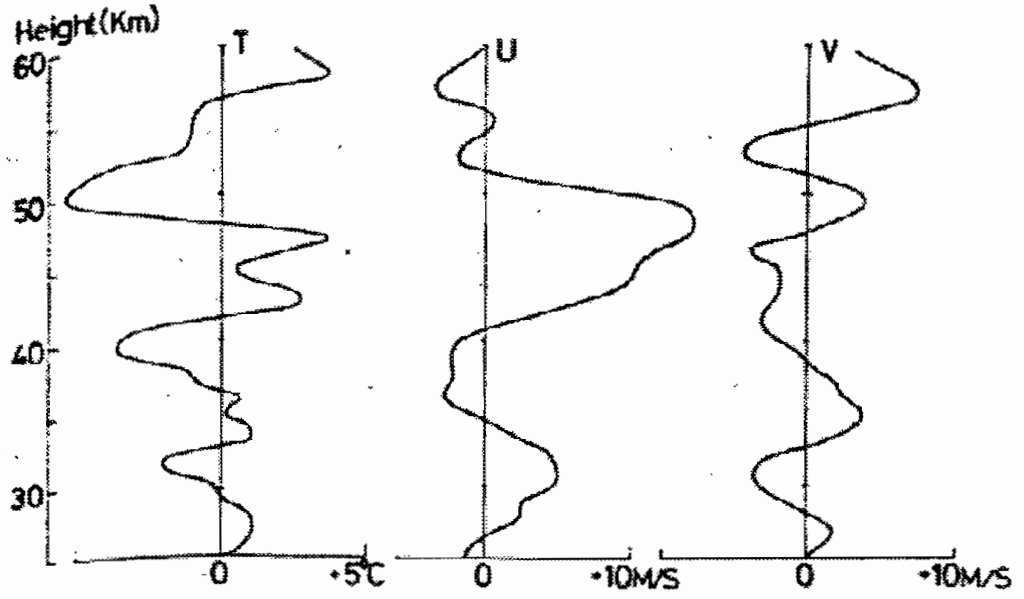


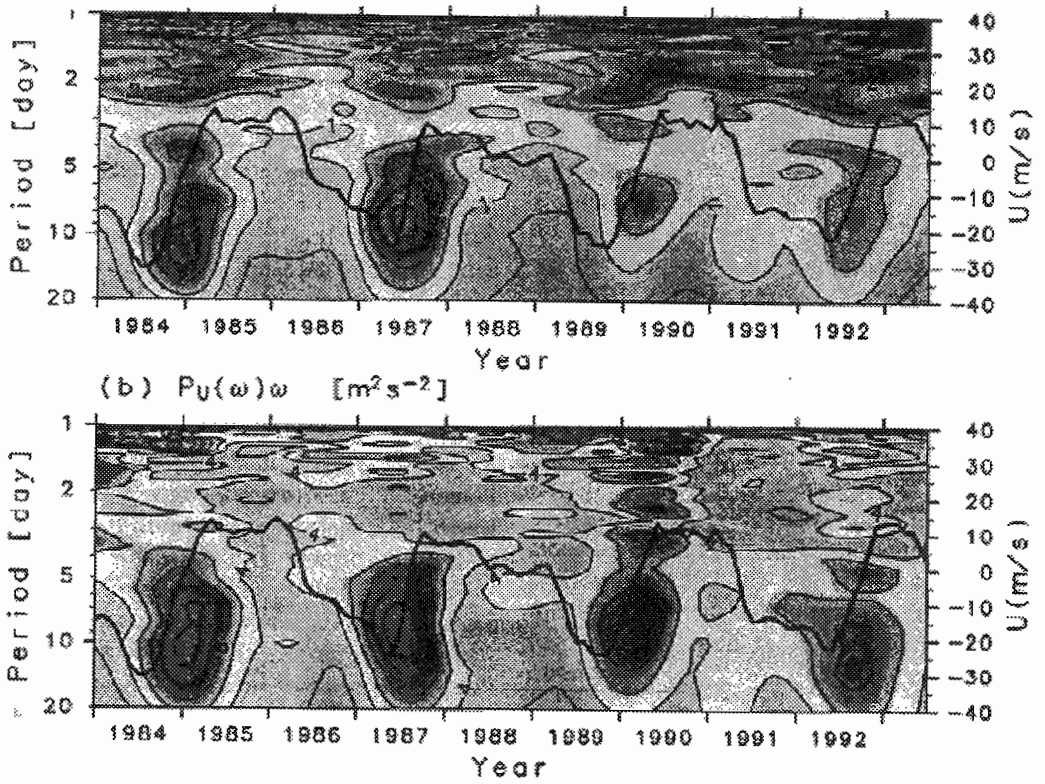
Time-height section of zonal wind at Balboa, Canal Zone (9°N). Isotachs are placed at intervals of 5 m sec^{-1}. Westerlies are shaded.



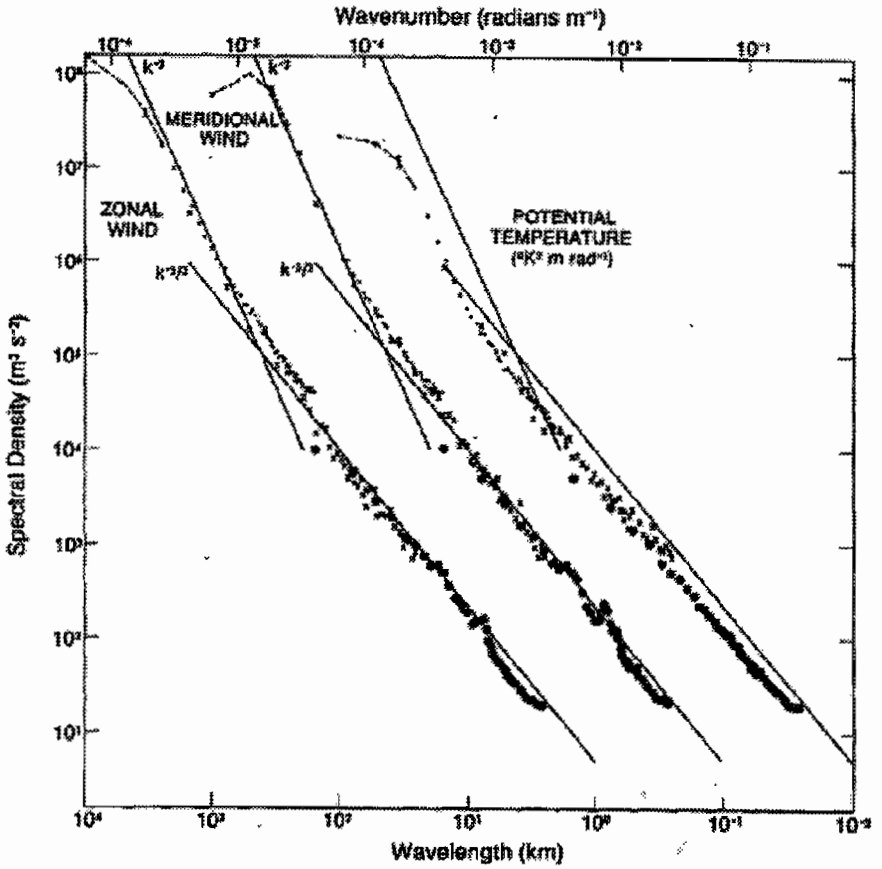
Time-height section of temperature at Balboa. Isotherms are placed at intervals of 2°C. Temperatures below −80°C are shaded. Dashed lines represent axes of the more prominent easterly fluctuations

2-DAY DIFFERENCE (Feb.17-19, 1971)

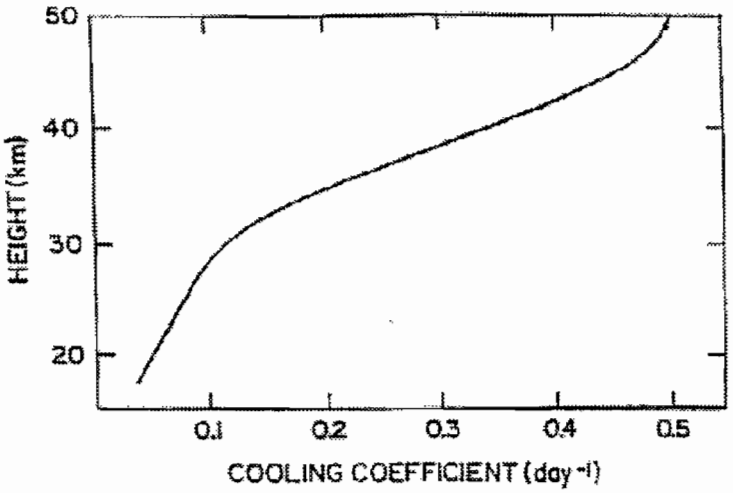




Power spectra for (a) T and (b) u fluctuations at Singapore as a function of time, averaged over 20–25 km. Contour interval 0.5 K^2, and 2 $(m\ s^{-1})^2$, respectively.

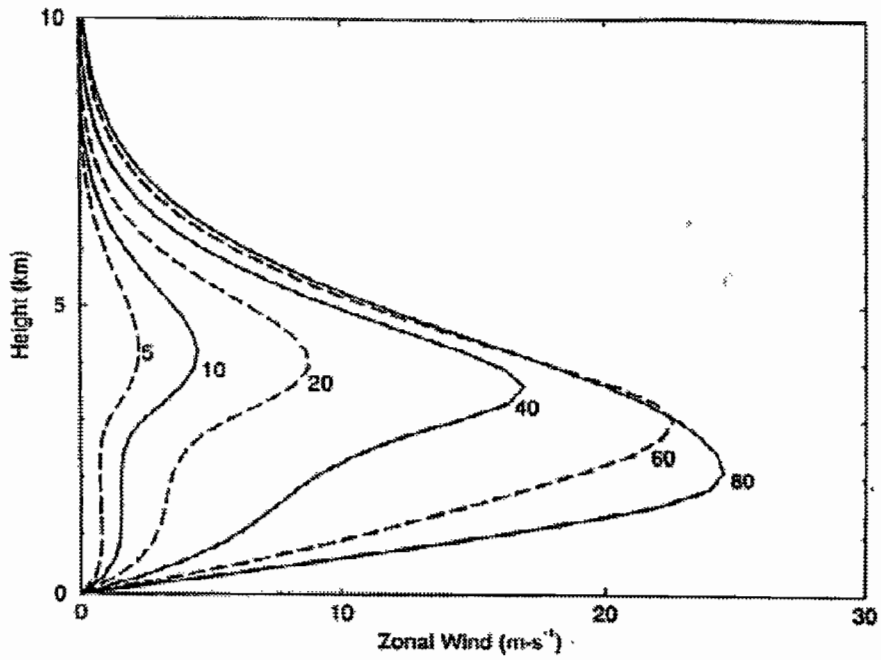


Variance power spectra of wind and potential temperature near the tropopause from GASP aircraft data. The spectra for meridional wind and temperature are shifted one and two decades to the right, respectively; lines with slopes -3 and $-\frac{5}{3}$ are entered at the same relative coordinates for each variable for comparison.



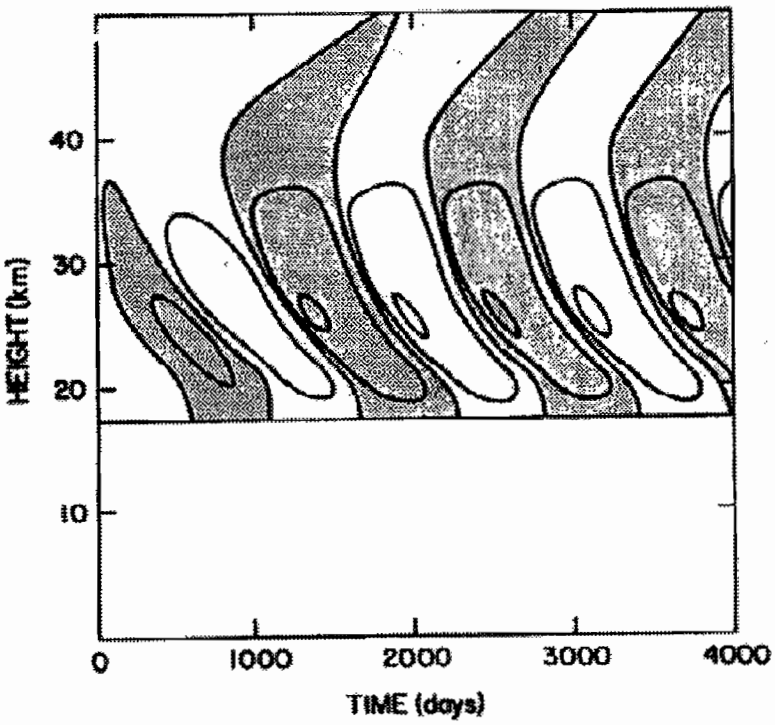
32

A reasonable representation of the damping rate of temperature perturbations due to radiative effects in the stratosphere, appropriate for temperature perturbations with long vertical scales. Reproduced from Hamilton (1982b) and based on the radiative-photochemical results of Blake and Lindzen (1973).

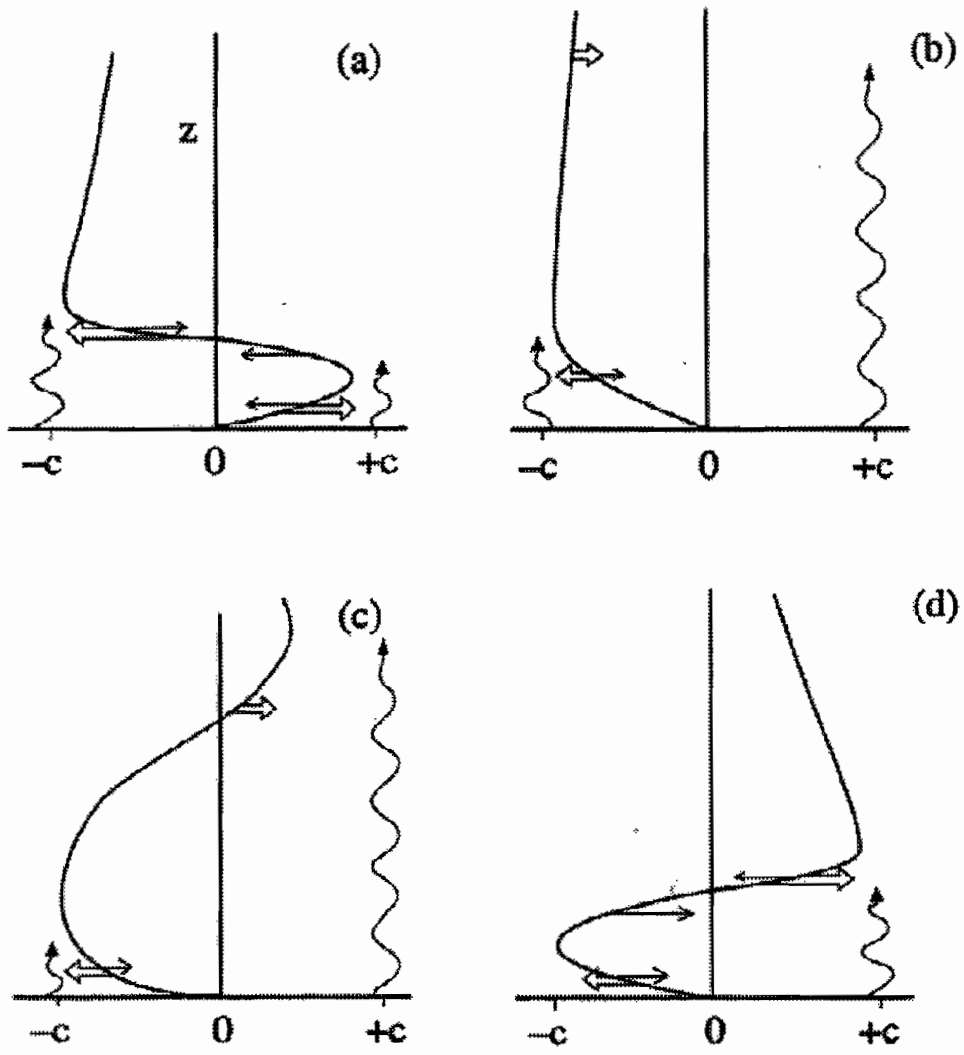


33

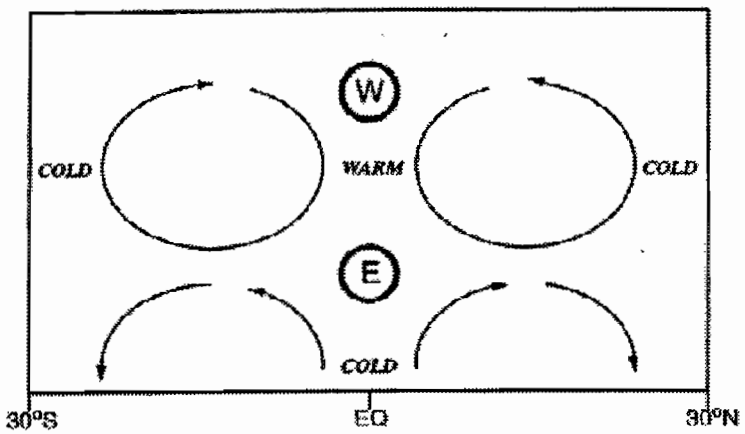
The mean zonal wind in a simple Boussinesq one-wave Plumb model plotted at 5, 10, 20, 40, 60 and 80 days of integration starting from zero mean wind. The parameters employed are $F(z=0) = 0.02\ \mathrm{m^2\ s^{-2}}$, $k = 2\pi/4 \times 10^7 \mathrm{m^{-1}}$, $k/\omega = 30\ \mathrm{m\ s^{-1}}$, mean flow diffusivity $0.3\ \mathrm{m^2\ s^{-1}}$, $N = 0.02\ \mathrm{s^{-1}}$, $\alpha = 0.01\ \mathrm{day^{-1}}$ for $z < 3$ km and $\alpha = 0.01\ (1 + 3.3z)\ \mathrm{day^{-1}}$ for $z > 3$ km.



Height-time evolution of the mean flow in a two-wave Plumb model. The model employed is a version that takes into account compressibility. The model is forced with two waves with phase +25 m s^{-1} and −25 m s^{-1} and the wave fluxes are specified at the lower boundary located at 17-km height. The contour interval is 10 m s^{-1} and regions of negative wind are shaded. Reproduced from Hamilton (1982b), where further details can be found.



Schematic representation of the evolution of the mean flow in *Plumb*'s [1984] analog of the QBO. Four stages of a half cycle are shown. Double arrows show wave-driven acceleration, and single arrows show viscously driven accelerations. Wavy lines indicate relative penetration of eastward and westward waves. After *Plumb* [1984].



Schematic view of how the QBO may affect the mean meridional circulation in the tropical stratosphere. The "E" and "W" refer to the maximum easterly and westerly zonal winds, and the peak warm and cold QBO anomalies at the equator are also marked. Adapted from Reed (1965).

Sources for Figures

Fig. 1. S. Chapman and R.S. Lindzen, "Atmospheric Tides", 1970. courtesy D. Reidel Publishing.

Fig. 2. Middle Atmosphere Program Handbook 16, 1985.

Fig. 3. H. Rhiel "Climate and Weather in the Tropics", 1979. courtesy Academic Press.

Fig. 4. World Meteorological Organization Technical Document 1043, 2000.

Fig. 5. World Climate Research Programme publicity brochure, 2000.

Fig. 6. World Meteorological Organization Technical Document 1043, 2000.

Fig. 7. J.M. Wallace, Rev. of Geophys. Space Phys., 11, 196, 1973.

Fig. 8. Courtesy Northwest Research Associates.

Fig. 10. "The Eruption of Mt. Krakatoa and Subsequent Phenomena" Trubner and Company, 1888.

Fig. 11. K. Hamilton, Atmosphere-Ocean, 36, 319, 1998. courtesy Canadian Meteorological and Oceanographic Society.

Fig. 12. R.J. Reed, US Navy Report, 1964.

Fig. 13. Provided by B. Naujokat.

Fig. 14. R.J. Reed, J.Atmos. Sci., 22, 331, 1965. courtesy American Meteorological Society.

Fig. 15. M. Baldwin et al., Rev. Geophys., 39, 179, 2001.

Fig. 16. NASA publicity brochure.

Fig. 17. T.J. Dunkerton and D.P. Delisi, J. Atmos. Sci., 42, 376, 1985, courtesy American Meteorological Society.

Fig. 18. J.M. Wallace, Rev. of Geophys. Space Phys., 11, 196, 1973.

Fig. 19. K. Hamilton "A Review of Observations of the Quasibiennial and Semiannual Oscillations of Wind and Temperature in the Tropical Middle Atmosphere", pp. 19-29 of "Transport Processes in the Middle Atmosphere", 1987. courtesy D. Reidel Publishing.

Fig. 20. M. Baldwin et al., Rev. Geophys., 39, 179, 2001.

Fig. 21. M. Salby and P. Callaghan, J. Climate, 13, 2652, 2000. courtesy American Meteorological Society.

Fig. 22. M. Baldwin et al., Rev. Geophys., 39, 179, 2001.

Fig. 23. D.P. Delisi and T.J. Dunkerton, J. Atmos. Sci., 45, 2772, 1988. courtesy American Meteorological Society.

Fig. 24. A.D. Belmont et al., Quart. J. Roy. Meteor. Soc., 100, 203, 1974. courtesy Royal Meteorological Society.

Fig. 25. K. Hamilton and J.D. Mahlman, J. Atmos. Sci., 45, 3212, 1988. courtesy American Meteorological Society.

Fig. 26. I. Hirota, J. Atmos. Sci., 35, 714, 1978. courtesy American Meteorological Society.

Fig. 27. R.R. Garcia et al., J. Geophys. Res., 102, 26019, 1997.

Fig. 28. J.M. Wallace, Rev. of Geophys. Space Phys., 11, 196, 1973.

Fig. 29. I. Hirota, J. Atmos. Sci., 35, 714, 1978. courtesy American Meteorological Society.

Fig. 30. M. Baldwin et al., Rev. Geophys., 39, 179, 2001.

Fig. 31. G.D. Nastrom and K.S. Gage, J. Atmos. Sci., 42, 950, 1985. courtesy

American Meteorological Society.

Fig. 32. K. Hamilton, Atmosphere-Ocean, 36, 319, 1998. courtesy Canadian Meteorological and Oceanographic Society.

Fig. 33. K. Hamilton, Atmosphere-Ocean, 36, 319, 1998. courtesy Canadian Meteorological and Oceanographic Society.

Fig. 34. K. Hamilton, J. Atmos. Sci., 39, 1881, 1982. courtesy American Meteorological Society.

Fig. 35. D.G. Andrews et al., "Middle Atmosphere Dynamics", 1987. courtesy Academic Press.

Fig. 36. K. Hamilton, Atmosphere-Ocean, 36, 319, 1998. courtesy Canadian Meteorological and Oceanographic Society.

Chapter 2

Observational Techniques

Robert Vincent,

University of Adelaide, Australia

1 Introduction

Improvement in our understanding of atmospheric processes goes hand-in-hand with the development and application of new techniques for observing the atmosphere. There has been rapid development in the past few decades as new methods for sounding the atmosphere and ionosphere have become available. Techniques range from highly detailed "single-shot" observations made with rocket- and balloon-borne instruments to continuous observations made with ground-based instruments using the radiowave and optical spectrum. Recently, satellite observations have begun to play an important role in providing a global perspective.

It should be recognized that each technique has its own advantages and disadvantages for sounding the atmosphere and it is now increasingly common to "cluster" different instruments to make use of the best attributes of each. The following discussion is meant to serve as an introduction to the more commonly used techniques; it is not meant to be exhaustive.

2 Radiosonde Observations

A radiosonde is a balloon-borne package that makes observations as a function of height of temperature, humidity, and pressure. Radiosondes released daily by weather services around the world provide the primary information for studying the thermal and dynamical state of the atmosphere up to heights near 30 km.

The sonde, which is approximately the size of a shoebox, is suspended by up to 100m below the balloon to reduce the effects of the wake as the balloon ascends. Using an onboard radio transmitter the information is telemetered regularly back to a ground receiving station. The rate at which data are received depends on the system used, but typically data are acquired at 10-sec intervals. Balloons ascend at about 5 ms^{-1} so the equivalent height resolution is about 50 m. However, data may not be finally reported and archived at this rate, and information at 'significant levels' only may actually be saved, which can reduce the effective height resolution to $\sim$0.5-1 km.

The type of devices used to measure temperature and humidity depend on the manufacturer. For example, the widely used Vaisala sondes use a capacitive bead to measure temperature, with an accuracy of 0.2-0.3°C, and a thin film capacitor to measure humidity. Some stations also have the capability of measuring winds in the upper atmosphere if the position of the balloon is known as it ascends. One way is

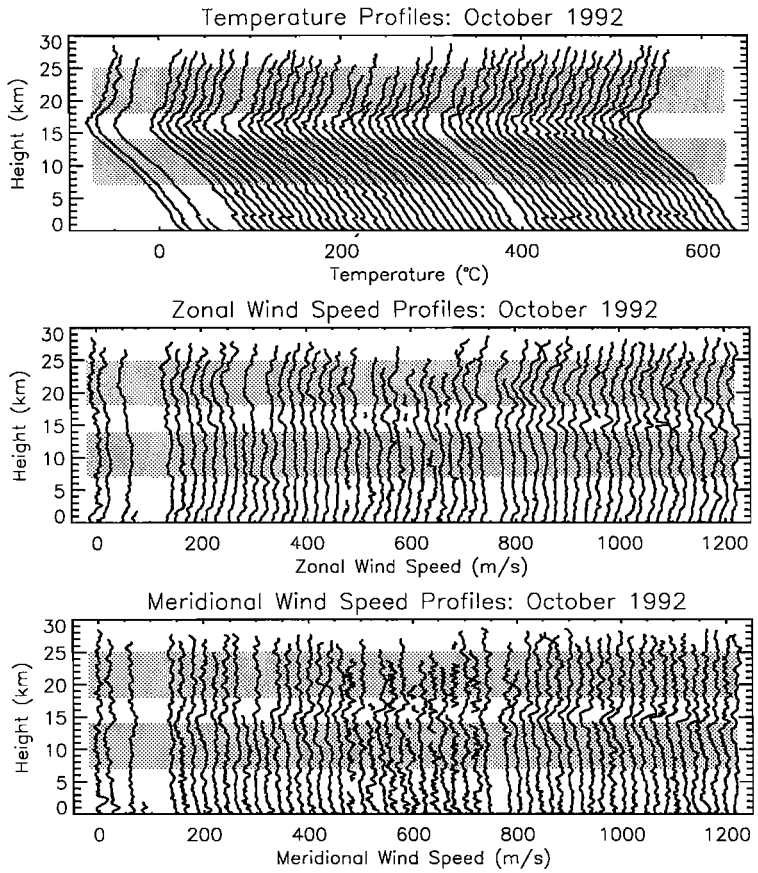


Figure 1: A sequence of radiosonde observations from Cocos Islands (12°S, 97°E) (After Allen, 1996)

to use a ground-based radar to track the sonde. Alternatively, the sondes themselves carry a radio receiver capable of receiving signals from various navigation systems, such as Loran or GPS, and transmitting their position to the ground station, which computes the wind speed and direction.

Weather agencies make synoptic observations twice a day at 0000 UTC and 1200 UTC from a network of stations spread across the globe. The actual height reached by a balloon before it bursts depends on such factors as the quality of the balloon and meteorological factors such as icing. Figure 1 shows a sequence of synoptic soundings made from the Cocos Island (12°S, 97°E) and illustrates some of the features of the tropical troposphere and lower stratosphere, including the cold and high tropopause at ~17 km. It shows that the upper level of most soundings is between 25 and 30 km, but that some balloons burst at much lower heights. Small wave-like features evident in the soundings, especially in the stratosphere, are interpreted as being due to atmospheric waves and radiosonde measurements have proved to be an important

source of information on waves in the tropical lower atmosphere (e.g. Tsuda et al. 1994; Vincent and Alexander, 2000).

3 Radars

Ground-based radars (*RA*dio *D*etection and *RA*nging) enable atmospheric dynamics to be studied with a time resolutions as short as a few minutes. The most powerful systems can obtain information from near the ground to heights near 100 km, albeit with a 'gap' between 25 and 60 km. In principle radars can measure the three-dimensional wind field and other important quantities, such as turbulence parameters. They are an important source of information on atmospheric waves, such as gravity waves and tides.

There are many types of radar, each with their own particular characteristics. Here we concentrate on describing the types most commonly used for studies of the atmosphere from the ground up to heights near 100 km, with an emphasis on radars used to study the middle atmosphere. This height range includes the lower ionosphere, which extends upward from about 50 km during the day and from about 80 km at night. Table 1 shows the main radar types, together with typical frequency range of operation and their nominal height coverage. The frequency ranges shown are only approximate, but cover the majority of systems. In the table MST stands for *M*esosphere-*S*tratosphere-*T*roposphere. Less powerful radars of this type are called Stratosphere-Troposphere (ST) radars.

The elements of a typical radar are shown in schematic form in Figure 2. Basically, a radar consists of a master oscillator/synthesiser to generate the frequency of the transmitted radiowave and other reference frequencies, a power amplifier for the transmitter, a transmitting antenna, one or more receiving antennas to collect the backscattered signals, one or more receiver to amplify the signals and a data acquisition system to digitise, integrate and store the signals for further computer manipulation. Radars are fully phase coherent so both phase and amplitude information is acquired. Most atmospheric radars have computing systems (often PCs) to analyse and present the data in almost real time. The majority of radars use co-located transmitting and receiving antennas. Many different antenna configurations are possible. A single antenna can be used for both transmission and reception using a transmit/receive (T/R) switch, or the sub-modules of the transmitting antenna might be used, or separate transmitting and receiving antennas are used in some

Table 1: Atmospheric radars

Type	Frequency (MHz)	Height Coverage (km)
Medium Frequency (MF)	2-3	60-100 (day) 80-100 (night)
Meteor wind radar (MWR)	30-50	80-105
MST	50	2-25 50-100 (day)
Incoherent Scatter (ISR)	430	60-100 (day)

systems, in which case T/R switches are not required.

Almost all atmospheric radars operate in pulsed mode so that the overall time of flight of a pulse determines the distance or range to a given echoing region (Figure 3). Pulses of duration Δt_{T_x} are transmitted at an interpulse period of T_{ipp}. A pulse that takes a time T_R to travel to and from the radar travels a distance cT_R, where $c = 3 \times 10^8 \mathrm{ms}^{-1}$ is the speed of light, and the range is $R_o = cT_R/2$. Bandwidth limitations mean that the voltage detected at the receiver output is averaged over a time interval Δt_{Rx}. The receiver bandwidth is matched to the transmitted pulse length ensure the optimum signal to noise, so that $\Delta t_{Tx} = \Delta t_{Rx} = \Delta t$. The echo pulse then contains information about the range interval from $R - c\Delta t/2$ to $R + c\Delta t/2$, so that the effective range resolution is $\Delta R = c\Delta t/2$.

Clear air radar echoes come from fluctuations in refractive index n with scale of $\lambda/2$, where λ is the wavelength. At VHF (30-300 MHz) frequencies and higher the refractive index of air is given by

$$n - 1 = 0.373\frac{e}{T^2} + 77.6 \times 10^{-6}\frac{p}{T} - 40.3\frac{N_e}{f^2} \quad (1)$$

where e is the partial pressure of water vapour in hPa, p is the total atmospheric pressure in hPa, T is the temperature in K, N_e is the electron density in m^{-3}, and f is the radar frequency in Hz. Water vapour has a small scale height, so the first or 'wet' term is most important in the lowest part of the atmosphere. The second or 'dry'

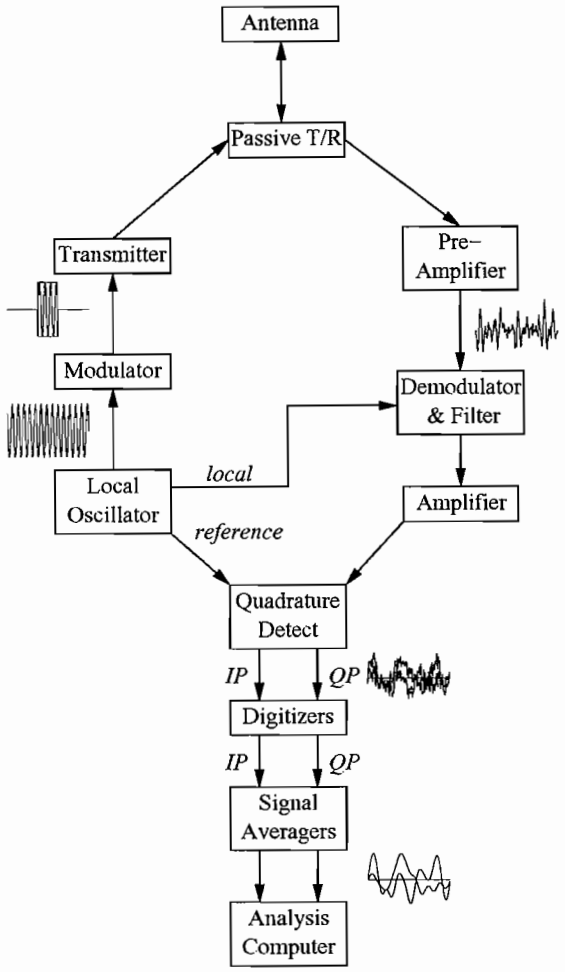


Figure 2: A radar in schematic form together with signal waveforms (after Tsuda, 1989)

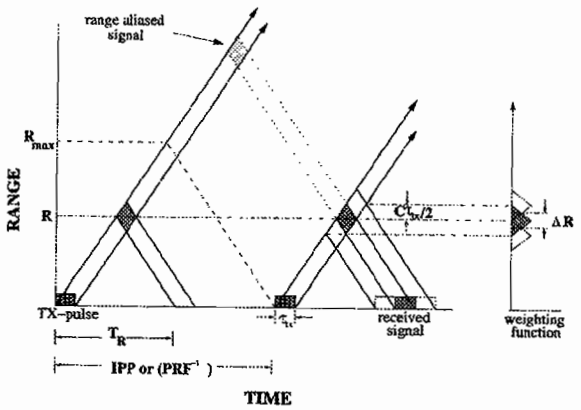


Figure 3: Range-time diagram of pulsed radar operation.

term includes the effects of temperature fluctuations, and dominates in the middle atmosphere. In the lower ionosphere (D-region), fluctuations in ionization determine the refractive index. The relative contribution of these three terms is illustrated in Figure 4 for a radar operating near 50 MHz.

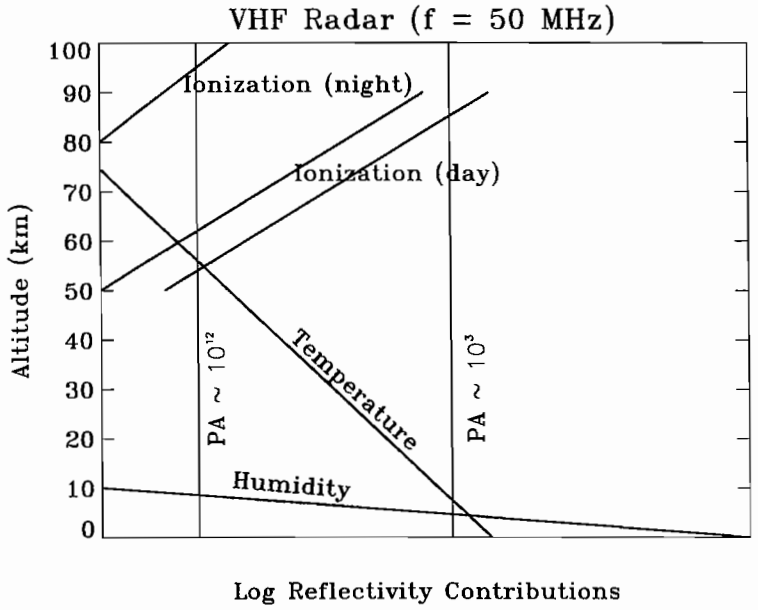


Figure 4: Altitude variation of relative refractivity contributions for VHF radars (after Röttger, 1984).

The actual received power and hence height coverage of a radar, depends on a number of factors. These include instrumental factors, such as the transmitter power and size of the antennas, and on atmospheric factors, such as the strength of turbulence. Two limiting processes may be considered for echoes received from a range R using a pulse length of ΔR. First, in the situation known as volume scatter, isotropic scatterers fill the volume defined by the radar beam and the pulse length the received power is given by the radar equation:

$$P_r = \frac{\pi P A \alpha \Delta R}{64 R^2} \eta \tag{2}$$

where P is the mean transmitted power, A is the effective antenna area, α is the overall efficiency of the radar, and η is the volume reflectivity. Sometimes, however, echoes come from a sharply bounded discontinuity in refractive index ("Fresnel irregularity"), in which case the radar equation is:

$$P_r = \frac{P A^2}{4 \lambda^2 R^2} |\rho|^2 \tag{3}$$

with λ as the radar wavelength and ρ is the amplitude reflection coefficient (Gage and Balsley, 1980). Partial reflections from horizontally stratified irregularities will be most important when the radar beam is pointing vertically. In between the extreme situations described by (2) and (3) different scattering/reflecting mechanisms can occur. Surprisingly, there is still much that we do not know about the mechanisms that give rise to the irregularities that cause radar scatter.

3.1 MST Radars

The common factor of PA or power-aperture product evident in (2) and (3) is often used as a figure of merit and used as a guide to the height coverage for a given radar. Two situations are shown in Figure 4 for a small system with a PA of $\sim 10^3$ Wm^2 and a very large radar with a PA of $\sim 10^{12}$ Wm^2. The small system would be able to make measurements only up to heights near 10 km, while the very large system can observe throughout much of the atmosphere. At present only the 50 MHz MST radar at Jicamarca, Peru, has this capability. However, it is important to realize that at any given moment the actual height coverage depends on the presence of suitable scattering irregularities, which means on the strength of, for example, turbulence which can be highly variable in both space and time. The pointing direction of the radar beam is also a factor. As the beam is steered away from the zenith the received power decreases since the effects of specular reflections becomes less important and isotropic scattering dominates.

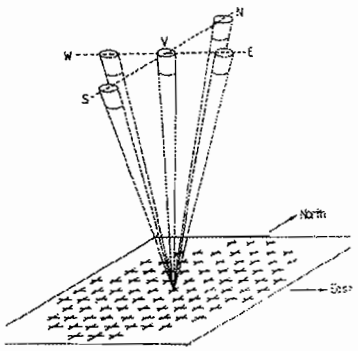


Figure 5: Conceptual VHF radar array and beams

Unlike weather radars, which use mechanically steered dish antennas, atmospheric radars use fixed arrays of smaller antennas. The beam is pointed in different directions by introducing suitable electronic phasing between rows of antennas. For a given

pointing direction atmospheric motions induce Doppler shifts in the backscattered signals. For a radar with a sufficiently narrow beam pointing at an angle θ from the zenith at an azimuth of ϕ measured clockwise from North, the component of velocity along the beam is

$$v_r = u \sin\phi \sin\theta + v \cos\phi \sin\theta + w \cos\theta. \tag{4}$$

The use of at least three non-coplanar beams results in simultaneous equations for v_r, which can then be solved to determine the three-dimensional wind vector $\underline{V} = (u, v, w)$. An atmospheric radar using the Doppler technique uses a minimum of three beams to determine $\underline{V}$, but often 5 beams are used (Figure 5). Typical values for θ lie between 7 and 15° for routine observations at VHF. Since different volumes are being probed there is an assumption that the wind field is horizontally homogeneous.

The radar echoes can be analyzed in either the frequency or time domains. Analysis in the frequency domain requires an estimate of the power spectrum $S(f)$, an idealized example of which is shown in Figure 6. The atmospheric return has a Gaussian form with a Doppler shift of f_d and width Δf. P_N indicates the noise level. The radial velocity is related to f_D by

$$v_r = -\frac{4\pi}{\lambda} f_D \tag{5}$$

In the time domain the complex autocorrelation $\rho(\tau)$, which is Fourier transform of the normalised power spectrum, is used, with the slope of the phase of ρ giving the Doppler shift (Woodman and Guillen, 1974).

Other important information about the atmosphere can be derived from either $S(f)$ or $\rho(\tau)$. The width of the spectrum Δf or equivalently the width of the autocorrelation function $\Delta\tau$ are partly determined by the strength of the random motions or turbulence. Extracting this information requires care, as the strength of the background wind and its shear with height also influence the spectral width (e.g. Hocking, 1985). Alternatively, the power returned by the scatterers can be used to find the rate of energy dissipation by turbulence.

Radars that use the Doppler technique to measure winds need narrow beams. Hence they require antennas with dimensions large compared with the radar wavelength. For MST-type radars operating at frequencies near 50 MHz (wavelength $\sim$ 6 m) antenna arrays have dimensions larger than 50 m on a side. The Jicamarca radar has dimensions 150 m, producing a beam width of $< 1°$.

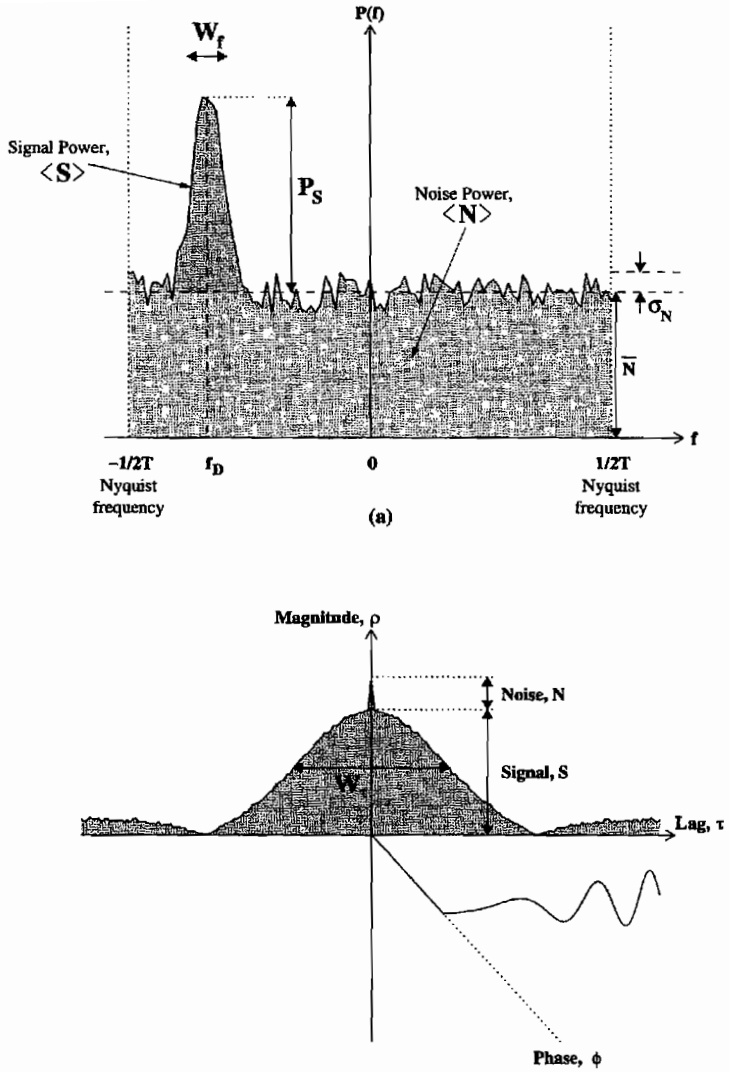


Figure 6: Idealized Doppler spectrum of atmospheric returns.

3.2 MF Radars

Much of the discussion concerning MST and ST radars in the previous section also applies to radars operating in the MF band near 2 MHz. These radars detect echoes from fluctuations in electron density in the lower ionosphere. However, it should be recognized that birefringence in the ionosphere due to the presence of Earth's magnetic field becomes very important at these frequencies, which means that radiowaves can propagate in either of the ordinary or extraordinary characteristic modes. Equation (1) has to be modified to account for this effect.

The approximately f^{-2} factor in (1) means that ionospheric reflections are much easier to observe at MF than they are at VHF, especially at night. This means that is less expensive to build an MF radar to study the dynamics of the mesosphere and lower thermosphere (Vincent, 1984). However, generating narrow beams at MF, where the wavelength is $\sim 100 - 150$ m, is difficult because of the large amounts of real estate required. Hence MF radars generally use another technique, the spaced-antenna method, to measure the wind velocity. One antenna is used for transmission and three non-collinear antennas are used for reception. The receiving antennas are usually arranged in the form of an equilateral triangle and all antennas are phased to point vertically. Radiowaves scattered from a given height form a diffraction pattern over the ground. If the atmosphere is in motion then the pattern also moves and the signal strengths at each receiving antenna show relative time displacements from which the velocity of the pattern and hence atmosphere can be calculated. While the Doppler and spaced-antenna techniques appear very different in practice they both make use of the Doppler shifts inherent in the backscattered signals and it is possible to derive the same information on winds and turbulence using both techniques (Briggs, 1980). From a practical point of view, each technique has its advantages and disadvantages.

3.3 Meteor Radars

Radiowaves from meteor wind radars are scattered from ionized trails in the upper mesosphere and lower thermosphere (MLT). The trails, formed as meteors ablate as they enter the atmosphere, usually last only a fraction of a second. The backscattered echoes suffer a Doppler shift as the trails are carried along by the neutral wind. Combining Doppler shift measurements of trails observed in different regions of the sky enables the wind velocity to be inferred. Locating the position of the trails

requires either the use of narrow beam radars (often ST radars) or multiple receive antennas arranged as an interferometer. Usually it is assumed that the vertical wind component is small and can be neglected in comparison with the horizontal wind components. Observations of other atmospheric and astronomical interest can also be made. Diffusion causes the trails to expand, so measurements of the echo amplitude decay times allows diffusion coefficients and temperature to be inferred.

3.4 Incoherent Scatter Radars

Free electrons in the ionosphere act as Thompson scatterers when the plasma frequency is smaller than the radar frequency. As the electrons in a particular volume are in random motion the scattered waves add randomly. The ionosphere is a weak plasma, so the heavy ions constrain the scattering process and the high frequency of collisions between charged and neutral particles means that the scattering process is collision dominated. ISR have to be very powerful as the incoherent scatter cross-section is very small. For example the most powerful ISR, and the only one in tropical regions, the 430 MHz Arecibo (19°N) radar has peak power of a few MW. ISR can be used to study ion drift (neutral) velocities, temperatures, electron and neutral densities.

4 Lidars

Lidars (*LI*ght *D*etection and *RA*nging) are the optical equivalent of electromagnetic radars. Monochromatic optical pulses from a laser are backscattered from atmospheric molecules and detected as a function of range. At optical wavelengths the pulses can be very short, so the range resolution is usually very good (1.5 to 150 m). Various types of lidar operation are possible. The backscattered wavelength can either the same as the transmitted wavelength (Rayleigh, Mie and resonant scattering) or at a higher wavelength (Raman and fluorescent scattering). Here we concentrate on Rayleigh and resonant scattering, which allow density and temperature measurements in the middle atmosphere.

A Rayleigh lidar consists of a laser transmitting vertically, a telescope to collect the returned light signals, which are then passed through a narrow band filter to reject unwanted stray light and focussed on a photomultiplier. Signals from the middle atmosphere are so weak that individual photons are counted. The problem or

measuring signal amplitude as a function of height reduces to counting the number of pulses from the photomultiplier in a series of range gates (Kent and Wright, 1970).

Rayleigh scattering occurs when light waves are scattered by atmospheric molecules; Mie scattering occurs when atmospheric aerosols are involved. Above heights of $\sim$ 30 to 35 km aerosol scattering is negligible compared to molecular scattering. In the regime where Rayleigh scattering dominates the lidar equation for a vertically pointing lidar is:

$$N(h) = \frac{N_o A \epsilon \mathcal{T}(0,h)}{4\pi h^2} n_o(h) \beta_R \Delta h \tag{6}$$

where N is the number of detected photons per laser pulse from a layer of thickness Δh and height h, N_o is the number of photons emitted for each laser pulse, A is the area of the collecting system (telescope), and ϵ is the overall efficiency of the optical system, including the quantum efficiency of the photomultiplier. The factors $\mathcal{T}$, n_o and β_R are, respectively, the total transmisivity of the atmosphere between the ground and height h, the number density of atmospheric molecules, and the Rayleigh backscatter cross-section.

Equation (6) is inverted to find the atmospheric density ρ at height h

$$\rho(h) = N(h) K \frac{h^2}{\Delta h} \mathcal{T}^{-1} - n(h). \tag{7}$$

Here K is a normalization coefficient that depends on the lidar-specific factors N_o, A, and ϵ. $n(h)$ is the number of photons due to electronic noise in the receiver and counting system and to sky background.

It is often difficult to calibrate lidar systems accurately so, in order to measure the absolute density profile, it is necessary to compare with an independent measurement at some convenient reference height. Usually, the calibration is made at heights near 30 km or above (where aerosol concentrations are small) using data acquired from a nearby radiosonde sounding.

Atmospheric temperature can also be obtained from the scattering profile. This is achieved through the use of the equation of state $p = R\rho T/M$ and hydrostatic equation $dp = -\rho g dz$, where p is pressure, M is the mean molecular mass, g the acceleration due to gravity and R is the universal gas constant. Eliminating p gives the temperature as a function of height

$$T(h) = \frac{\rho_1 T_1 + \frac{Mg}{R} \int_{h_1}^{h} \rho dh}{\rho(h)} \tag{8}$$

Note that an initial guess of the temperature T_1 at the top of the profile has to be made. This is usually done by using the appropriate value from a reference atmosphere such as the COSPAR international reference atmosphere (CIRA, 1986). As the integration proceeds downward any error due to incorrect choice of T_1 reduces out rapidly, so the actual starting value becomes of little consequence.

Resonant scattering occurs when the wavelength of light is close to that of an absorption line. A photon is absorbed and reemitted at the same wavelength if the life time of the excited state is short. Resonant scattering from sodium atoms in the 80-110 height region is an example where this processes is very useful for atmospheric studies. Although the abundance of Na atoms is small compared with the neutral atmospheric abundances the backscatter cross-section is some 10^{14} times larger than for an atmospheric molecule. A Na lidar with only modest power-aperture product can be used. Figure 7 shows lidar observations of both Rayleigh and Na lidar

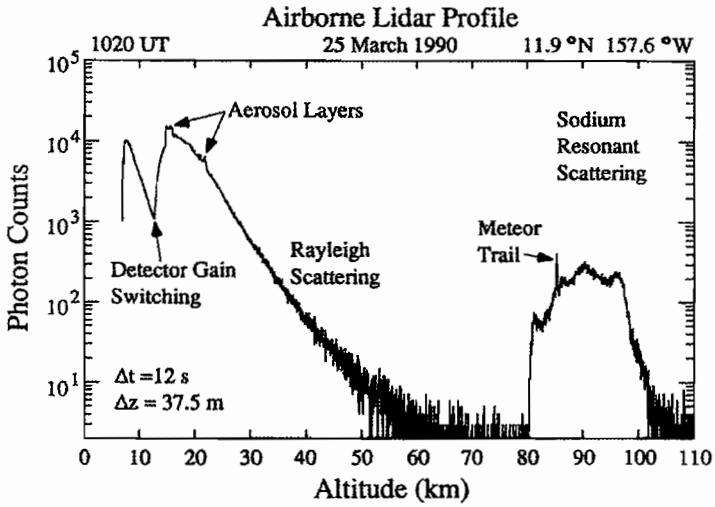


Figure 7: Lidar photon count profile from a Na/Rayleigh lidar mounted on an aircraft (after Gardner, 1991).

measurements taken from systems mounted on an aircraft during the ALOHA-90 campaign to study low-latitude atmospheric processes and aeronomy; the campaign is described in a special section of the July 1991 edition of Geophysical Research Letters (GRL). A later campaign (ALOHA-93) is highlighted in the October 15, 1995 edition of GRL. The papers in these issues provide a good introduction to the diversity of ground and airborne techniques for studying the MLT.

In an important development She et al (1990) extended Na lidar studies to include temperatures in the mesopause region. Using a two-frequency narrowband lidar sys-

tem they measured laser-induced fluorescence at two frequencies in the D_2 spectrum that are sensitive to temperature changes. One frequency is at the peak of the Na D_{2a} line and the other is at the crossover between the D_2 lines. This technique is becoming more widely used. Simultaneous observations of temperatures made with

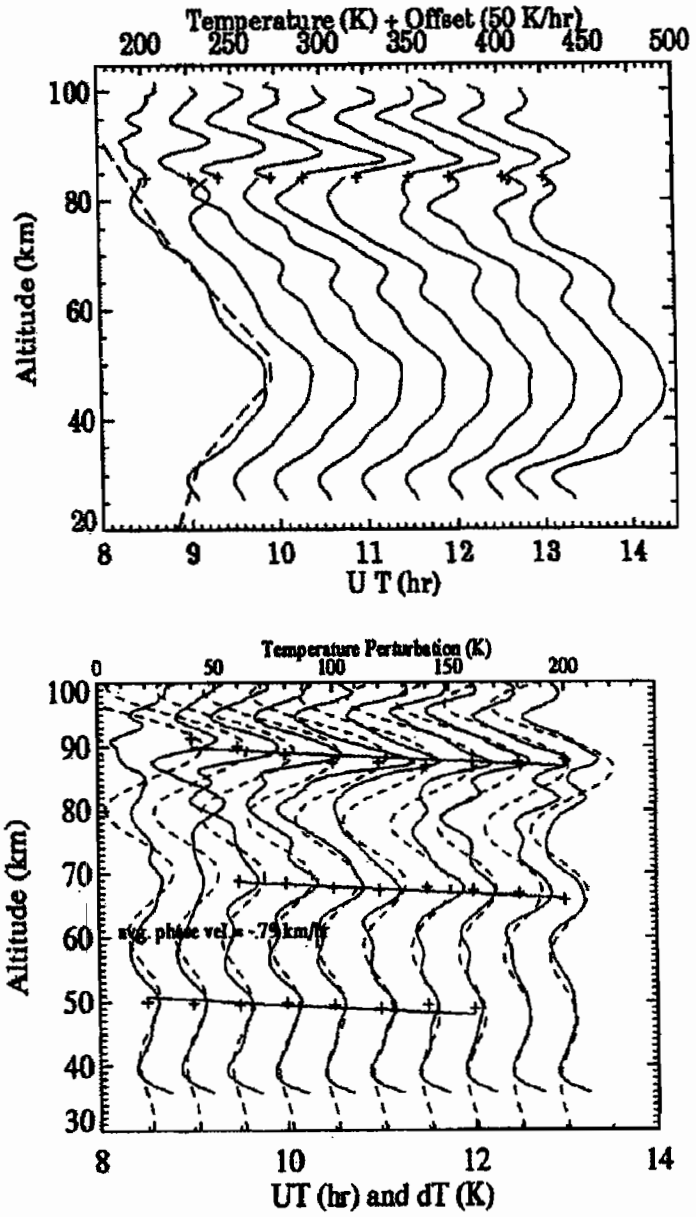


Figure 8: (Top) Composite temperature lidar profiles at 30-min intervals formed from Rayleigh and Na temperature measurements during ALOHA-93. (Bottom) Temperature perturbation profiles (After Dao et al., 1995).

both the Rayleigh and Na lidar techniques are shown in Figure 8. Taken during the ALOHA-93 campaign in Maui, Hawaii, they show strong perturbations that may be

due to tides and/or gravity waves or to gravity wave-tidal interactions (Dao et al., 1995).

5 Rocket Observations

Rocket observations provided some of the first information on the state of the middle atmosphere, especially at low latitudes. Although rocket measurements are now less routinely made they still have an important role for specialized in-situ measurements in the middle atmosphere and lower thermosphere of processes that cannot be made by remote sensing from either ground or space. The main techniques of concern here use payloads deployed from rockets just as they reach apogee, and then tracked as they fall.

A rocketsonde is an instrumented package, rather like a radiosonde, but suspended from a parachute. Temperature and pressure measurements made on the way down are telemetered back to a ground station. Tracking the position of the sonde with a high precision radar enables the winds to measured. The typical upper height limit of such a system is about 60 km. Rocketsondes formed the backbone of the meteorological rocket network (MRN) that made coordinated observations on a synoptic basis during the 1970s and 1980s. With a height resolution of $\sim$500 m they provided an important source of information on winds and waves in the stratosphere, especially of gravity waves (e.g. Eckermann et al., 1995).

The falling sphere technique is an important method for measuring densities and temperatures. Using more powerful rockets than used for rocketsondes the sphere, which is made of metallic-coated Mylar, is carried to heights in excess of 100 km and inflated to a diameter of about 1 m after release. Initially, the sphere falls at supersonic speeds, but atmospheric drag slows the fall until the fall rate goes subsonic at heights near 70 km. The sphere collapses at heights near 30 km. If the acceleration of the sphere is is measured by a high-precision tracking radar then atmospheric density can be determined with knowledge of the sphere's drag coefficient. Once the density profile is known then atmospheric temperature can be found from the equations of state and hydrostatics, in the same way as with a Rayleigh lidar (see equation 8). Tracking radar observations of the horizontal accelerations give the horizontal winds.

6 Satellite Techniques

6.1 UARS

Satellite measurements have played an increasingly important role in developing our understanding of middle atmosphere processes. In particular, instruments on the Upper Atmosphere Research Satellite (UARS) launched in September 1991 have been crucial for providing a global picture of prevailing winds, planetary-scale waves and tides and their effects on atmospheric composition. Nine instruments were deployed on UARS; details of the instruments and their capabilities can be found in the June 1993 edition of Journal of Geophysical Research. Of these instruments two were devoted specifically to direct measurements of the dynamics of the middle atmosphere and thermosphere.

The high resolution Doppler imager (HRDI) and wind imaging interferometer (WINDII) both measure Doppler shifts of airglow emission lines, although with rather different methods. Both are "limb viewing" instruments so that the weak omissions are observed against the cold background of space (Andrews et al., 1988). HRDI observes absorption and emission lines in the O_2 atmospheric band, scanning vertically to get a height profile,The instrument has a giving a height coverage from about 60 to 115 km. Vector winds are obtained by observing the same atmospheric volume from nearly two orthogonal directions. WINDII works in a different way to HRDI. By observing a number of airglow lines it has a height coverage from about 80 km to 300 km. Both techniques require the inversion of limb-scanned data sets so they have similar sampling conditions. Their height resolution is $\sim$2 km, but spatial averaging inherent in the limb-viewing geometry means that their horizontal resolution is a few hundred kilometers.

6.2 GPS Occultation Techniques

Monitoring of signals from the constellation of global positioning system (GPS) satellites is becoming a powerful tool for atmospheric sounding. GPS satellites, which orbit the earth twice a day at an altitude of $\sim$20,200 km, continuously transmit at frequencies of 1.2 GHz (L2) and 1.6 GHz (L1). As the signals propagate through the ionosphere and atmosphere they are refracted by vertical and horizontal variations in refractive index (equation 1). Ground-based observations of GPS signals can be used to measure both the total electron content (TEC) of the ionosphere and the

integrated water vapour content of the atmosphere above the receiver, which has applications in improving weather forecasting. Businger et al (1996) and Ware et al (1996) provide an overview of both Earth- and satellite-based GPS monitoring of the atmosphere.

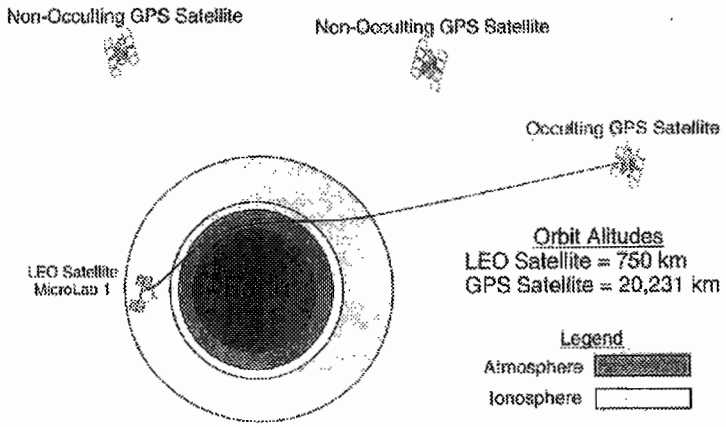


Figure 9: Schematic of a GPS/MET sounding of the Earth's atmosphere and ionosphere (after Ware et al., 1996)

The basic principles of spaced-based GPS techniques are illustrated in Figure 9, which shows a satellite in low Earth orbit (LEO) receiving signals from an occulting GPS satellite. The signals are progressively delayed as they are refracted through increasingly longer paths through the ionosphere and neutral atmosphere. After accounting for the relative motion of the LEO and GPS satellites and drift of the clocks in the GPS transmitters it is possible to extract the delays caused by the refractive bending and propagation velocity changes caused by the ionosphere and atmosphere. The path delays are then converted to refractivity profiles, which can be sampled at a high rate down to almost the surface

As (1) shows the ionospheric component is dispersive (the L1 and L2 signals travel at different velocities) so the ionospheric delays can be removed. Equation (1), together with the hydrostatic equation, is then used to convert the refractivity profiles to profiles of pressure and temperature. Rather like the falling sphere and Rayleigh lidar techniques an initial temperature guess from a suitable climatology is used to initialize the inversion process at high altitudes. Any errors due to this guess are small at heights below ~50 km. Below about 5 to 10 km the moist term in (1) dominates and moisture profiles can be derived if p and T are known. More details can be found in Rocken et al. (1997)

The vertical resolution is limited to about 1.5 km by effects such as diffraction and horizontal atmospheric inhomogeneities, while the long slant path between the LEO and GPS satellites means that the horizontal resolution is ~300 km. The relatively small estimated temperature errors of 1°C in the 5 to 40 km height range and global coverage makes GPS measurements from space a very suitable technique for producing gravity wave climatologies, a potential that is now being realized (Tsuda et al., 2000).

7 Airglow Imager Observations

Weak optical emissions have been studied for many years to study the state of the upper atmosphere. It is not the purpose here to discuss the processes by which airglow is produced, but rather to note that they provide a means by which waves and their effects can be directly visualized.

Figure 10: An image of OH airglow (843.0 nm) obtained on February 13, 2000 at the Buckland Park field station near Adelaide. The field of view is 50×75° (equivalent to ~80×135 km), with north at the top and west to the right. The image exposure was for 1 minute and the brightest OH emission appears white (courtesy of J. H. Hecht).

Airglow imagers are essentially cameras that focus light onto a CCD array so that the output can be stored electronically. Narrow-band filters isolate particular emissions, so sequential changing of filters enables airglow from different height regions to be studied. For example, OH Meinal, O_2 atmospheric bands and 557.7 nm

airglow emissions come from heights of about 87, 93 and 97 km. Comparing intensities from different lines in a give wavelength band also makes it possible to measure temperature.

An example of fluctuations in OH airglow intensity is given in Figure 10. The image, made from a 1 minute exposure, shows features with horizontal scales of about 40 km. Movies from sequential exposures allows the direction and speed of travel to be estimated, and if independent wind measurements from a radar, say, are available then the intrinsic wave properties can be found. These observations were made with an imager that had a relatively narrow field of view, but some instruments provide images over almost the complete sky. Instruments with wide and narrow fields of view have their own advantages and disadvantages.

8 References

Allen, S. J., 1996, *Gravity wave motions in the troposphere and lower stratosphere*, Ph. D, thesis, University of Adelaide, pp 253.

Andrews, D.G., J.R. Holton, and C.B. Leovy, 1987, *Middle Atmosphere Dynamics*, 489 pp., Academic, San Diego, Calif.

Briggs, B.H., 1980, Radar observations of atmospheric winds and turbulence: A comparison of techniques, *J. Atmos. Terr. Phys.*, **42**, 823-833.

Businger, S., and coauthors, 1996, The promise of GPS in atmospheric monitoring, *Bull. Amer. Meteor. Soc.*, **77**, 5-18.

Gardner, C. S., 1991, Instroduction to ALOHA-90: The airborne lidar and observations of the Hawaiian airglow, *Geophys. Res. Lett.*, **18**, 1313-1316.

Dao, P.D., R. Farley, X. Tao, and C.S. Gardner, Lidar observations of the temperature profile between 25 and 103 km: Evidence of strong tidal perturbation, 1995, *Geophys. Res. Lett.*, **22**, 2825-2828.

Eckermann, S.D., I. Hirota, and W.K. Hocking, 1995, Gravity wave and equatorial wave morphology of the stratosphere derived from long-term rocket soundings, *Q. J. R. Meteorol. Soc.*, **121**, 149-186.

Gage, K.S., and B.B. Balsley, 1980, On the scattering and reflection mechanisms contributing to clear air radar echoes from the troposphere, stratosphere and mesosphere, *Radio Sci.*, **15**, 243-257.

Hocking, W.K., 1985, Measurements of turbulent energy dissipation rates in the middle atmosphere by radar techniques, *Radio Sci.*, **20**, 1403-1422.

Kent, G.S., and R.W.H. Wright, 1970, A review of laser radar measurements of atmospheric properties, *J. Atmos. Terr. Phys.*, **32**, 917-943.

Rocken, C., and coauthors, 1997, Analysis and validation of GPS/MET data in the neutral atmosphere, *J. Geophys. Res.*, **102**, 29,849-29,866.

Röttger, J., 1984, The MST radar technique, in *Handbook for Middle Atmosphere Program*, (ed. by R.A. Vincent), 187-232, SCOSTEP Secretariat, University of Illinois.

She, C.Y., H. Latifi, J.R. Yu, R.J. Alverez, R.E. Bills, and C.S. Gardner, 1990, Two-frequency lidar technique for mesospheric Na temperature measurements, *Geophys. Res. Lett.*, **17**, 929-932.

Tsuda, T., Data acquisition and processing, 1990, in *Handbook for the Middle Atmosphere Program*, (ed. S. Fukao), 151-189, SCOSTEP Secretariat, University of Illinois.

Tsuda, T., Y. Murayama, H. Wiryosumarto, S. W. B. Harijono, and S. Kato, 1994, Radiosonde observations of equatorial atmosphere dynamics over Indonesia, 2, Characteristics of gravity waves, *J. Geophys. Res.*, **99(D5)**, 10,507-10,516.

Tsuda, T., M. Nishida, C. Rocken, and R. Ware, 2000, Global Morphology of Gravity Wave Activity in the Stratosphere Revealed by the GPS Occultation Data (GPS/MET), *J. Geophys. Res.*, **105**, 7257-7274.

Vincent, R.A., 1984, MF/HF radar measurements of the dynamics of the mesosphere region - A review, *J. Atmos. Terr. Phys.*, **46**, 961-974, 1984.

Vincent, R. A., and M. J. Alexander, 2000, Gravity waves in the tropical lower stratosphere: An observational study of seasonal and interannual variability, *J. Geophys. Res.*, **105(D14)**, 17,971-17,982.

Ware, R., and coauthors, 1996, GPS sounding of the atmosphere from low Earth orbit: Preliminary results, *Bull. Amer. Meteor. Soc.*, **77**, 19-40.

Woodman, R. F., and A. Guillen, 1974, Radar observations of winds and turbulence in the stratosphere and mesosphere, *J. Atmos. Sci.*, **31**, 493-505.

Chapter 3

Equatorial Waves

Rolando Garcia,

NCAR, Boulder, U.S.A.

Waves in the Tropics

Rolando R. Garcia
National Center for Atmospheric Research
Boulder, Colorado, U.S.A.

1 Introduction

Dynamical processes in the tropical atmosphere differ fundamentally from those that dominate in extratropical regions because the Coriolis parameter, $f = \sin\theta$, becomes small in the vicinity of the Equator ($\theta \to 0$).

The difference is exemplified by the behavior of the zonal-mean zonal wind, $\bar{u}$, in the Tropics and extratropical regions. The top panel of Fig. 1 shows the evolution of $\bar{u}$ at 60 N over a period of five years. The temporal variation is mainly annual because, to a first approximation, $\bar{u}$ is in thermal wind balance with the temperature field,

$$f\frac{\partial \bar{u}}{\partial z} = -\frac{R}{H}\frac{\partial \bar{T}}{\partial y} \tag{1}$$

and the seasonal cycle of temperature is driven by the annual variation in radiative heating.

On the other hand, the behavior in the Tropics can be rather different, as shown in the bottom panel of Fig. 1. The temporal behavior of $\bar{u}$ in the tropical stratosphere is not annual; instead, it displays quasi-regular variability with a period of about 27 months, which is known as the quasi-biennial oscillation (QBO; see, e.g., Baldwin et al., 2001).

These differences can be understood by considering the terms that contribute to the zonal momentum balance at low frequencies. These are summarized below:

Zonal-Mean Equations

$$\begin{array}{rcll} \bar{u}_t + \bar{v}^*\bar{u}_y + \bar{w}^*\bar{u}_z - f\bar{v}^* & = & \mathcal{F} & \text{zonal momentum} \\ \bar{v}^*_y + \rho^{-1}(\rho\bar{w}^*)_z & = & 0 & \text{continuity} \\ f\bar{u} & = & -\bar{\phi}_y & \text{meridional momentum} \\ \bar{\phi}_{zt} + \bar{v}^*\bar{\phi}_{zy} + \bar{w}^*S & = & -\alpha\bar{\phi}_z & \text{thermodynamics} \end{array} \tag{2}$$

Quasi-steady motions in Extratropical Regions

- momentum balance $-f\bar{v}^* \simeq \mathcal{F}$
- this implies $(\bar{v}^*, \bar{w}^*)$ are determined by $\mathcal{F}$
- in turn this determines $\bar{\phi}$ via $\bar{w}^*S = -\alpha\bar{\phi}_z$
- and $\bar{u}$ via $f\bar{u} = -\bar{\phi}_y$

Low-frequency motions in the Tropics

- $f = 2\Omega\sin\theta \rightarrow \left(\frac{2\Omega}{a}\right)\cdot a\theta \equiv \beta y$
- f does not dominate, so all terms in the zonal momentum equation are potentially important:

$$u_t + v^*u_y + w^*u_z - \beta y v^* = \mathcal{F} \tag{3}$$

- $\mathcal{F}$ can give rise to long-period oscillations in $\bar{u}$ (SAO, QBO) instead of a quasi-steady-state
- $\bar{u}$ remains in geostrophic balance $\beta y\bar{u} = -\bar{\phi}_y$

The term $\mathcal{F}$ on the RHS of the momentum equation represents any zonally-averaged force acting along the zonal direction. For present purposes, we are concerned mainly with the force due to the dissipation of a variety of waves that propagate vertically from the lower to the upper atmosphere. In this case, $\mathcal{F}$ represents the divergence of the Eliassen-Palm (EP) flux:

$$\mathcal{F} = \frac{1}{\rho\cos\theta}\nabla\cdot\mathbf{F} \tag{4}$$

where the EP flux vector is given by:

$$\mathbf{F} = \rho_o\cos\phi\left[\bar{u}_z\frac{\overline{v'\theta'}}{\bar{\theta}_z} - \overline{w'v'}, \left(f - \frac{1}{\cos\phi}\frac{\partial\bar{u}\cos\phi}{\partial y}\right)\frac{\overline{v'\theta'}}{\bar{\theta}_z} - \overline{w'w'}\right] \tag{5}$$

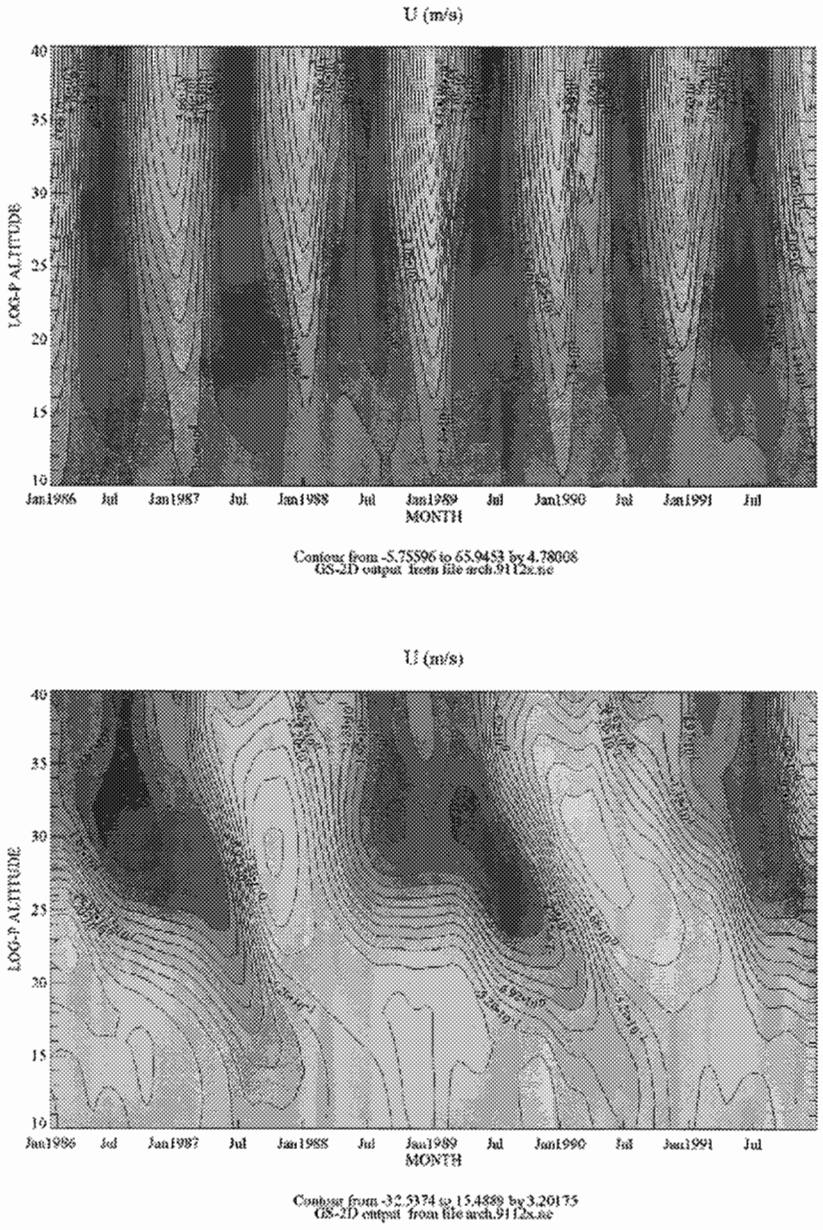
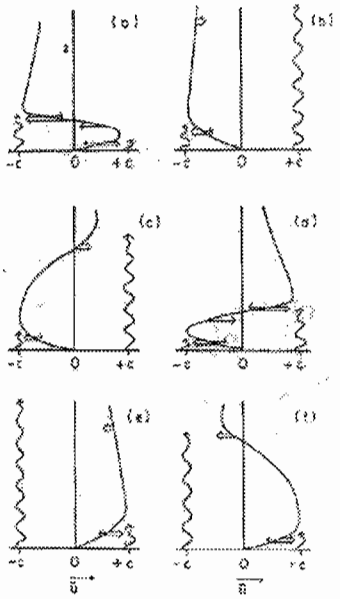


Figure 1: Altitude-time plot of the seasonal variation of the mean zonal wind at 60 N (top) and at the Equator (bottom) over a period of 5 years. Darker shades of gray correspond to negative values.

Fig. 2 shows schematically how the divergence of $\mathbf{F}$ due to dissipating waves can produce alternating regimes of easterly and westerly mean zonal winds in the Tropics. The spectrum of waves emanating from the lower atmosphere contains both easterly ($c < 0$) and westerly ($c > 0$) components. These are absorbed and deposit their momentum near the altitudes where the zonal wind approaches the phase speed of the wave, i.e,, where $(\bar{u} - c) \to 0$. Waves whose phase velocity is of opposite sign propagate freely through such regions, but deposit their momentum at higher altitudes. In panel (a) of the figure this mechanism has produced an easterly shear layer near the bottom of the domain. In panel (b), this layer descends until westerly winds disappear altogether, and waves of $c > 0$ propagate freely to higher altitudes, where they begin to produce a westerly shear layer, as shown in panel (c). This westerly layer descends as shown in panel (d) until only westerly winds remain and the cycle begins anew with waves of $c < 0$ propagating to higher altitudes (panels (e) and (f)).



Schematic representation of the evolution of the mean flow in Plumb's analog of the QBO. Six stages of a complete cycle are shown. Double arrows show wave-driven accelerations and single arrows show viscously driven accelerations. Wavy lines indicate relative penetration of easterly and westerly waves. See text for details. [After Plumb (1984).]

Figure 2: (from Andrews et al, 1987)

2 Waves on the Equatorial β-plane

Given the importance of vertically-propagating waves on the circulation of the tropical atmosphere, it is useful to acquire a basic understanding of the dynamics of these waves. At small scales, on the order of tens to a few hundred km, wave motions are nearly of the pure gravity wave type, where the effects of the Earth's rotation are relatively unimportant. These "meso-scale" gravity waves wll be discussed in due course, but first it is appropriate to consider waves of horizontal scale of order 1000 km and larger, for which the effects of rotation are important. Because of the smallness of the Coriolis parameter near the Equator these waves have characteristics that set them appart from large-scale motions outside the Tropics, as summarized below:

- **Extratropical large-scale waves are quasi-geostrophic and quasi-non-divergent**
  - their restoring force is due to the variation of f with latitude
  - gravity (which acts through the divergence field) is not of first order importance
  - the motion is described by the quasi-geostrophic vorticity equation
  - the waves described by this equation are extratropical Rossby waves
- **Tropical large-scale waves respond both to variations in f and to gravity (buoyancy) forces**
  - divergence is important for tropical waves
  - tropical waves are equatorially trapped, the trapping being most effective at low frequencies

The last point is important, as it implies that wave activity associated with tropical waves will be confined to the Tropics and, insofar as the waves are vertically-propagating, will be able to affect the circulation of the tropical middle and upper atmosphere.

The basic properties of tropical waves can be described under an approximation to the primitive equations known as the **equatorial beta-plane**. For latitudes θ close to the equator, the Coriolis parameter is approximated as:

$$f = 2\Omega a \sin\theta \simeq \left(\frac{2\Omega}{a}\right) a\theta \equiv \beta y \tag{6}$$

which yields the following linearized system of **equatorial beta-plane equations:**

$$\begin{aligned} \frac{\partial u'}{\partial t} - \beta y v' &= -\phi'_x \\ \frac{\partial v'}{\partial t} + \beta y u' &= -\phi'_y \\ \frac{\partial u'}{\partial x} + \frac{1}{\rho}\frac{\partial(\rho w')}{\partial z} &= 0 \\ \frac{\partial \phi'_z}{\partial t} + w' N^2 &= 0 \end{aligned} \tag{7}$$

Assuming a solution of the form,

$$(u', v', w', \phi') = (\hat{u}, \hat{v}, \hat{w}, \hat{\phi}) \cdot \exp(z/2H) \cdot \exp i(kx + mz - \omega t) \tag{8}$$

then gives the Equatorial β-plane system:

$$\begin{aligned} -i\omega\hat{u} - \beta y\hat{v} &= -ik\hat{\phi} \\ -i\omega\hat{v} + \beta y\hat{u} &= -\hat{\phi}_y \\ ik\hat{u} + \hat{v}_y + im\hat{w} &= 0 \\ -i\omega\hat{\phi}'_z + \hat{w}N^2 &= 0 \end{aligned} \tag{9}$$

where it has also been assumed that $m^2 >> (1/4H)^2$ (the Boussinesq approximation).

2.1 The Equatorial Kelvin Wave

A particularly interesting solution of the set (9) is obtained when the meridional velocity, v', is zero. The equations then become:

$$\begin{aligned} -i\omega\hat{u} &= -ik\hat{\phi} \\ \beta y\hat{u} &= -\hat{\phi}_y \\ ik\hat{u} + im\hat{w} &= 0 \\ -i\omega\hat{\phi}'_z + \hat{w}N^2 &= 0 \end{aligned} \tag{10}$$

Note that the geopotential gradient in the zonal direction, $\hat{\phi}_x = ik\hat{\phi}$, is balanced by a zonal acceleration (as in a pure gravity wave) while the gradient in the meridional direction, $\hat{\phi}_y$, is balanced by the Coriolis force (as in Rossby waves).

These equations have the simple solution

$$\hat{\phi} \propto \exp\left[-\frac{y^2}{2L^2}\right] \tag{11}$$

with $L^2 = \omega/\beta k$, which follows upon elimination of $\hat{u}, \hat{w}$. The complete solution is:

$$\phi' \propto \exp\left[-\frac{y^2}{2L^2}\right] \cdot \exp(z/2H) \cdot \exp i(kx + mz - \omega t) \tag{12}$$

Combining (10) and (12) leads to the dispersion relation:

$$\omega^2 = \frac{k^2 N^2}{m^2} \tag{13}$$

or

$$\omega = -kN/m, \qquad m < 0 \tag{14}$$

where $m < 0$ so that $\omega > 0$ and $L^2 = \omega/\beta k > 0$, and the negative root of (13) is chosen to insure that the vertical group velocity:

$$c_{gz} = \frac{\partial \omega}{\partial m} = \frac{kN}{m^2} \tag{15}$$

is upward. Note that L^2 can be then written as $L^2 = (N/\beta|m|)$. Other wave fields follow directly from the β-plane equations:

$$\hat{u} \propto \left(\frac{ik}{\omega}\right)\hat{\phi} \tag{16}$$

$$\hat{w} \propto -\left(\frac{\omega}{N^2}\right)\left(m - \frac{i}{2H}\right)\hat{\phi} \tag{17}$$

An example of the structure of the Kelvin wave is shown in Fig. 3.

2.2 Other Equatorial Waves

Other wave soulutions can be obtained from the beta-plane equations by droppping the requirement that $v' = 0$. Assuming solutions of the form:

$$(u', v', w', \phi') = (\hat{u}, \hat{v}, \hat{w}, \hat{\phi}) \cdot \exp(z/2H) \cdot \exp i(kx + mz - \omega t) \tag{18}$$

the **general β-plane equations** can be combined to give:

$$\left[\frac{d^2}{dy^2} + \left(\frac{m^2\omega^2}{N^2} - k^2 - \frac{\beta k}{\omega}\right) - \frac{\beta^2 m^2}{N^2} y^2\right] \hat{v} = 0 \tag{19}$$

This is formally identical to the "quantum harmonic oscillator" equation, which has solutions in terms of Hermite polynomials:

$$\hat{v} = \hat{v}_o \exp\left(-\frac{y^2}{2L^2}\right) H_n(y/L) \qquad L^2 = \frac{N}{\beta|m|} \tag{20}$$

where H_n are the Hermite polynomials, whose index n must satisfy:

$$\frac{m^2\omega^2}{N^2} - k^2 - \frac{\beta k}{\omega} = (2n+1)\frac{\beta|m|}{N} \tag{21}$$

which is also the dispersion relation for the waves.

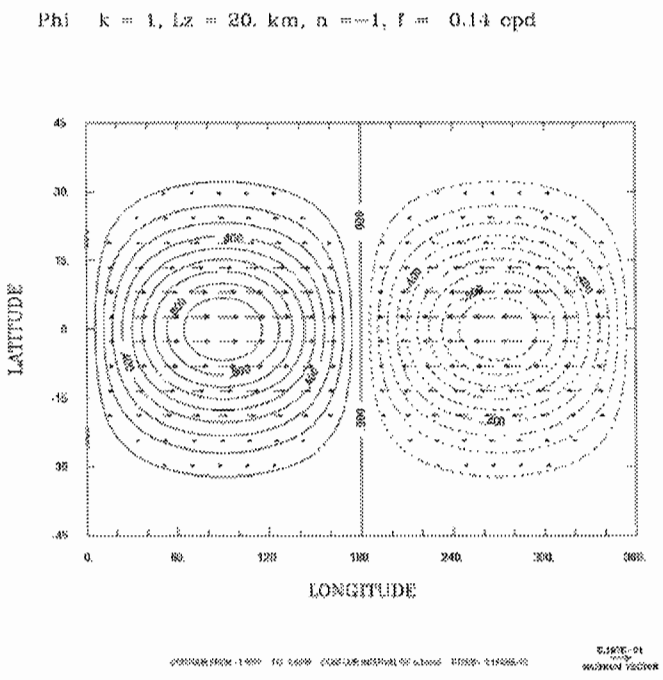


Figure 3: Horizontal (latitude-longitude) structure of the zonal wavenumber 1 Kelvin wave according to the solution of the beta-plane equations with $v' = 0$. The vertical wavenumber has been assumed to be 20 km, which yields a frequency of 0.14 cpd, according to the dispersion relation (14). The velocity and geopotential fields are in phase, as per Eqs. (12) and (16).

Using the recursion relations for Hermite polynomials:

$$H_{n+1}(y/L) = 2(y/L)H_n(y/L) - 2nH_{n-1}(y/L) \tag{22}$$

$$\frac{d}{dy}H_n(y/L) = (2n/L)H_{n-1}(y/L) \tag{23}$$

solutions for other fields follow from the β-plane equations:

$$\hat{u} = i\hat{v}_o(\beta m N)^{1/2}\left[\frac{0.5H_{n+1}}{|m|\omega - Nk} + \frac{nH_{n-1}}{|m|\omega + Nk}\right]\cdot\exp\left(-\frac{y^2}{2L^2}\right) \tag{24}$$

$$\hat{\phi} = i\hat{v}_o\left(\frac{\beta N^3}{m}\right)^{1/2}\left[\frac{0.5H_{n+1}}{|m|\omega - Nk} - \frac{nH_{n-1}}{|m|\omega + Nk}\right]\cdot\exp\left(-\frac{y^2}{2L^2}\right) \tag{25}$$

The solutions can be categorized under several types, according to their dispersion relationships:

- eastward-propagating and westward-propagating gravity waves
- westward-propagating equatorial Rossby waves
- Rossby-gravity waves (these resemble Rossby waves for $\omega < 0$ and Kelvin waves for $\omega > 0$)

Detailed discussions of these solution types can be found in Gill (1982), Andrews et al (1987), and Matsuno (1966). Here we limit ourselves to summarizing the salient properties of these equatorial waves.

2.2.1 Waves with $n > 0$

These waves obey the general dispersion relation:

$$\frac{m^2\omega^2}{N^2} - k^2 - \frac{\beta k}{\omega} = (2n+1)\frac{\beta|m|}{N} \tag{26}$$

which has two solution branches (Gill, 1982), an **upper (high-frequency) branch** for which

$$\omega \cong -\frac{N}{m}\left[k^2 + (2n+1)\frac{\beta|m|}{N}\right]^{1/2} \tag{27}$$

which corresponds to eastward-propagating inertia-gravity waves ($\omega > 0$) when $m < 0$, and to westward-propagating waves ($\omega < 0$) when $m > 0$, and **a lower (low-frequency) branch** for which

$$\omega = \frac{-\beta k}{\left[k^2 + (2n+1)\frac{\beta|m|}{N}\right]} \tag{28}$$

which corresponds to equatorial Rossby waves.

Note that at large zonal wavenumber k (small horizontal scale), the frequency approaches the following limits:

$$\omega \to \pm\frac{kN}{|m|} \tag{29}$$

for the inertia-gravity waves, and

$$\omega \to -\beta/k \tag{30}$$

for the Rossby waves.

2.2.2 Waves with $n = 0$

In this case, dispersion relation is:

$$\frac{|m|\omega}{N} - k = \frac{\beta N}{\omega} \tag{31}$$

which has solutions

$$\omega = \frac{kN}{2|m|} \cdot \left[1 \pm \left(1 + \frac{4\beta|m|}{k^2 N}\right)^{1/2}\right] \tag{32}$$

These waves are known as **Rossby-gravity waves** because they behave like inertia-gravity waves for $\omega > 0$ (the positive root in (32)), and as Rossby waves for $\omega < 0$ (the negative root).

The large-k limit for the eastward $n = 0$ wave is:

$$\omega \to \frac{kN}{|m|} \tag{33}$$

while, for the eastward $n = 0$ wave,

$$\omega \to -\beta/k \tag{34}$$

These are the same limits as for the $n > 0$ eastward-propagating inertia-gravity wave and the $n > 0$ Rossby wave, respectively. Thus, at large k, eastward-propagating waves of $n = 0$ behave exactly like inertia-gravity waves, while westward-propagating waves behave like Rossby waves. Fig. 4 shows plots of the dispersion relations for various types of equatorial wave, assuming a buoyancy frequency $N = 0.02$ sec^{-1} and a vertical wavelength of 20 km.

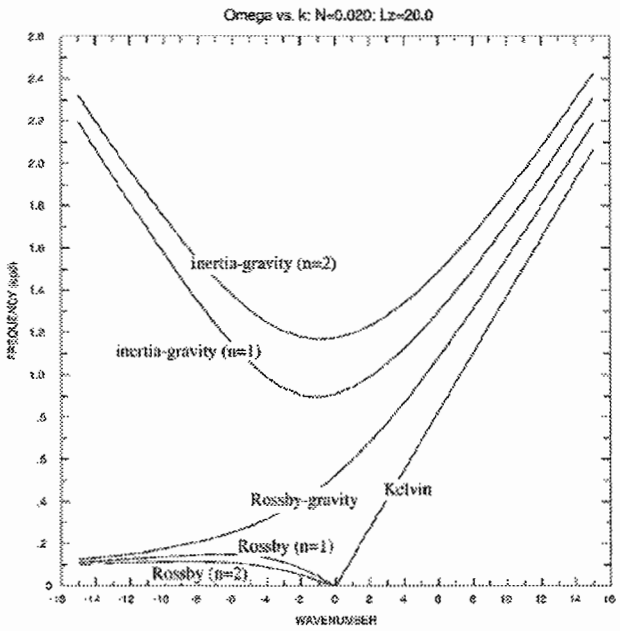


Figure 4: Dispersion curves for equatorial waves: Kelvin wave, Rossby-gravity wave, inertia-gravity waves, and Rossby waves. For the latter two types, the first two modes, $n = 1, 2$ in Eqs. (27) and (28), are shown.

Fig. 5 shows examples of the horizontal structure of the eastward-propagating and westward-propagating Rossby-gravity wave ($n = 0$). Figs. 6 and 7 show the structures of the $n = 1$ (eastward-propagating) inertia-gravity and (westward-propagating) Rossby waves, respectively.

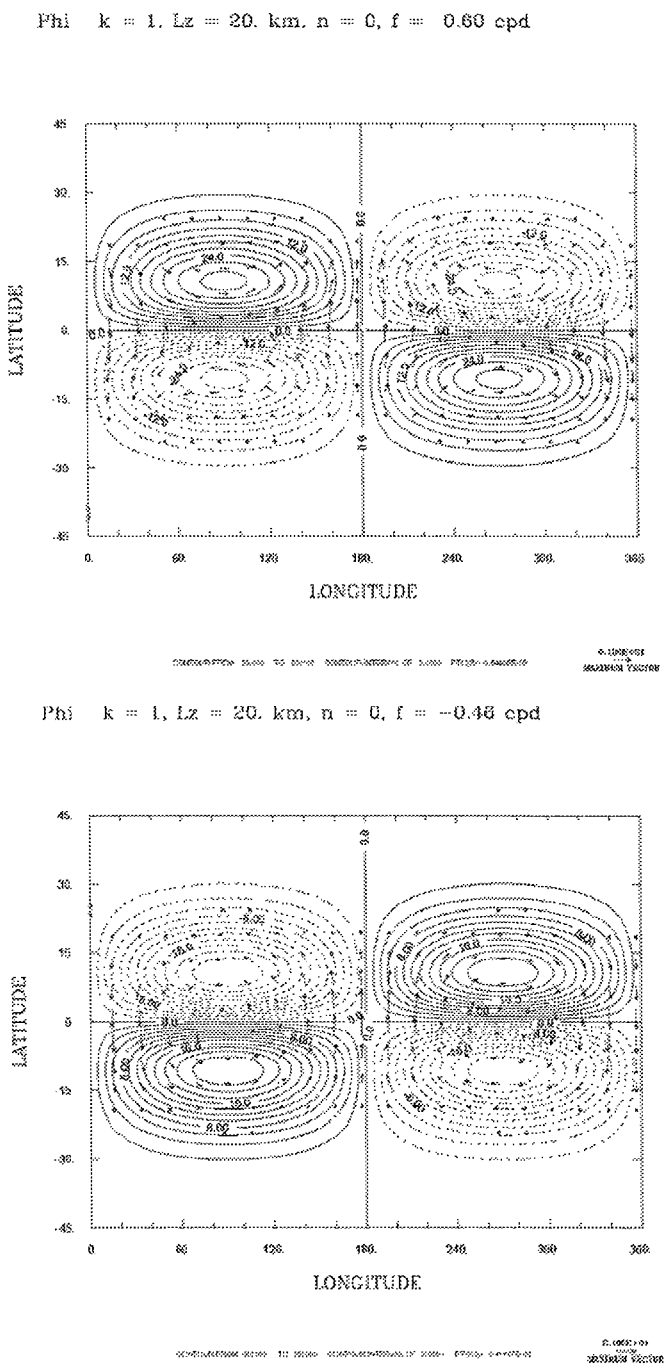


Figure 5: Horizontal structure of eastward-propagating (top) and westward-propagating (bottom) Rossby-gravity waves of $k = 1$ and $\lambda_z = 20$ km, with $N = 0.02$ sec^{-1}.

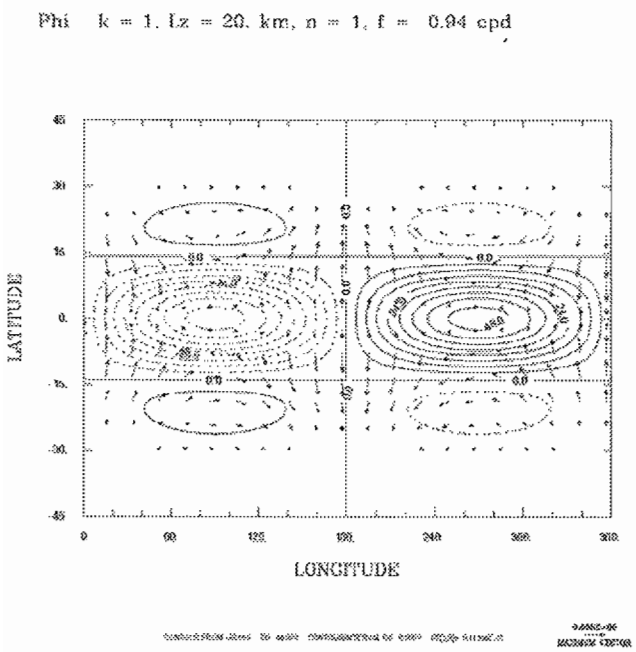


Figure 6: Structure of the eastward-propagating inertia-gravity wave ($n = 1$).

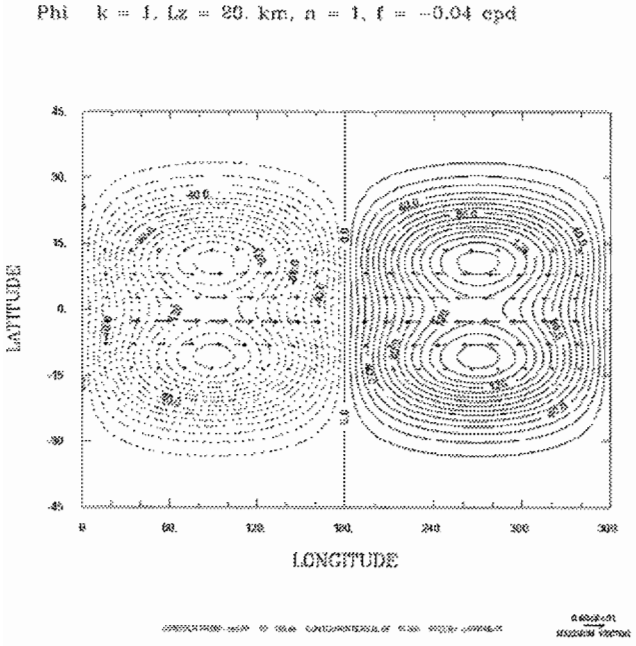


Figure 7: Structure of the westward-propagating Rossby wave ($n = 1$).

A few additional points are worth noting:

- Kelvin waves behave like eastward-propagating gravity waves
- Previous results can be generalized to case with a uniform background wind $\bar{u}$; then $c \to \hat{c} = c - \bar{u}$
- Thus, for gravity waves,

$$\hat{\omega} = k\hat{c} = k(c - \bar{u}) = \pm\frac{kN}{|m|} \tag{35}$$

- It follows that

$$m = \pm\frac{N}{k|\hat{c}|} \tag{36}$$

and, therefore,

$$m \to \pm\infty \qquad \text{when } |\hat{c}| \to 0 \tag{37}$$

The result (36)-(37) is of interest, since it shows that the vertical wavelength ($\lambda_z = 2\pi m^{-1}$) vanishes when $(c - \bar{u}(z)) \to 0$. The altitude where this happens is known as a **critical level** for the wave in question. Waves are dissipated strongly at their critical levels, as we have seen earlier in connection with Fig. 2.

3 Small-scale gravity waves

When the spatial and temporal scales of wave motions are small enough, the Coriolis force does not play a major role in determining the characteristics of the waves. For such waves, the main restoring force is gravity, acting through the convergence/divergence field. In order to describe these waves, it is possible to use Eq. (19), this time ignoring the terms related to the Earth's rotation:

$$\left[\frac{d^2}{dy^2} + \left(\frac{m^2\omega^2}{N^2} - k^2\right)\right]\hat{v} = 0 \tag{38}$$

Since the coefficient is now constant, we can assume that

$$\hat{v} \propto \exp(ily) \tag{39}$$

which yields the dispersion relation

$$\omega = \pm\frac{N}{|m|}\left(k^2 + l^2\right)^{1/2} \tag{40}$$

This has the same form as the dispersion relation for inertia-gravity waves (27) discussed earlier:

$$\omega = \pm \frac{N}{|m|} \left[k^2 + (2n+1) \frac{\beta m}{N} \right]^{1/2} \tag{41}$$

where $(2n+1)\,\beta m/N$ plays the role of a meridional wavenumber

Note that the two roots of the dispersion relation (41) are identical in the large-k limit:

$$\omega = \pm \frac{Nk}{|m|} \tag{42}$$

Small-scale gravity waves are ubiquitous in the Earth's atmosphere. They are known to play a central role in the dynamics of the extratropical mesosphere; they may also play a leading role in forcing the zonal mean wind oscillations (QBO and SAO) that characterize the tropical middle atmosphere.

4 Phase and Group Velocity

Further insight into the way that waves propagate in the tropical atmosphere can be gained from the concept of **group velocity**. The group velocity characterizes the rate at which wave energy can propagate, and it is generally different from the more familiar concept of **phase velocity**.

The phase velocity of a wave is given cimply by $c = \omega/k$; this is the speed at which crests and throughs of individual frequency components move. In general, $c = c(k)$, a function of wavenumber. An initial wave disturbance of arbitrary shape, or **wave packet**, can be expressed as a superposition of frequency and wavenumber components:

$$\psi(x,t) = \Sigma A_{k,\omega} \exp i(kx - \omega t) \tag{43}$$

The subsequent development of the wave packet does not, in general, preserve the initial shape since different components propagate at different rates. Therefore, c is not the velocity at which wave energy propagates; i.e., it is not the velocity of the wave packet.

Consider the following simple example:

$$\psi(x,t) = Re\left\{ e^{i[(k+\Delta k)x - (\omega+\Delta\omega)t]} + e^{i[(k-\Delta k)x - (\omega-\Delta\omega)t]} \right\} \tag{44}$$

that is, a combination of two waves of slightly different wavenumber,k, and frequency, ω. This can be written as:

$$\psi(x,t) = Re\left\{e^{i(kx-\omega t)} \cdot \left[e^{i(\Delta kx-\Delta\omega t)} + e^{-i(\Delta kx-\Delta\omega t)}\right]\right\} \tag{45}$$

which simplifies to:

$$\psi(x,t) = 2\cos(kx-\omega t)\cos(\Delta kx-\Delta\omega t) \tag{46}$$

or,

$$\psi(x,t) = 2\cos(kx-\omega t)\cos(\Delta k(x-c_g t)) \tag{47}$$

where:

$$c_g = \frac{\Delta\omega}{\Delta k} \rightarrow \frac{\partial\omega}{\partial k} \tag{48}$$

is the group velocity.

Waves for which $c \neq c(k)$ are called nondispersive. In this case $\omega = kc$, $c_g = \partial\omega/\partial k = c$, and the wave packet retains its initial shape. Otherwise, the envelopes of waves with different k travel at different rates, and the waves disperse as they propagate away from the source. The group velocity result can be generalized to more than one dimension. Thus, in general,

$$\mathbf{c}_g = \frac{\partial\omega}{\partial\mathbf{k}} \tag{49}$$

The difference between phase velocity and group velocity can be appreciated in Fig. 8.

4.1 Horizontal Group Velocity: Trapping

The solution to the equation

$$\left[\frac{d^2}{dy^2} + \left(\frac{m^2\omega^2}{N^2} - k^2 - \frac{k\beta}{\omega}\right) - \frac{\beta^2 m^2}{N^2}y^2\right]\hat{v} = 0 \tag{50}$$

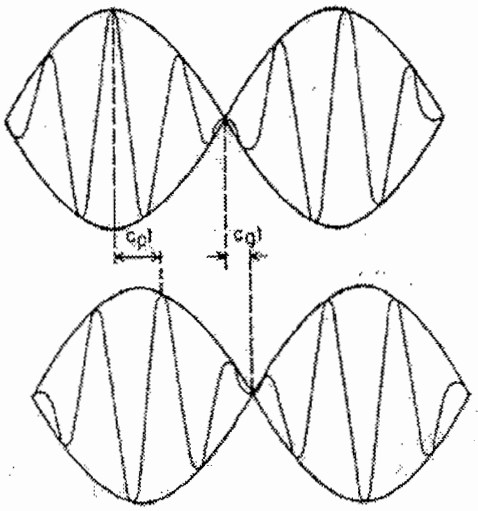
can be written as follows in the WKB approximation:

$$\hat{v} = \frac{1}{\sqrt{l}}\exp i\left[kx + \int l(y')dy' + mz - \omega t\right] \tag{51}$$

where

$$\begin{aligned} l^2(y) &= \frac{m^2\omega^2}{N^2} - k^2 - \frac{k\beta}{\omega} - \frac{\beta^2 m^2}{N^2}y^2 \\ &= \frac{\beta^2 m^2}{N^2}(y_c^2 - y^2) \end{aligned}$$

Dispersion



Superposition of two sinusoidal waves, illustrating the difference between the phase speed (the speed of the wave crests) and the group speed (the speed of the envelope, or wave packet).

Figure 8: Group and phase velocity. In the time interval t, the wave crests are displaced a distance $c_p t$, while the disturbance envelope travels a distance $c_g t$. From Gill (1982).

and

$$y_c^2 = \frac{\omega^2}{\beta^2} - \frac{k^2 N^2}{\beta^2 m^2} - \frac{k N^2}{\beta m^2 \omega} \tag{52}$$

The solution is wavelike ($l^2 > 0$) as long as y is less than the "critical latitude" y_c. This leads to wave trapping in $-y_c < y < y_c$.

The horizontal components of group velocity are:

$$c_{gx} = \frac{\partial \omega}{\partial k} = \frac{2k + \beta\omega}{2\omega m^2/N^2 + \beta k/\omega^2} \tag{53}$$

$$c_{gy} = \frac{\partial \omega}{\partial l} = \frac{2l}{2\omega m^2/N^2 + \beta k/\omega^2} \tag{54}$$

Then, the ray path in the x-y plane is defined by:

$$\frac{dy}{dx} = \frac{c_{gy}}{c_{gx}} = \frac{l}{k + \beta/2\omega} \tag{55}$$

$$\frac{d(y/y_c)}{1 - (y/y_c)^2} = \frac{\beta m/N}{k + \beta/2\omega} \cdot dx \tag{56}$$

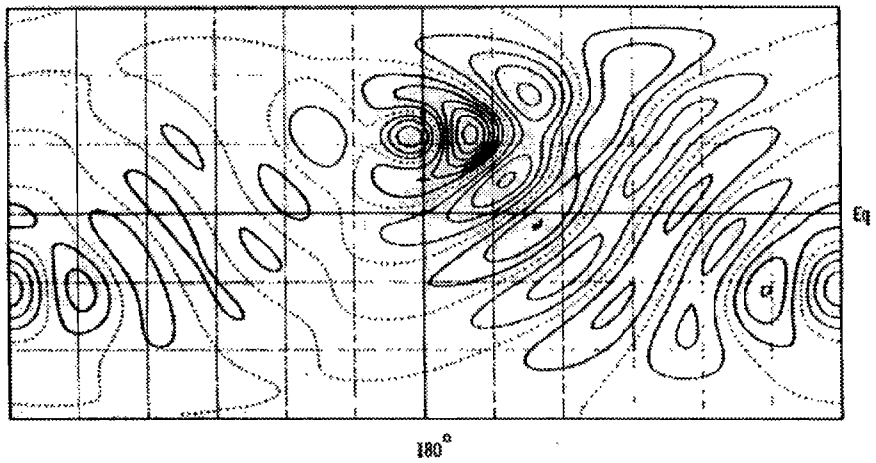
or

$$d\left[\sin^{-1}(y/y_c)\right] = \frac{\beta m/N}{k+\beta/2\omega} \cdot dx \tag{57}$$

Upon integration, this gives:

$$y = y_c \sin\left(\frac{\beta m/N}{k+\beta/2\omega} \cdot x\right) \tag{58}$$

See Gill (1982) for a detailed discussion. The meridional trapping of tropical waves is illustrated in Fig. 9.



Planetary wave propagation on a sphere, as found in a numerical experiment of Grose and Hoskins (1979). Contours are of perturbation vorticity, and disturbances to a superrotation zonal flow (i.e., an eastward flow with uniform angular velocity about the earth's axis) are produced by a circular mountain centered at 30°N and 180° longitude, and with radius equal to 22.5° of latitude. Waves travel backward and forward across the equator along ray paths that are curved because of variation in the Coriolis parameter f with latitude. The equatorial trapping effect is evident. The amplitude of the wave decays with distance because of dissipative effects included in the model. [From Grose and Hoskins (1979, Fig. 3a).]

Figure 9: Trapping of equatorial waves. From Gill (1982).

4.2 Vertical Group Velocity

Waves in the Tropics are excited by a variety of sources, including:

- Direct aborption of solar radiation by atmospheric gases (ozone in the stratosphere, water vapor in the troposphere). Most important for the excitation of the thermal tides

- Latent heat release by convection. This excites a broad spectrum of waves that are believed to be responsible for the driving of the QBO and SAO

From the point of view of middle atmosphere dynamics, the most important waves are those that can propagate readily in the vertical. The effectiveness of vertical propagation depends on the waves' vertical group velocity,

$$c_{gz} = \frac{\partial \omega}{\partial m} \tag{59}$$

Waves with large c_{gz} are less subject to dissipation as they propagate vertically, so they can exert an influence at high altitudes. Based upon these considerations, Kelvin and gravity waves should be of greatest importance for middle atmosphere dynamics. For these waves,

$$c_{gz} \simeq kN/m^2 = \omega^2/kN \tag{60}$$

Figs. 10-13 show how group velocity varies with horizontal wavenumber and vertical wavelength for the various equatorial waves discussed above.

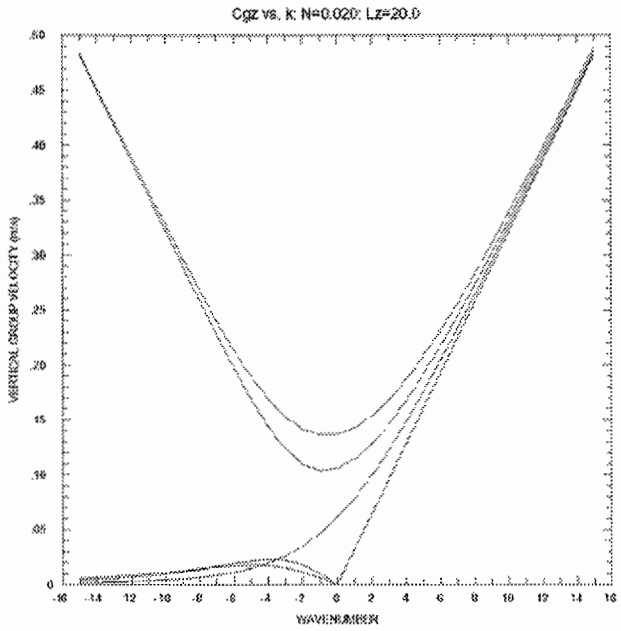


Figure 10: Vertical group velocities for equatorial waves of λ_z =20 km.

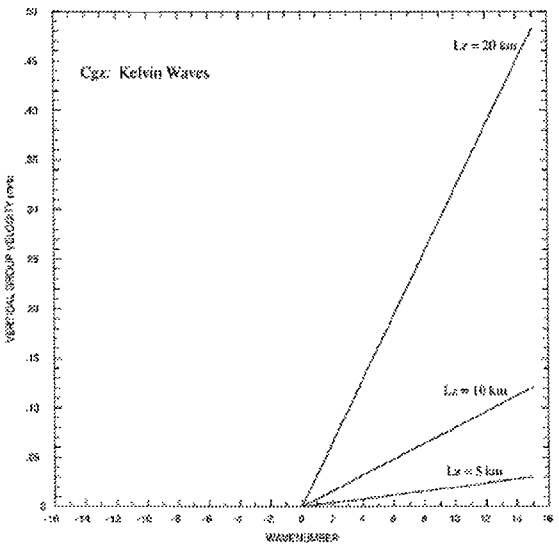


Figure 11: Vertical group velocities for Kelvin waves of various λ_z.

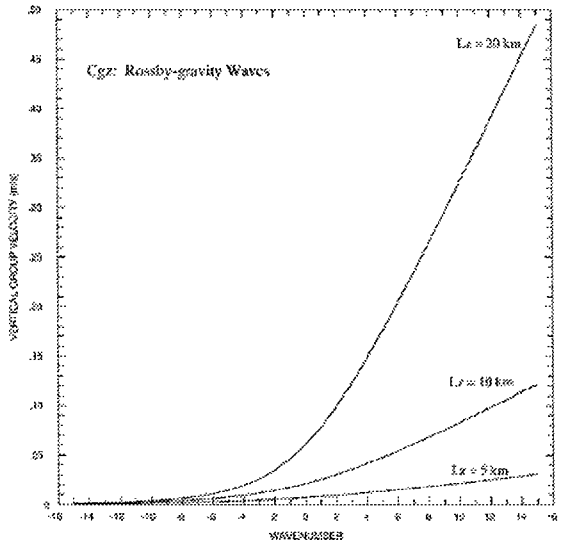


Figure 12: Vertical group velocities for Rossby-gravity waves of various λ_z.

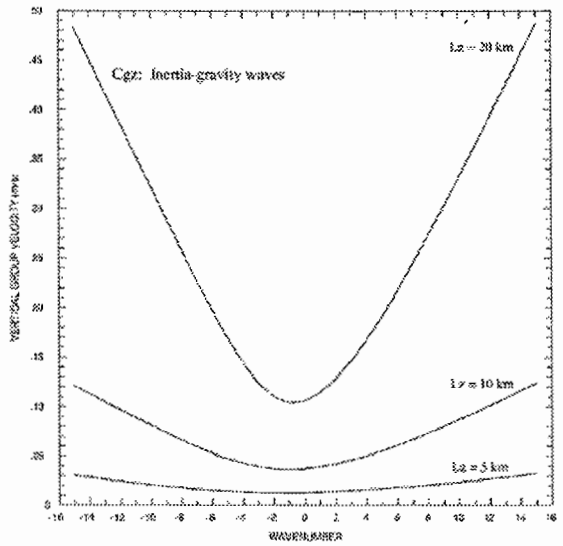


Figure 13: Vertical group velocities for inertia-gravity waves of various λ_z.

4.3 Vertical phase and group propagation

Kelvin and gravity waves waves have the dispersion relationship:

$$\omega = -\frac{kN}{m} \tag{61}$$

This is exact for Kelvin waves (cf. Section 2.1, but also applies for gravity waves except at the largest scales (cf. Section 2.2.1). For vertically-propagating waves with tropospheric sources, the requirement that the group velocity be directed upward implies that:

- Eastward-propagating waves ($\omega > 0$) have $m < 0$
- Westward propagating waves ($\omega < 0$) have $m > 0$
- Kelvin waves behave like eastward-propagating gravity waves

Given these relationships, it follows that constant-phase lines in the $z - x$ plane obey the equation:

$$kx + mz = C \tag{62}$$

where C is an arbitrary constant. Hence,

$$z_p = -\frac{k}{m} \cdot x_p = \begin{cases} -\dfrac{k}{|m|} \cdot x_p & (\omega < 0) \\ \dfrac{k}{|m|} \cdot x_p & (\omega > 0) \end{cases} \tag{63}$$

(since $\omega > 0$ when $m < 0$ and $\omega < 0$ when $m > 0$). Thus, (63) describes the locus (x_p, z_p) of a constant phase line in the $x - z$ plane (where x denotes the direction of horizontal phase propagation). Fig. 14 illustrates this relationship.

Similarly, constant phase lines in the $z - t$ plane obey:

$$mz - \omega t = C \tag{64}$$

whence

$$z_p = \frac{\omega}{m} \cdot t = -\left|\frac{\omega}{m}\right| \cdot t_p \tag{65}$$

so that (65) is the loci of lines of constant phase in the $z - t$ plane, as shown in Fig. 15.

Finally, note that the group velocities are given by:

$$c_{gz} = \frac{\partial \omega}{\partial m} = \frac{kN}{m^2} \tag{66}$$

in the vertical direction, and

$$c_{gx} = \begin{cases} -\dfrac{N}{|m|} \cdot t & (\omega < 0) \\ \dfrac{N}{|m|} \cdot t & (\omega > 0) \end{cases} \tag{67}$$

in the horizontal direction. Then ray paths in the $x - z$ plane can be obtained from:

$$\frac{dz}{dx} = \frac{c_{gz}}{c_{gx}} \tag{68}$$

so that, from (66) and (67):

$$z_g = \begin{cases} -\dfrac{k}{|m|} \cdot x_g & (\omega < 0) \\ \dfrac{k}{|m|} \cdot x_g & (\omega > 0) \end{cases} \tag{69}$$

describes the direction of ray paths in the $z - x$ plane.

Comparing this with the behavior of phase lines in the $z - x$ plane (63), we see that the group velocity in the $z - x$ plane is parallel to lines of constant phase. This is also shown in Fig. 14.

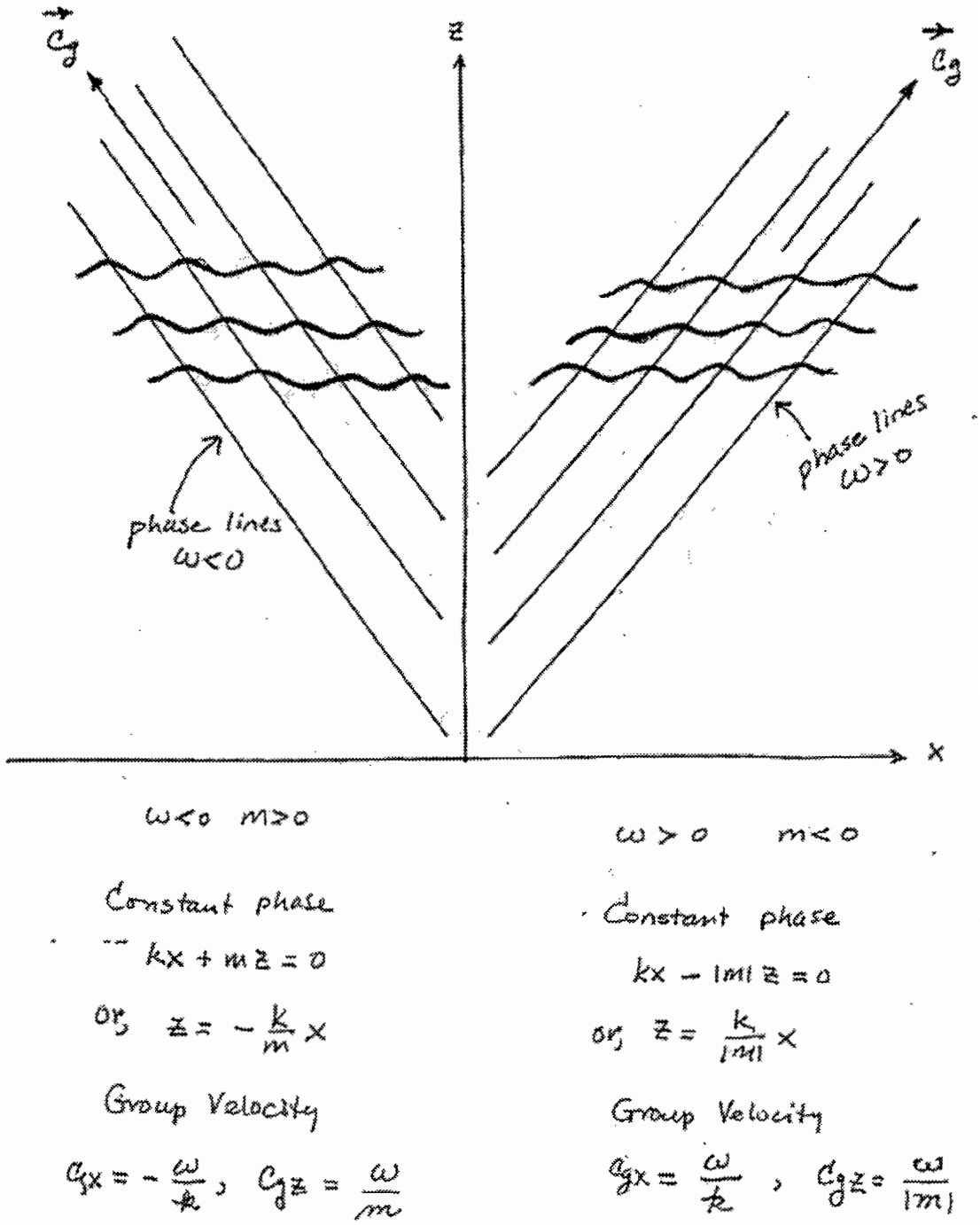
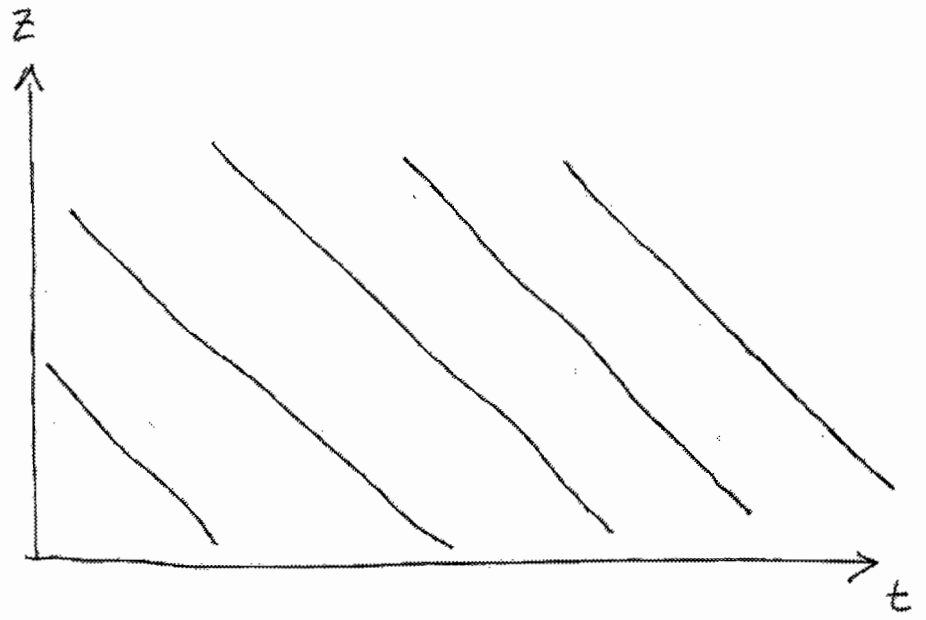


Figure 14: The behavior of constant phase lines in the $x - z$ plane, where z is altitude and x denotes the direction of horizontal wave propagation. The relationship between phase behavior and group velocity is also shown.



$z = \frac{\omega}{m} \cdot t$ $\quad m < 0$ for $\omega > 0$, $m > 0$ for $\omega < 0$

$\Rightarrow$ Phase lines of upward-propagating waves slope downward with time:

$z = -\left|\frac{\omega}{m}\right| \cdot t$

Figure 15: The behavior of constant phase line sin the $z - t$ plane. Upward-propagating waves are characterized by phase descent as a function of time.

5 Observations of Waves in the Tropics

A variety of techniques are used to infer the presence of waves in the tropical atmosphere:

- Large-scale waves are observable by conventional networks (radiosondes, rocketsondes) or by satellites (LIMS, UARS, etc)
- Large-scale waves in the Tropics were documented earliest and more thoroughly because
  - They generally occur at lower frequencies so sampling is easier
  - They have larger amplitudes at lower levels (lower, middle stratosphere)
  - They are detectable without aliasing by satellite platforms (these typically resolve up to 6 zonal wavenumbers and diurnal or longer periods)
- Smaller scale (and higher frequency) waves are harder to document, especially on a global basis
  - They often have small amplitudes at lower levels
  - Theyt are aliased in satellite observations
  - They are detectable by lidar/radar/sonde observations, but geographical coverage is limited
  - Nevertheless, small scale and higher frequency waves are likely quite important for driving the QBO and SAO

The following figures show examples of wave observations in the Tropics. After the publication of Matsuno's (1966) and other theoretical works on tropical waves, the existence of **large-scale Kelvin and Rossby-gravity waves** was documented in several studies:

- Wallace and Kousky (1968), see Fig. 16,
  - used radiosonde measurements of u and T from several tropical stations (Balboa, Trinidad, Kwajalein)
  - established the presence of a large-scale oscillation, and related it to the Kelvin wave predicted by theory (signal observed in zonal but not meridional wind)

- the wave period was $\sim$ 15 days, and the zonal wavenumber was $k = 1 - 2$ (see Fig. 17)
- the waves were estimated to have vertical wavelength of 6-10 km and phase velocity of 20 m s^{-1}

- Yanai and Maruyama (1966), see Fig. 18,
  - used radiosonde wind from several tropical Pacific stations
  - found a westward-propagating wave in the lower stratosphere, in westerly background winds
  - wave was present to about 22 km, where the winds became easterly
  - $\tau \sim 5$ days, $k \sim 4 - 5$, $c \sim -23$ m s^{-1}

Other large-scale waves, at higher altitudes in the tropical middle atmosphere, were detected using rocketsonde observations. It will be recalled that higher frequency waves are expected to propagate more readily to high altitudes because they are dissipated less readily than their lower frequecny counterparts. Thus, for, example,

- Hirota (1978), see Fig. 19,
  - used rocketsonde data from Ascension Island, giving information on u, v, T through the lower mesosphere
  - the data covered a wide range of altitudes in the tropical stratosphere and lower mesosphere
  - wavelike signals were found to be prominent in *uiandT*, but not in v, suggesting that the observed oscillation was a Kelvin wave
  - identified the oscillation as a high-frequency Kelvin wave (now known as the "Hirota Kelvin wave", or fast Kelvin wave)
  - the wave identified had $\lambda_z \simeq 15 - 20$ km, and $c \simeq 55 - 70$ m s^{-1}
  - the wave amplitude was found to depend on the phase of the SAO, suggesting that this wave play a role in forcing this oscillation in the zonal-mean zonal wind (Fig. 20)

Since the late 1970's **observations from polar-orbiting, limb-scanning satellites** have provided global, synoptic views of large-scale waves of low to moderate frequency:

- Coy and Hitchman (1984)
  - used observations of T by the LIMS instrumnet on Nimbus-7
  - observed a $k = 2$, eastward-propagating wave, identifiable as a Kelvin wave
  - wave parameters were estimated to be as follows: $\tau = 6$ days, $\lambda_z = 15$ km, $c_{gz} = 2.8$ km day^{-1}, $c \simeq 40$ m s^{-1}
- Salby et al (1984)
  - used temperature observations from the LIMS satellite, covering altitudes to about 65 km
  - found concentration of spectral variance in the Tropics, consistent with Kelvin wave structure, Fig. 21
  - wave periods become shorter at higher altitudes
  - documented the presence of $k = 1, 2$ waves
  - fast waves in lower mesosphere, $c \simeq 120$ m s^{-1}, $\lambda_z = 40$ km
  - wave structures were found to be consistent with theoretical expectations for Kelvin waves (Fig. 22)
- Canziani et al (1994), Fig. 23
  - temperature data from the MLS (microwave limb sounder) instrument on the UARS polar-orbiting satellite
  - clearly established the presence of Kelvin waves, $k = 1-2$, throughout the stratosphere
  - phase velocities were in the range $c \simeq 30 - 115$ m s^{-1}

Intermediate-scale, inertia-gravity waves have also been observed by a variety of techniques. Note, however, these waves are not observable by satellite instruments because their wavelengths (too small) and/or frequencies (often too high) cannot be resolved with the sampling patterns of polar-orbiting satellites. Instead, identification is made from a variety of local measurements, for example:

- Cadet and Teitelbaum (1979), Fig. 24
  - obtained observations of u, v from radar-tracked balloons during GATE (GARP Tropical Atlantic Experiment)

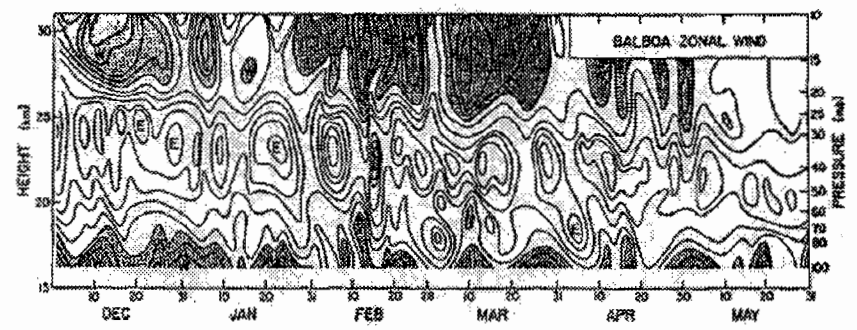
  - identified an inertia-gravity wave, consistent with meridional index $n = 1$
  - the wave was westward-propagating, had small vertical wavelength, and period $\tau \simeq 1.5$ days

- Tsuda et al. (1994)
  - detected oscillations in radiosonde data from Indonesia
  - waves present up to 35 km
  - eastward-propagating, $c \simeq 12$ m s^{-1}, small vertical wavelength (< 3 km)

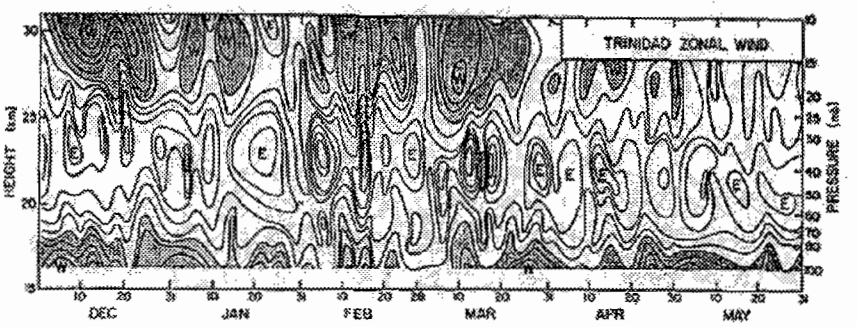
- Sato and Dunkerton (1997), Fig. 25
  - used Singapore radiosonde data
  - identified several Kelvin and inertia-gravity wave modes
  - zonal wavelengths ranging from about 1000 to perhaps 10,000 km
  - the presence of both eastward- and westward-propagating waves was inferred from the observations

The study of Sato and Dunkerton is especially interesting in that they used a novel technique involving co-spectra and quadrature spectra of zonal wind and temperature to deduce the momentum fluxes in the eastward and westward directions. They then identified the wave types responsible for these momentum fluxes by comparing observed parameters with theoretical expectations derived from the dispersion relations for the various types of wave.

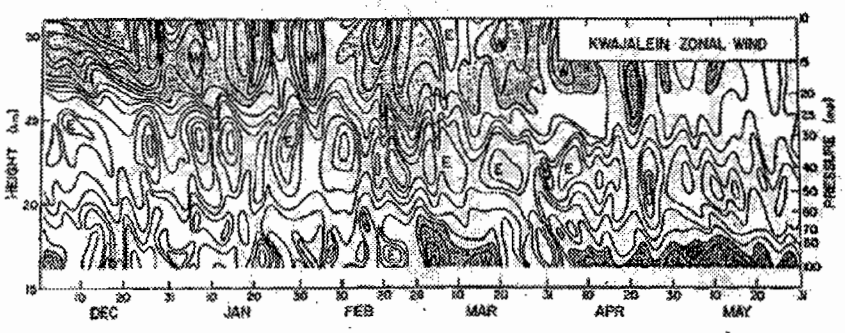
WALLACE AND KOUSKY (1968)



Time-height section of zonal wind at Balboa. Isotachs are placed at intervals of 5 m sec^{-1}. Westerlies are shaded.



Time-height section of zonal wind at Trinidad. Isotachs are placed at intervals of 5 m sec^{-1}. Westerlies are shaded.



Time-height section of zonal wind at Kwajalein. Isotachs are placed at intervals of 5 m sec^{-1}. Westerlies are shaded.

Figure 16: Altitude vs. time observations of the zonal wind at various tropical locations used to identify the presence of Kelvin waves. From Wallace and Kousky (1968).

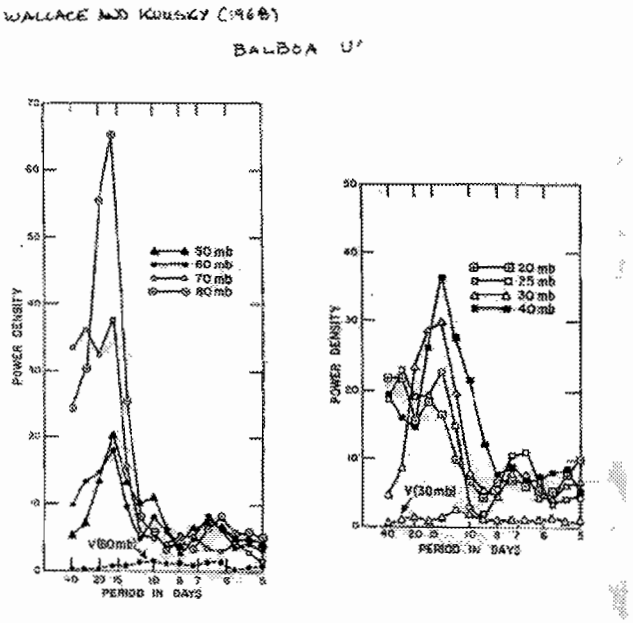


Figure 17: Spectra of the wind at Balboa, showing peak response in the zonal wind near 15 days over a wide range of altitudes. Note very small power for the meridional wind component (v). From Wallace and Kousky (1968).

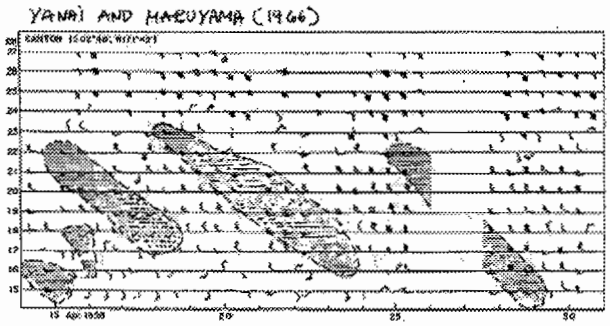
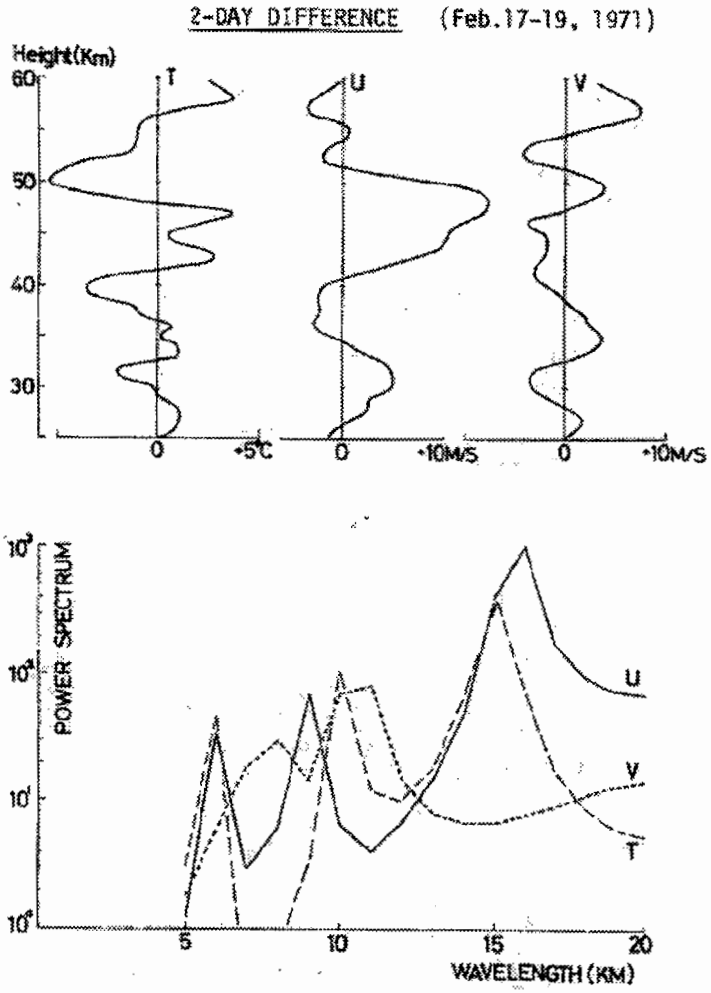


Figure 18: Phase behavior (altitude vs. time) of the zonal wind at Canton Island used to identify westward-propagating Rossby-gravity waves. Southerly winds are shaded. Note the phase descent with time of phase lines, characteristic of upward-propagation. From Yanai and Maruyama (1966).

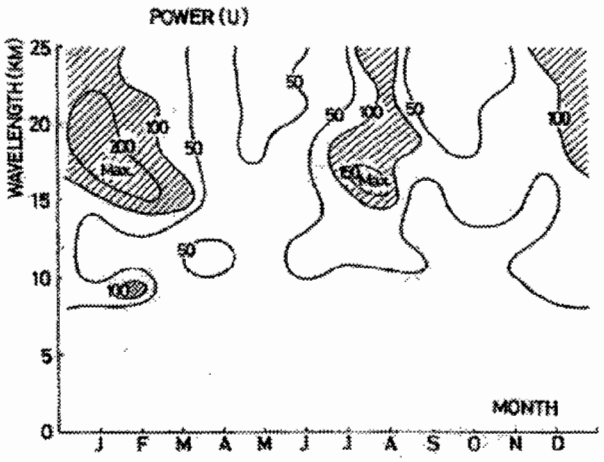
HIROTA (1978)



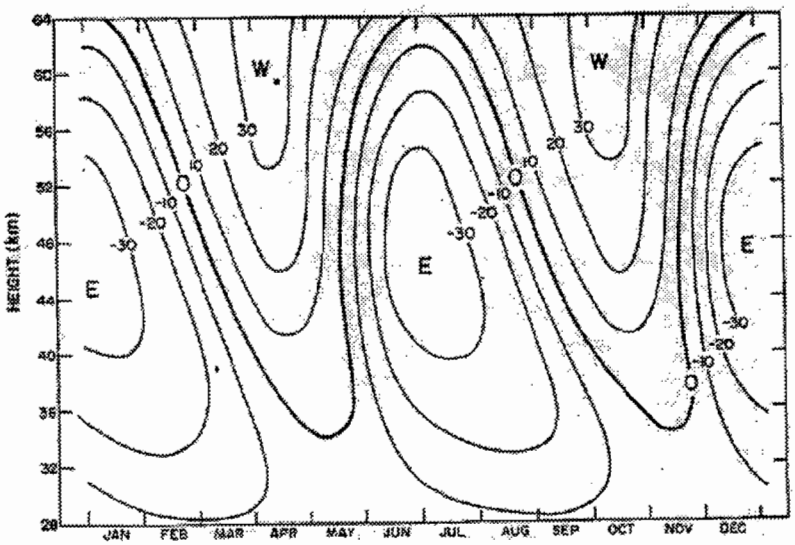
An example of two-day difference of temperature (T) and zonal (U) and meridional (V) wind component (above) at Ascension Island for 17–19 February 1971. Units are K day^{-1} and m s^{-1} day^{-1}. Power spectra of each component are shown as a function of vertical wavelength (below). Units are (K day^{-1})^2 km and (m s^{-1} day^{-1})^2 km.

Figure 19: Rocketsonde observations used by Hirota to identify the fast Kelvin wave at Ascension Island. From Hirota (1978).

HIROTA (1978)



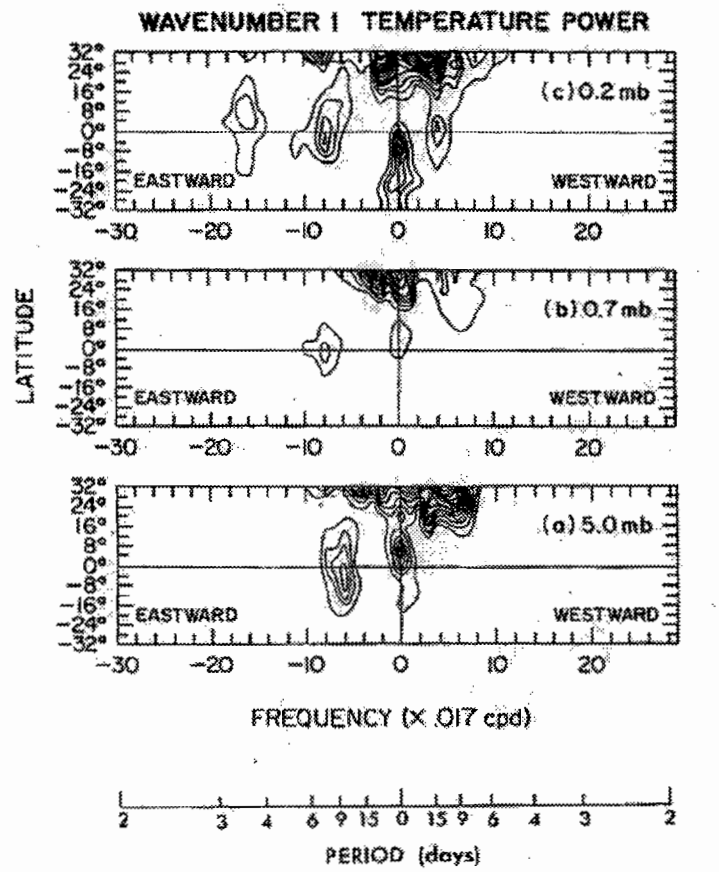
Seasonal variation of power spectral density for the zonal wind component as a function of the vertical wavelength during the period of 1971 and 1972. Units are $(m\ s^{-1}\ day^{-1})^2$ km.



Estimated variation of zonal wind (in m sec^{-1}) with month and altitude at the equator.

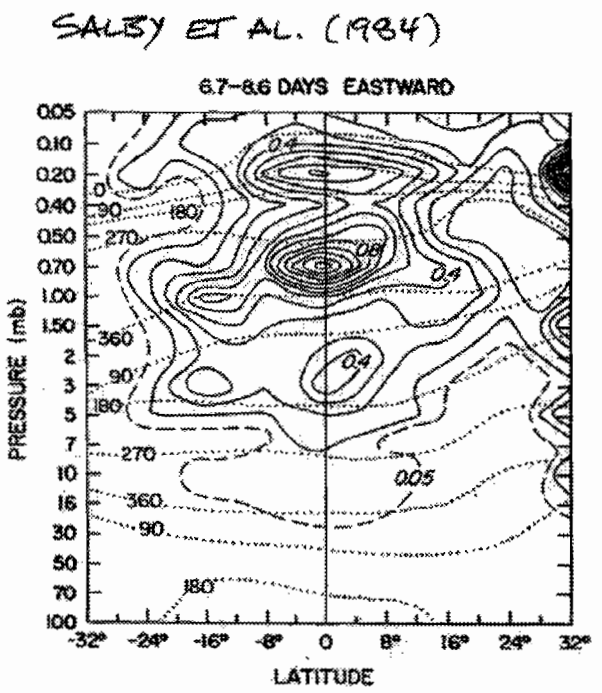
Figure 20: Relationship between power density of the zonal wind component (top) and the zonal-mean zonal wind (bottom) at Ascencion Island. From Hirota (1978).

SALBY ET AL (1984)

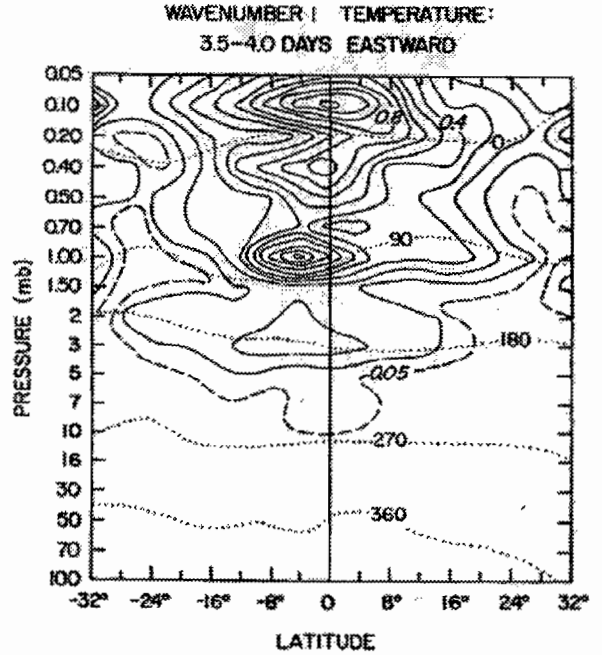


Temperature power for wavenumber 1 as a function of frequency and latitude for the January–February sample period at (a) 5.0, mb contour increment = 0.02 K^2, (b) 0.7 mb, contour increment = 0.10 K^2, and (c) 0.2 mb, contour increment = 0.04 K^2.

Figure 21: Power spectra from LIMS satellite observations used to identify the presence of Kelvin waves in the tropical stratosphere and mesosphere. From Salby et al. (1984).



Wavenumber 1 power (solid) and phase (dotted) structures, corresponding to the eastward spectra integrated between 6.7 and 8.6 days, (54–69 m s^{-1}) as functions of latitude and pressure.



As in Fig. 7, but for the faster wavenumber 1 feature, integrated between 3.5 and 4.0 days (115–135 m s^{-1}). Contour increments as for Fig. 7.

Figure 22: Structure (amplitude and phase) of Kelvin waves derived by Salby et al, from LIMS data. From Salby et al. (1984).

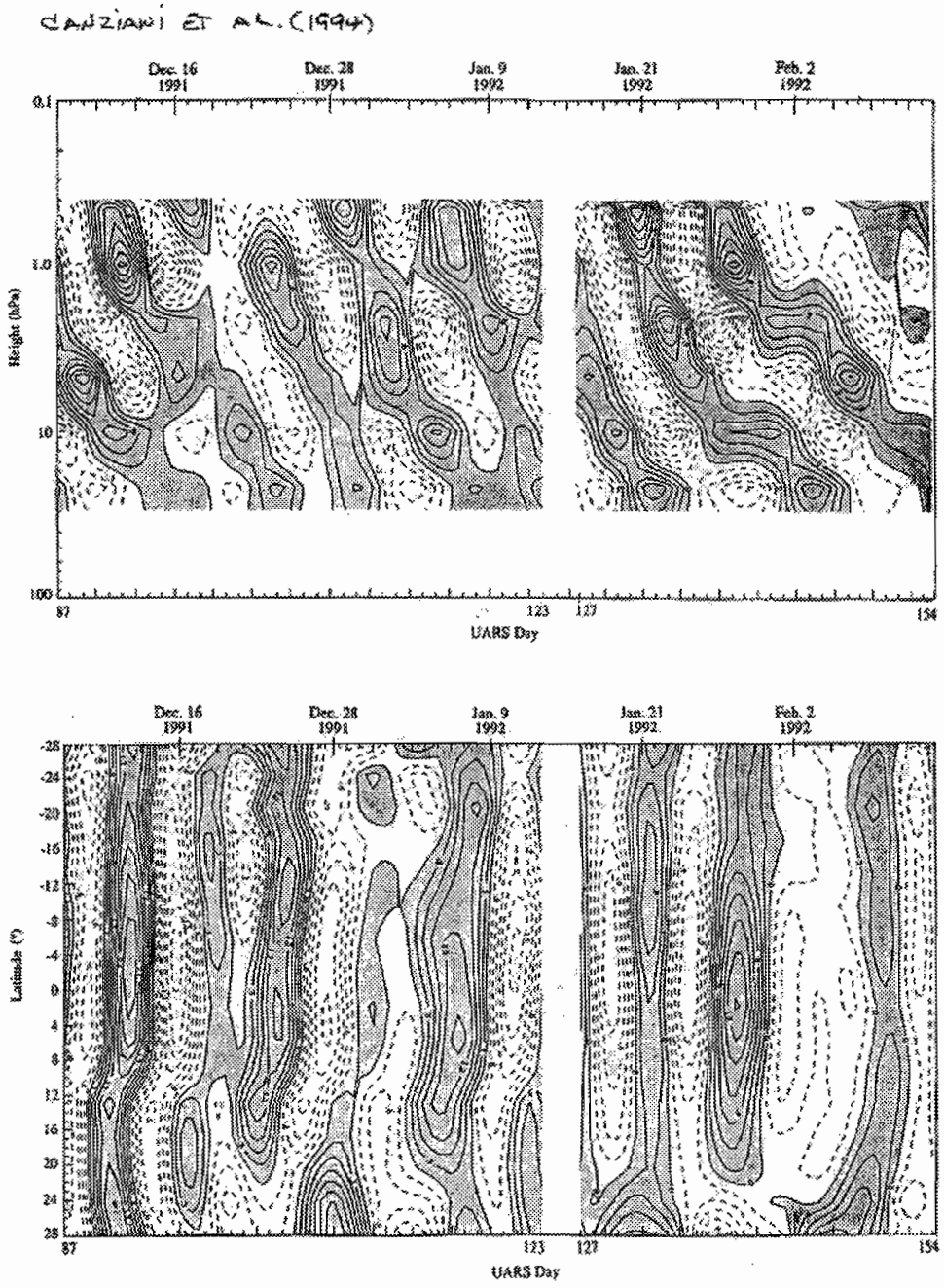


Figure 23: Top panel: time-height plot of temperature at the Equator, band-passed between 5.5 and 20 days. Phase descent with time implies upward propagation. Bottom panel: time-latitude plot at 1 mb (about 45 km) showing meridional structure suggestive of a Kelvin wave. From Canziani et al. (1994).

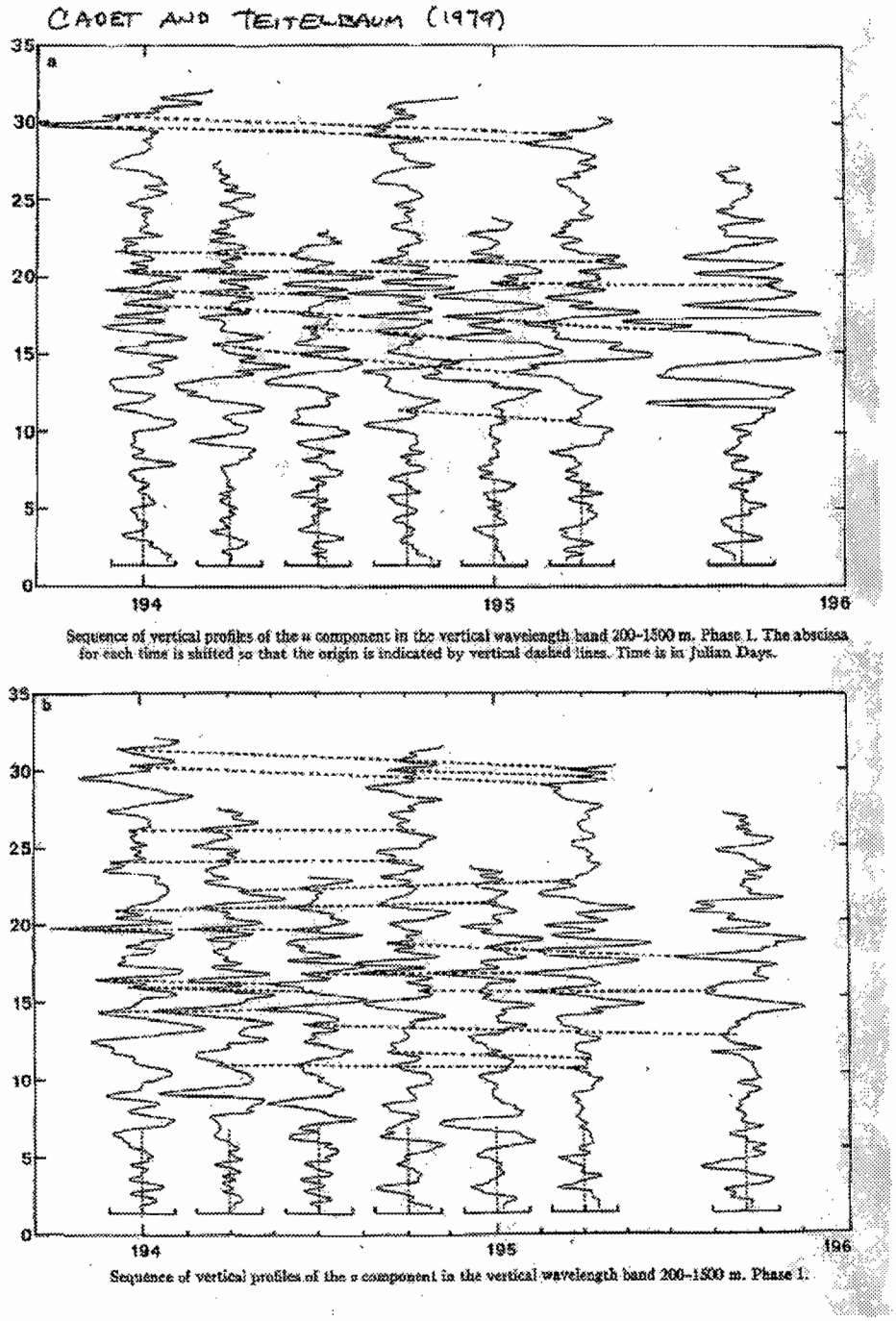


Figure 24: Vertical profiles of the u (top) and v (bottom) components of the wind obtained from radar-tracked balloons, showing the presence of short vertical wavelength inertia-gravity waves. From Cadet and Teitelbaum (1979).

SATO AND DUNKERTON (1997)

Table 1. Parameters of Possible Equatorial Waves for Observed 1–3 Day Components in Westerly Shear

Mode	c_r, m s^{-1}	λ_x, km	λ_z, km	y_0, $\times 10^3$ km
$n = -1$ Kelvin waves	>6.7	>870	>2.0	>550
$n = 1$ eastward IGW	7.7–19.6	1000–2500	2.0–4.2	550–790
$n = 3$ eastward IGW	9.4	1200	2.0	550
$n = 1$ westward IGW	−25 to −31.8	3200–4100	5.7–6.7	920–1010

The zonal phase velocity, zonal wavelength, vertical wavelength, latitudinal scale factor are denoted by c_r, λ_x, λ_z, and y_0, respectively. IGW, inertia-gravity waves.

Table 2. Same as Table 1 but for Easterly Shear

Mode	c_r, m s^{-1}	λ_x, km	λ_z, km	y_0, $\times 10^3$ km
$n = -1$ Kelvin waves	>15	>1900	>2.7	>820
$n = 1$ eastward IGW	15–44.1	1900–5700	3.4–6.6	710–970
$n = 3$ eastward IGW	15–20.6	1900–2700	2.6–3.1	630–660
$n = 1$ westward IGW	−7.2 to −61.8	930–8000	2.0–9.3	550–1180
$n = 3$ westward IGW	−8.6 to −23.8	1100–3100	2.0–3.6	550–730
$n = 5$ westward IGW	−11.7 to −14.7	1500–1900	2.0–2.2	550–570

Figure 25: Estimates of typical parameters for equatorial waves under westerly and wasterly shear conditions in the stratosphere. From Sato and Dunkerton (1997).

6 The Excitation of Equatorial Waves

Upward-propagating waves in the Tropics are believed to be excited principally by release of latent heat in deep convective systems. Deep convection is a major source of energy input to the tropical atmosphere. In a time-mean sense, convection redistributes heat and drives the Hadley and Walker cells; in addition, the unsteadiness of convection is the major excitation mechanism for transient motions in the tropics. Fig. 26 shows schematically how convective heating can excite wave motions. The heating, $Q'(\mathbf{x}, t)$ constitutes a forcing term in the thermodynamic equation. Depending on the spatial and temporal characteristics of Q', waves will be more or less efficiently excited by such forcing. Recall from Section 4.2 that high-frequency waves are more likely to propagate to higher altitudes, so high-frequency variability of Q' is an important factor in the excitation of upward-propagating waves.

Although the importance of convective excitation is is well established, the study of its effects in General Circulation Models (GCM) is hampered by the fact that convection occurs on spatial scales too small to be resolved by any practical GCM, so its effects must be parameterized somehow. That is, a method must be devised that produces an estimate of convective effects (e.g., precipitation rate, rate of latent heat release) in terms of fields resolved by the model (e.g., temperature, wind divergence, water vapor mixing ratio).

The parameterizations of convection is an "art", relying not just on physical considerations, but also on experience an empirical "tuning" of convection codes. In general, convective parameterizations produce reasonable agreement between computed annual or seasonal mean precipitation and observations thereof. On the other hand, the variability of parameterized convection has not received much attention until fairly recently, but it is now clear that:

- Different convective parameterizations produce widely different variability, especially at high frequencies (periods < 5 d).
- Convective variability determines the nature of the vertically propagating wave spectrum produced by GCMs
- Some of the difficulties encountered in the simulation of the tropical wind oscillations can be ascribed to insufficient wave forcing by parameterized convection

Clearly the most desirable way of determining the effects of convection would utilize a method that observes directly the convective processes in

Wave Excitation by Convective Heating: Physical Mechanism

$$\left\{\frac{\partial}{\partial t}+\bar{u}\frac{\partial}{\partial x}\right\}T' + w'\frac{HN^2}{R} = Q'_{conv}$$

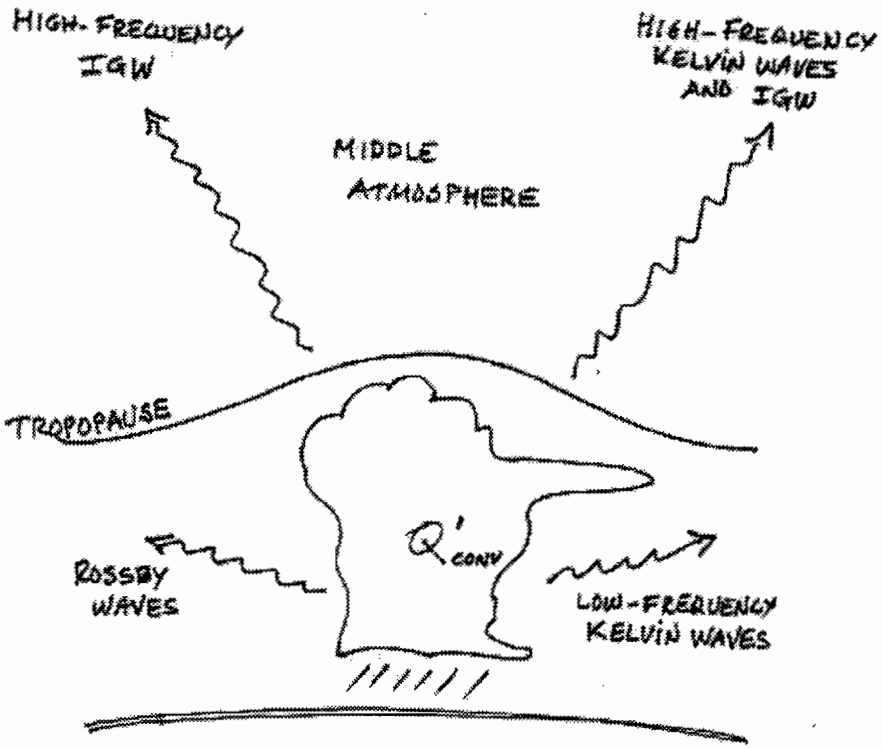


Figure 26: Schematic representation of the excitation of Tropical waves by convective heating.

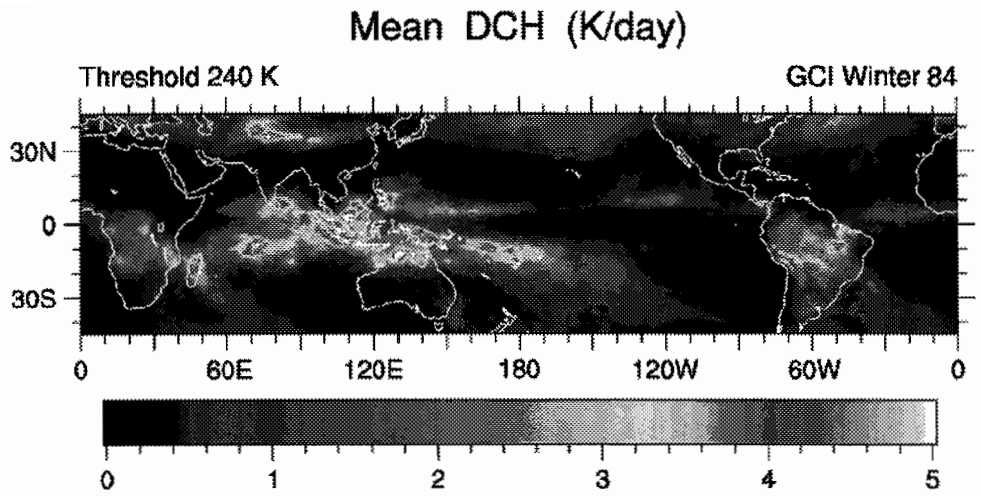
question. For example, tropical precipitation can be measured by a network of raingages or observed by radar. Since precipitation is the result of the condensation of atmospheric water vapor, precipitation rate is directly related to the release of latent heat of condensation. Thus, reliable observations of the former are also a good indicator of the latter. However, observations of precipitation are often not available on a global basis, so it is necessary to use other, indirect means of estimating precipitation and the attendant latent heat release.

OLR observations have been made from satellites for many years. The International Satellite Cloud Climatology Project (ISCCP) gathers data from several polar orbiting and geostationary satellites (Schiffer and Rossow, 1983). Some of these data has been reprocessed by M. Salby and colleagues (Tanaka et al., 1991; Salby et al., 1991) into a dataset of OLR covering the entire globe, with horizontal resolution of 0.35° of latitude by 0.7° of longitude, and time resolution of 3 hours. Hendon and Woodberry (1993) have shown how OLR observations can be used to infer convective activity, and this method has been used by Bergman and Salby (1994) and Ricciardulli and Garcia(2000) to estimate global spectra of convective heating.

It should be noted, however, that OLR methods may overestimate the variance of convection. Recent work by Horinouchi (in press) suggests that overestimation of variance by the OLR method can be severe in some cases. On the other hand, the shape of the spectrum derived from OLR observations is similar to that obtained from redar measurements. See also Ebert et al (1996), and Ebert and Manton (1998).

In spite of this caveat, inferences about convective variability made form OLR observations turn out to be useful for characterizing the excitation of tropical waves by convective heating. Fig. 27 shows the time mean and standard deviation of convective heating inferred from OLR observations for winter of 1984. As indicated by the standard deviation map, convective variability is large over the oceanic Intertropical Convergence Zone (ITCZ) and also over the Tropical continents (South America, Africa and Oceania). Fig. 28 shows how the variance is distributed as a function of period. It is remarkable that, over the tropical land masses, most of the variability is concentrated at periods shorter than 2 days.

The wavenumber-frequency spectrum of convective heating inferred from OLR observations is red, with variance extending to large values of wavenumber and frequency. Superimposed on this red background, there are concentrations of variance at the diurnal frequency and its harmonics, a reflection of the importance of the diurnal cycle in tropical convection. Given the



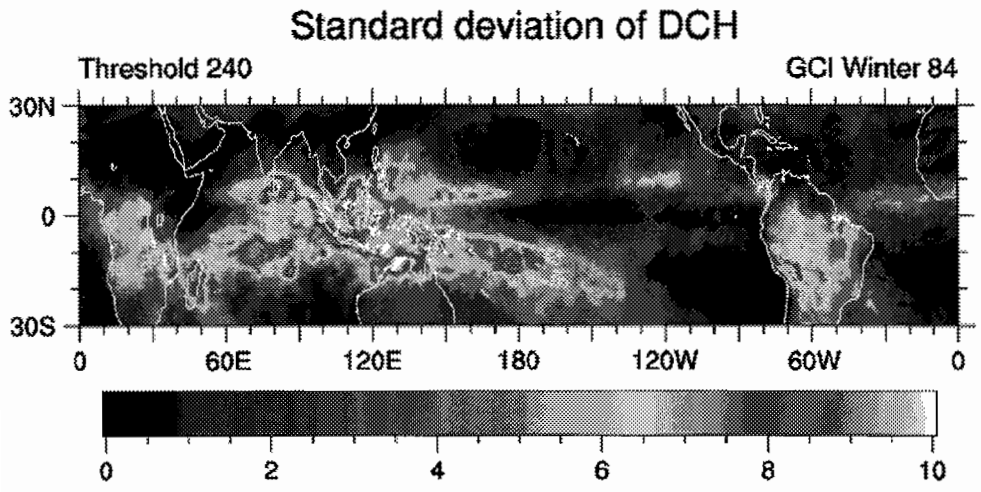


Figure 27: Mean and standrad deviation of convective heating in the global Tropics, as inferred from observations of OLR. From Ricciardulli and Garcia (2000).

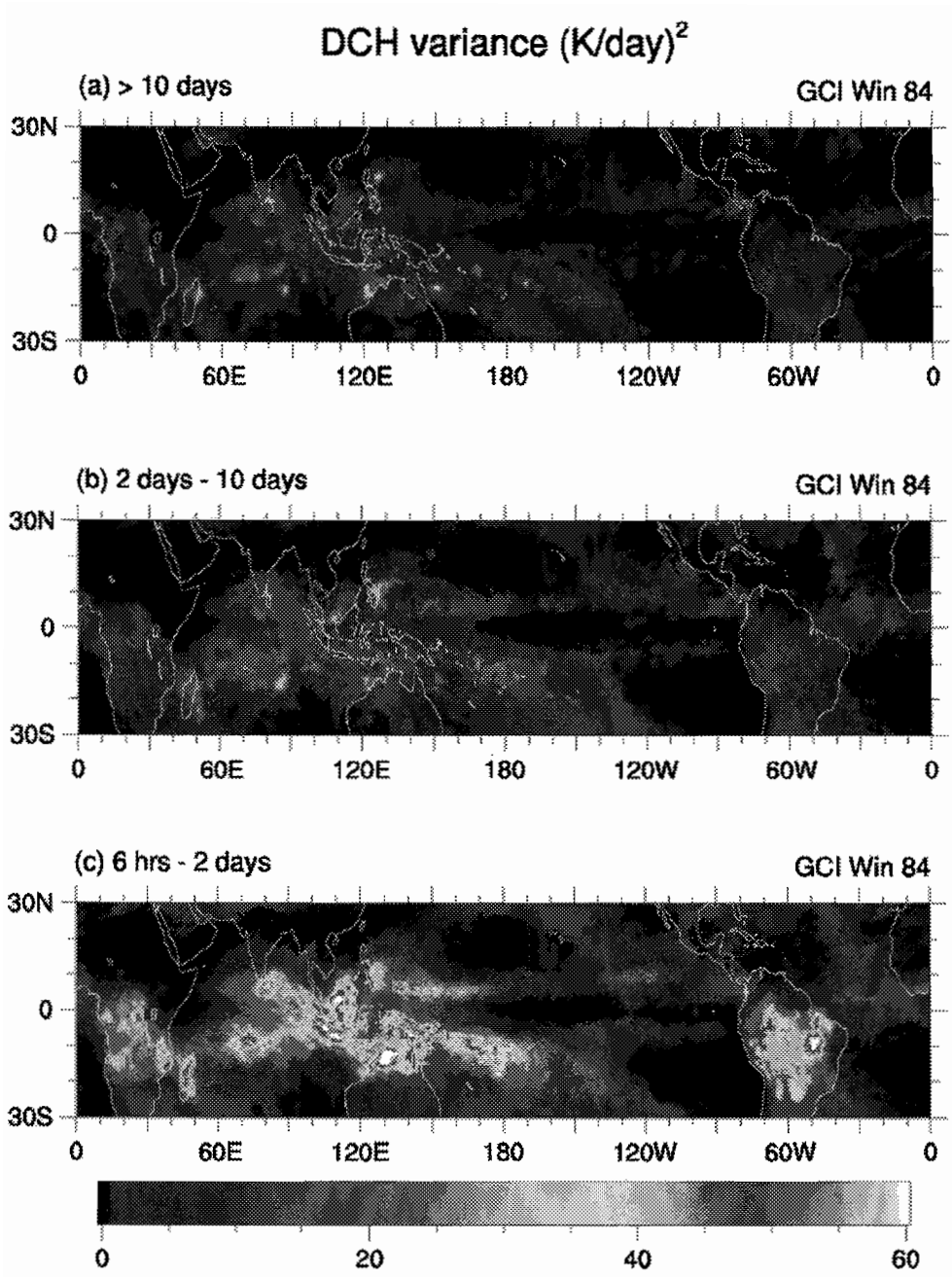


Figure 28: Mean and standard deviation of convective heating in the Tropics, as inferred from observations of OLR. From Ricciardulli and Garcia (2000).

distribution and variability of convective heating inferred from the OLR observations, it is possible to apply the method of Salby and Garcia (1987) to estimate the amplitude of the waves that are excited, and their contribution to the vertical component of EP flux, F_z, emanating from the troposphere.

Comparison of F_z inferred from OLR observations, and calculated in a similar way from various GCMs indicates that model performance is exteremly variable, depending on the method used to parameterize convection. Fig. 29 shows an example of the frequency distribution of F_z obtained from OLR observations and from the NCAR Community Climate Model (CCM3) run with two different convective parameterizations: those of Zhang-McFarlane (1995) and Hack (1994). It is clear that the two parameterizations differ from each other, and also from the OLR estimates. Both of the parameterizations shown here underestimate F_z at high frequencies, and the Zhang-MacFarlane parameterization does so also at low frequencies.

These differences have major effects on the GCM simulations. For example, Fig. 30 shows that, when the model is run with the Zhang-McFarlane parameterization, an SAO is produced only when an additional source of forcing is included, whereas with the Hack parameterization, an SAO can be driven just by the waves excited by the oenvective parameterization. Fig. 31 shows that the Madden-Julian Oscillation (see, e.g., Madden and Julian, 1994), a major observed feature of the tropical troposphere, is simulated realistically when the Hack parameterization is used in the GCM.

Summarizing the points discussed above:

- Convective parameterizations are usually designed to reproduce the mean distribution of convective heating
- Parameterized convective heating often misrepresents actual heating variability
- Convective heating is the main excitation mechanism for transient motions in the Tropics:
  - Vertically-propagating waves drive the QBO and SAO
  - High phase velocity waves are likely to be most effective for the SAO; lower phase velocity waves, for the QBO
  - The tropospheric Madden-Julian Oscillation appears to depend also on the availability of convective forcing (at lower frequencies)

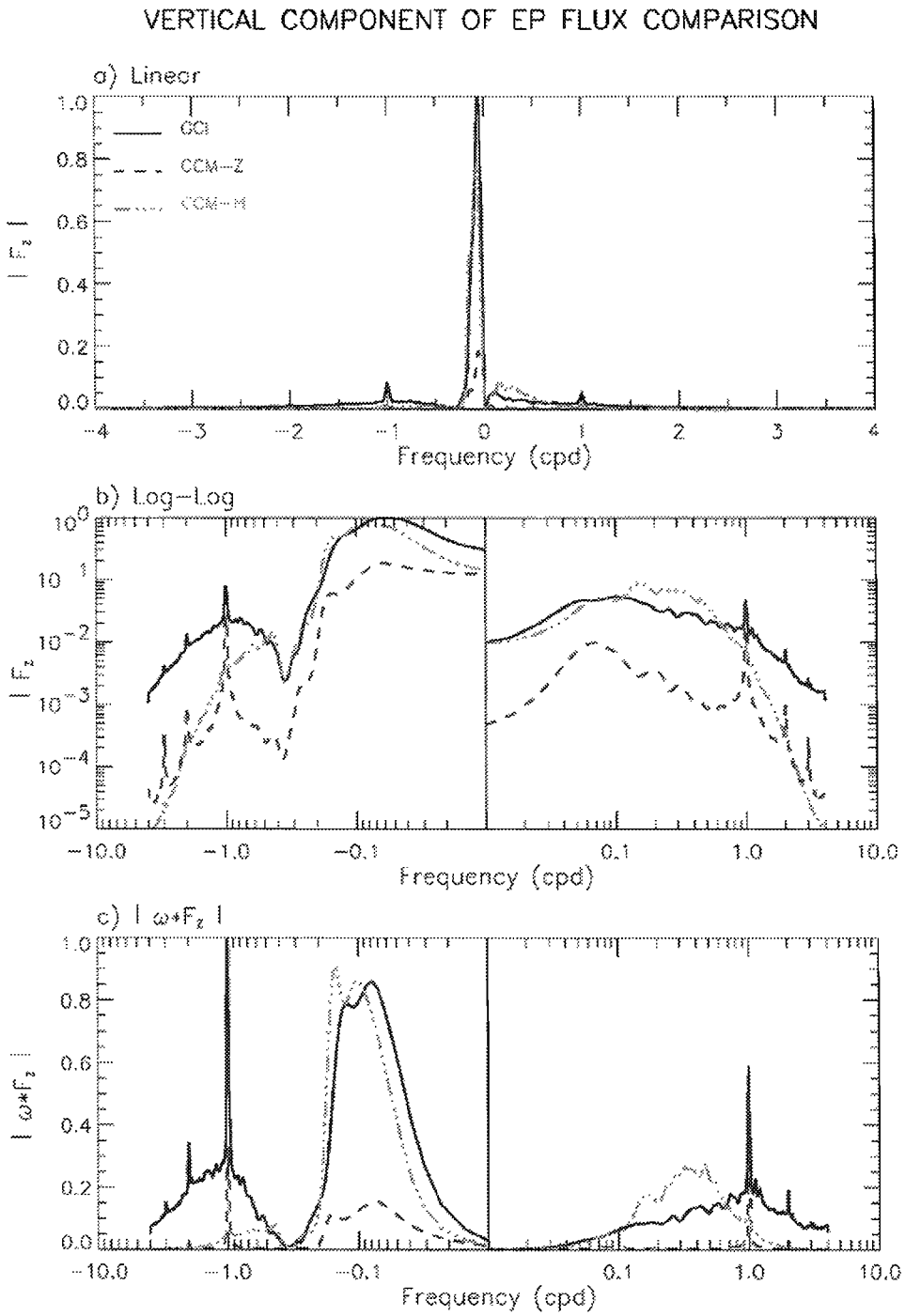


Figure 29: The vertical component of EP flux as a function of frequency derived from OLR observations, and from a GCM using two different convection parameterizations.

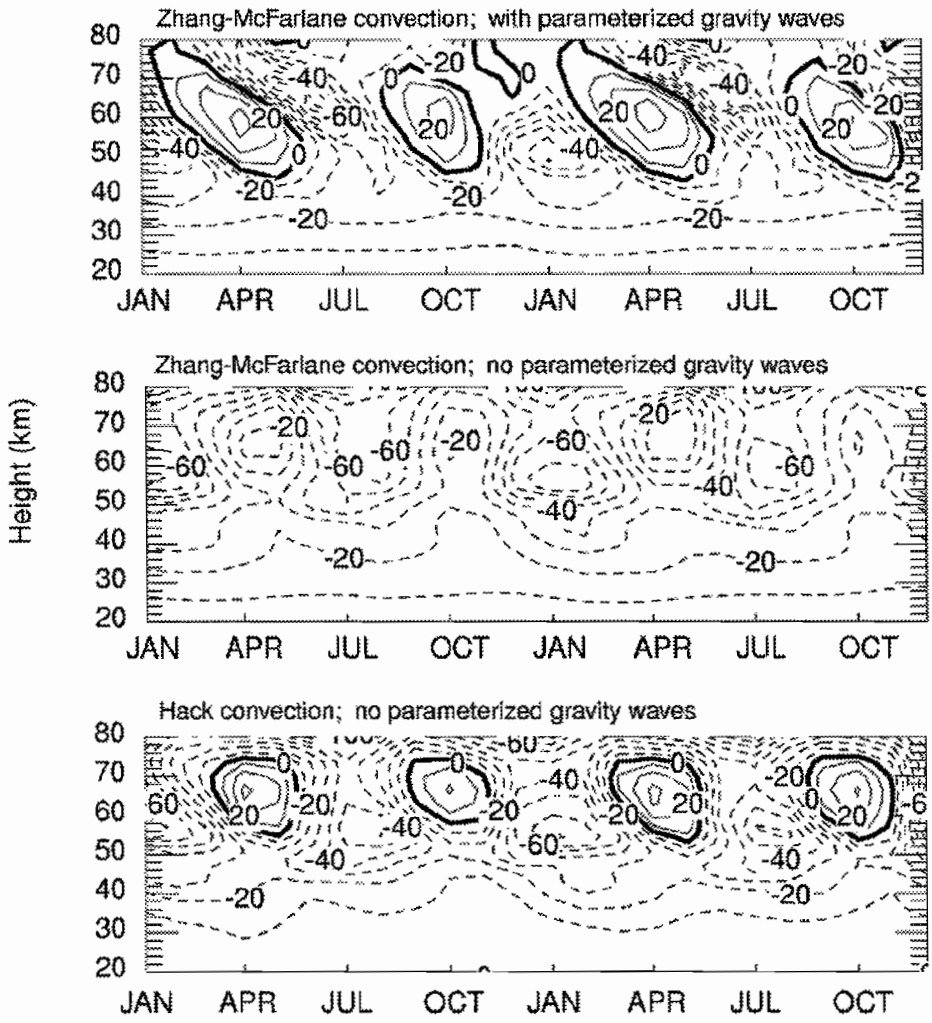


Figure 30: Annual cycle of the zonal-mean zonal wind at the Equator in the NCAR CCM. Top: model run with the Zhang-McFarlane parameterization of convection plus additional forcing by small-scale gravity waves. Middle: same as top, but with small-scale gravity wave forcing removed. Bottom: model run with the Hack parameterization and no additional forcing by small-scale gravity waves. A semiannual oscillation is obtained in the runs shown in the top and bottom panels, but for entirely different reasons.

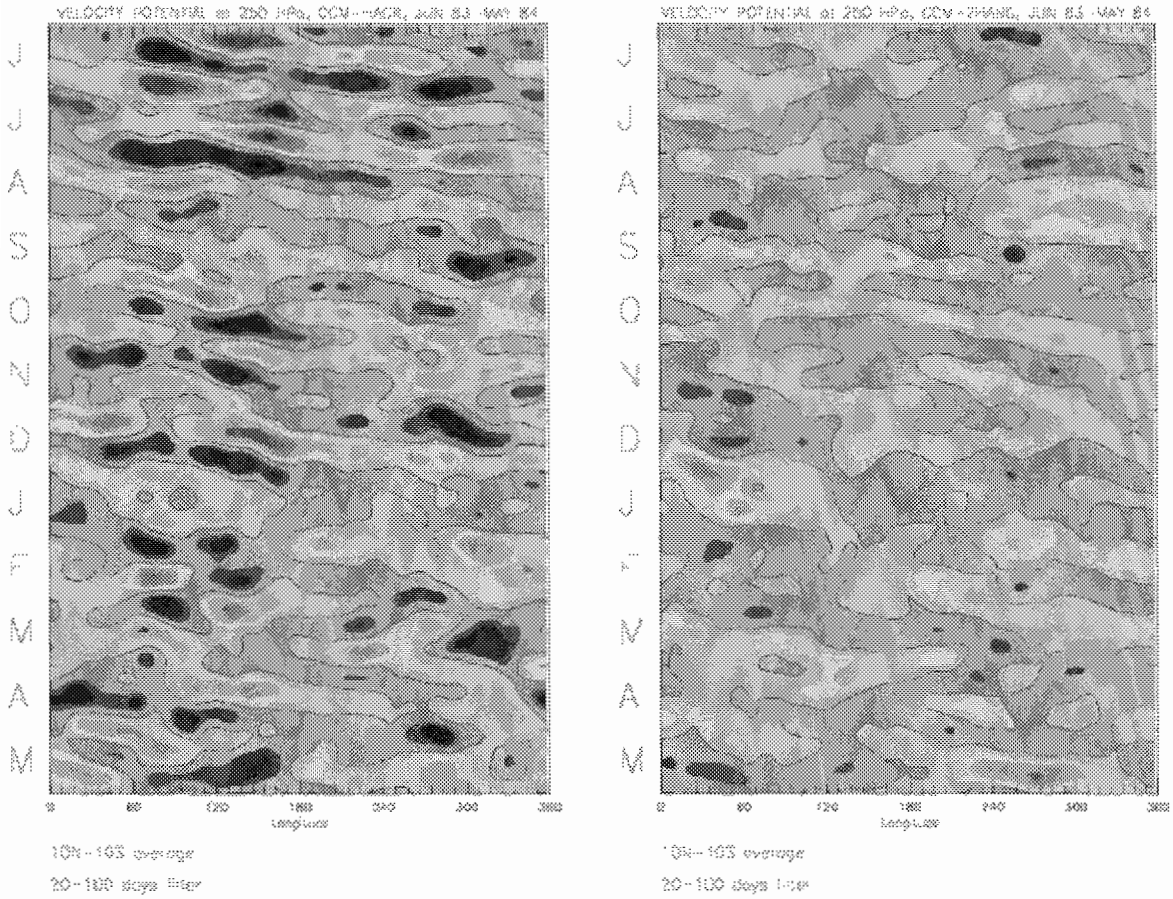


Figure 31: Velocity potential at 250 mb for runs using the Hack (left) and Zhang-McFarlane (right) parameterizations of convection. A much more realistic MJO is obtained with the Hack parameterization.

7 Summary

- Equatorial wave theory is able to define the characteristics of waves in the Tropics, at least under idealized conditions
- Theoretical results can be used to identity and interpret wave behavior in the real world
- Observations of tropical waves are abundant since the mid-1970's
- With satellite observations, the presence, morphology, and behavior of large-scale, low to moderate frequency waves is well documented
- Global forcing of the mean state due to large-scale waves can be estimated
- Smaller-scale, higher-frequency waves are more difficult to document because they are often aliased in satellite observations
- Nevertheless, observations of these waves (Kelvin, inertia-gravity) are now plentiful
- Observations of mesoscale ($L_x \sim 100$ km) gravity waves are also becoming common
- Remaining challenges involve
  - determining the effect of the smaller-scale waves on the mean state of the Tropcis
  - estimating the forcing of (resolved) tropical waves by convective heating in the Tropics

8 References

Andrews, D.G., J.R. Holton, and C.B. Leovy, 1987: **Middle Atmosphere Dynamics**, Academic Press, 489 pp.

Baldwin, M.P., et al., 2001: The quasi-biennial oscillation. *Rev. Geophys.*, **39**, 179-229.

Bergman, J.W., and M.L. Salby, 1994: Equatorial wave activity from fluctuations in observed convection. *J. Atmos. Sci.*, **51**, 3791-3806.

Cadet, D. and H. Teiteilbaum, 1979: Observational evidence of inertia-gravity waves in the tropical stratosphere. *J. Atmos. Sci.*, **36**, 892-907.

Canziani, P.O., J.R. Holton, E. Fishbein, L. Froidevaux, and J.W. Waters, 1994: Equatorial Kelvin waves: A UARS/MLS look. *J. Atmos. Sci.*, **20**, 3053-3076.

Coy, L. and M. Hitchman, 1984: Kelvin wave packets and flow acceleration: A comparison of modeling and observations. *J. Atmos. Sci.*, **41**, 1875-1880.

Ebert, E.E., M.J. Manton, P.A. Arkin, R.J. Allam, G.E. Holpin and A. Gruber, 1996: Results from the GPCP Algorithm Intercomparison Programme. *Bull. Amer. Meteor. Soc.*, **77**, 2875-2887.

Ebert, E.E. and M.J. Manton, 1998: Performance of satellite rainfall estimation algorithms during TOGA COARE. *J. Atmos. Sci.*, **55**, 1537-1557.

Gill, A.E., 1982: **Atmosphere-Ocean Dynamics**. Academic Press, London, 662 pp.

Hack, J.J., 1994: Parameterization of moist convection in the National Center for Atmospheric Research Community Climate Model (CCM2). *J. Geophys. Res.*, **99**, 5551-5568.

Hendon, H.H. and K. Woodberry, 1993: The diurnal cycle of Tropical convection. *J. Geophys. Res*, **98**, 16623-16637.

Hirota, I., 1978: Equatorial waves in the upper stratosphere and mesosphere in relation to the semiannual oscillation of the zonal wind. *J. Atmos. Sci.*, **35**, 714-722.

Hirota, I. and T. Niki, 1985: A statisitcal study of inertia-gravity waves in the middle atmosphere. *J. Meteor. Soc. Japan*, **63**, 2055-1066.

Madden, R.A., and P.R. Julian, 1994: Observations of the 40-50 day tropical oscillation – A review. *Mon. Wea. Rev.*, **122**, 814-837.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan*, **44**, 25-42.

Ricciardulli, L. and R.R. Garcia, 2000: The excitation of equatorial waves by deep convection in the NCAR Community Climate Model (CCM3). *J. Atmos. Sci.*, **57**, 3461-3487.

Salby, M.L., D.L. Hartmann, P.L. Bailey, and J.C. Gille, 1984: Evidence for equatorial Kelvin modes in Nimbus-7 LIMS. *J. Atmos. Sci.*, **41**, 220-235.

Salby, M.L. and R.R. Garcia, 1987: Transient response to localized episodic heating in the Tropics. Part I: Excitation and short-time near-field behavior. *J. Atmos. Sci.*, **44**, 458-498.

Salby, M.L., H.H. Hendon, K. Woodberry and K. Tanaka, 1991: Analysis of global cloud imagery from multiple satellites. *Bull. Amer. Meteor. Soc.*, **72**, 467-480.

Sato, K. and T.J. Dunkerton, 1997: Estimates of momentum flux associated with equatorial Kelivn and gravity waves. *J. Geophys. Res.*, **102**, 26,247-26,261.

Schiffer, R.A. and W.B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): the first project of the World Climate Research Program. *Bull. Amer. Meteor. Soc.*, **64**, 779-784.

Tanaka, K.H., K. Woodberry, H.H. Hendon and M.L. Salby, 1991: Assimilation of global cloud imagery from multiple satellites. *J. Atmos. Ocean. Tech.*, **8**, 613-626.

Tsuda, T., Y. Maruyama, H. Wiryosumarto, Sri Woro B. Harijono, and S. Kato, 1994: Radiosonde observations of equatorial atmospheric dynamics over Indonesia. 2. Characteristics of gravity waves. *J. Geophys. Res.*, **99**, 10,507-10,516.

Wallace, J.M. and V.E. Kousky, 1968: Observational evidence of Kelvin waves in the tropical stratosphere. *J. Atmos. Sci.*, **25**, 900-907.

Yanai, M. and T. Maruyama, 1966: Stratospheric wave disturbances propagating over the equatorial Pacific. *J. Meteor. Soc. Japan*, **44**, 291-294.

Zhang, G.J., and N.A. McFarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Centre General Circulation Model. *Atmos. Ocean*, **33**, 407-446.

Chapter 4

Gravity Waves

Robert Vincent,

University of Adelaide, Australia

1 Introduction

Waves usually produce small perturbations in basic atmospheric parameters such as wind and temperature. The problem is how best to determine the perturbation quantities and then to interpret them in terms of useful wave parameters, such as frequency, wavenumber and propagation direction. In addressing these issues it must also be recognized that observational selection is a significant problem. Different techniques are sensitive to different parts of the wave spectrum; no one technique can provide all the information that is necessary to fully characterize the wave field.

2 Basic Wave Parameters

Observations provide information on wave perturbations in wind, temperature, density etc. Ideally, one would like simultaneous, continuous measurements taken with high time and height resolution of the three-dimensional wind velocity (u, v, w) and temperature T over as wide a range of heights as possible. Unfortunately, virtually no instrument is capable of making such measurements, although very specialized wind/temperature lidars are becoming available (Gardner and Yang, 1998). Observations made with different techniques need to be compared and combined in order to assess the role of waves in determining the state of the atmosphere. Some techniques, such as radiosondes and rockets, provide nearly instantaneous vertical profiles of horizontal wind and temperature over a wide height range, albiet with poor temporal resolution. Other methods, such as radars, provide wind measurements with good time coverage, but not always over a wide height range.

Once the raw data are acquired there is then the need to reduce the data in some commonly accepted manner so that results from different sites can be compared. In this context it is useful to represent or parameterize the wave parameters so that they can be incorporated in numerical circulation models of the atmosphere. Different model parameterization schemes represent waves in different ways, but basically the need is for information on one or all of the following:

- Wave amplitudes or energy as a function of time and space.
- Propagation directions in the both the horizontal and vertical.
- The relative importance of different wave sources.
- Wave energy and momentum fluxes as a function of frequency and wavenumber.

2.1 Wave energy

An observed variable can be represented as a perturbation ($'$) superimposed on a basic state $\overline{()}$. In the following, the zonal, meridional, and vertical wind components, and temperature are, respectively, denoted by $\overline{u} + u'$, $\overline{v} + v'$ $\overline{w} + w'$ and $\overline{T} + T'$. An important problem in data analysis is deciding how to determine and remove the background state. One way is to high-pass filtered in either wavenumber or frequency space. This requires a judgment as to what constitutes a suitable cut-off wavenumber or frequency to remove non-wave components. Figure 1 illustrates the process for GPS/MET observations of temperature. Sometimes it is more convenient to fit and subtract a low-order polynomial, such as often used with radiosonde observations (e.g. Allen and Vincent, 1995; Vincent et al., 1997). Whatever the situation, some degree of bias will be introduced into the wave determination due to the choice of filter or polynomial (e.g. Vincent et al., 1997). Figure 2 illustrates the process of obtaining wind perturbations from radiosonde data using a third-order polynomial to extract the background winds.

After the perturbation or fluctuating components have been obtained the wave energy per unit mass can be computed:

$$E = \frac{1}{2}\left(\overline{u'^2} + \overline{v'^2} + +\overline{w'^2} + \frac{g^2\overline{\hat{T}'^2}}{N^2}\right) \tag{1}$$

where $\hat{T}' = T'/\overline{T}$ is the perturbation temperature normalized to the background temperature, g is the acceleration due to gravity and N is the buoyancy frequency. In (1) terms involving velocity components constitute the kinetic energy (KE) and the term involving T is the potential energy (PE). In practise w' is much smaller than either u' or v', so $\overline{w'^2}$ can be neglected in computations of the KE.

A time series of monthly mean wave energy is shown in Figure 3. Derived from six years of routine radiosonde data taken in the Cocos Islands (12°S, 97°E) it shows an annual cycle in energy, with maximum values observed at the time of year when convection is strongest. There is also strong interannual variabilty, which is related to the phase of the QBO in mean winds. This raises the question: Are the annual and QBO-like variations due to sources or wind-filtering effects? Vincent and Alexander (2000) and Alexander and Vincent (2000) discuss these issues further.

Radiosondes provide much of the basic information used for studies of gravity waves in equatorial troposphere and lower stratosphere. If high-vertical resolution wind and temperature data are available then they can be used to develop climatolo-

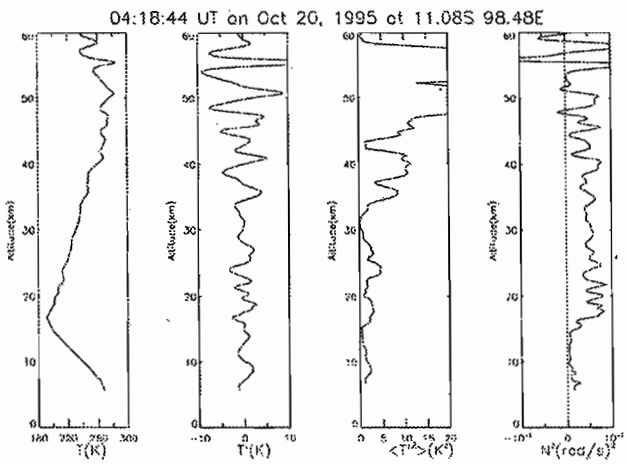


Figure 1: GPS/MET temperature profiles obtained at (11°S, 99°E), October 20, 1995. The panels show left to right the temperature profile, the temperature fluctuations after the high-pass filtering with a cut-off of 10km, temperature variance in 2km layers, and the square of the buoyancy frequency, N^2 (after Tsuda et al., 2000)

gies of wave activity. Routine soundings made by weather services are an important resource currently used in an international project coordinated by the WCRP project *Stratospheric Processes and their Role in Climate* (SPARC) to develop a global climatology of gravity wave parameters.

Where temperature-only data are available then the potential energy (PE) can be used to assess wave amplitudes. Allen and Vincent (1995) used high resolution radiosonde data to develop a wave climatology in the troposphere and stratosphere for the Australian region. Tsuda et al (2000) derived a more global PE climatology in the stratosphere from temperatures acquired with GPS occultation observations, as shown in Figure 4. Temperature perturbations are largest at low latitudes and seem to be linked to tropical convection sources.

3 Propagation Directions

Gravity waves may be propagating in preferred directions in both the vertical and horizontal. Wave sources may radiate waves in anisotropically, while waves propagating through a background that changes with height may be removed at critical levels where their phase speed equals the background wind (wind-filtering). It is important to quantify the degree of anisotropy.

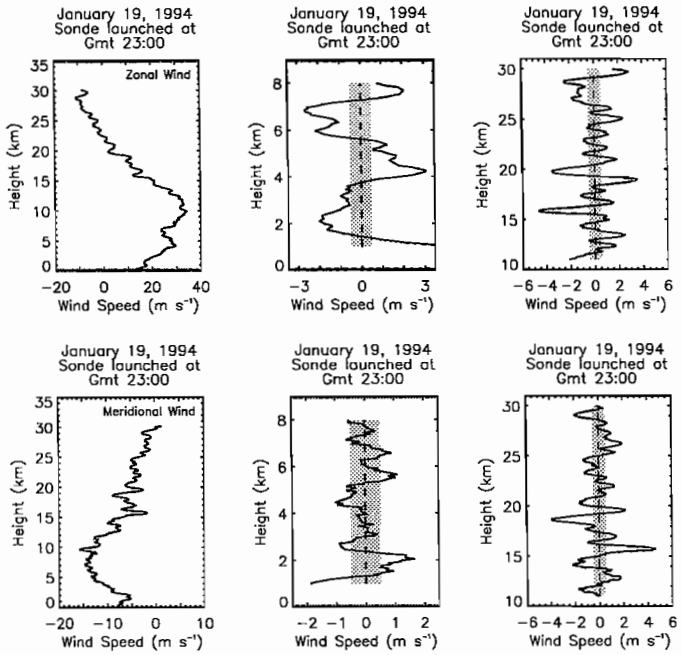


Figure 2: Zonal (top) and meridional (bottom) wind speed and fluctuations observed by radiosonde over Macquarie Island (55^{circ}S). Data from height ranges in the troposphere (middle) and stratosphere (right) are shown. The background values were determined by fitting third-order polynomials in the 2-9 km and 17-24 km height ranges (After Allen, 1995).

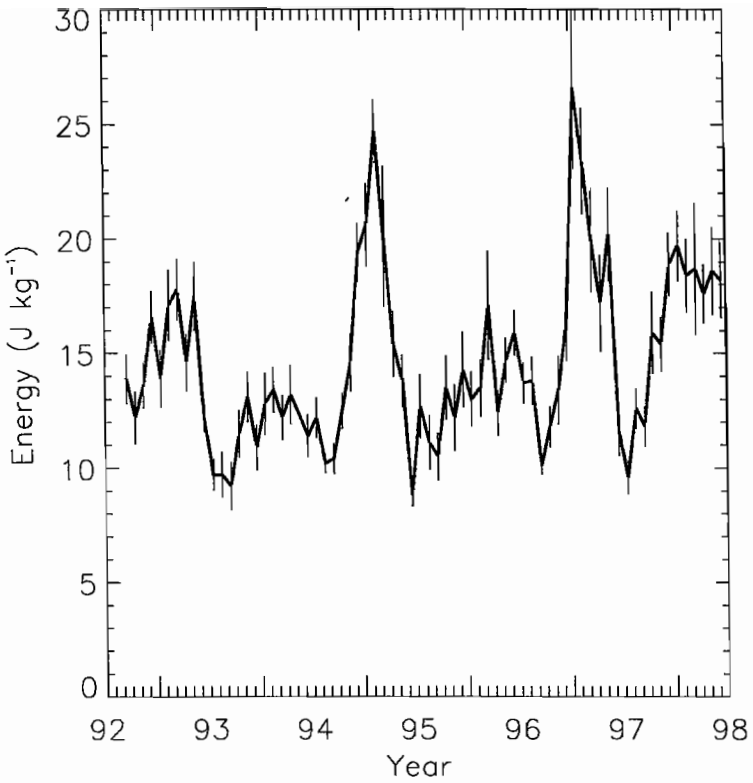


Figure 3: Time series of monthly-mean wave energy per unit mass averaged over the 18-25 km height range (after Vincent and Alexander, 2000).

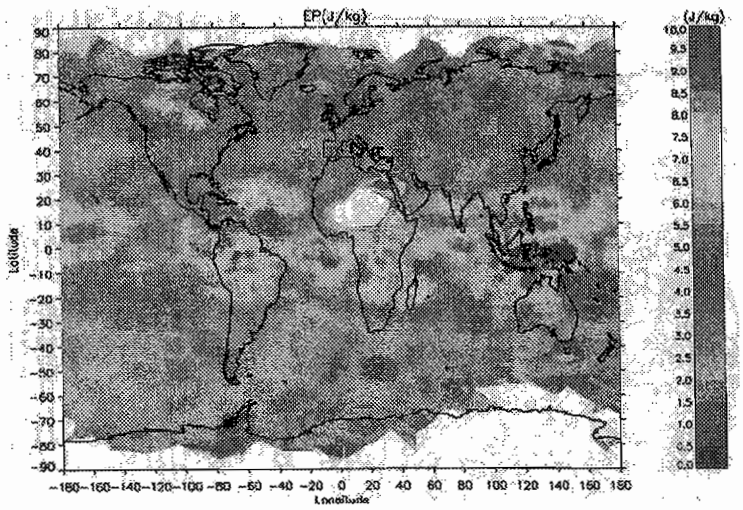


Figure 4: Global distribution of PE (after Tsuda et al., 2000)

Consider the polarization relations that relate the various fluctuations induced by an idealized monochromatic gravity wave propagating in an isothermal atmosphere with constant wind. If fluctuations are harmonic in time and space $u', v', w', T' \sim e^{i(kx+ly+mz-\hat{\omega}t)}$, then v' is expressed in terms of u' as

$$v' = \frac{\hat{\omega}l - ifk}{\hat{\omega}k + ifl}u'. \tag{2}$$

Here f is the local inertial frequency and $\hat{\omega}$ is the intrinsic frequency, the wave frequency measured in a reference frame moving with the background wind. The vertical wavenumber m is related to the horizontal wavenumbers k, l, through the non-hydrostatic dispersion relation

$$m^2 = \frac{N^2 - \hat{\omega}^2}{\hat{\omega}^2 - f^2}k_h^2 - \frac{1}{4H^2} + \frac{\hat{\omega}^2}{c_s^2} \tag{3}$$

where $k_h^2 = k^2 + l^2$, H is the density scale height and c_s is the speed of sound. Usually, it is possible make the so-called Boussinesq approximation ($m > 1/2H$) and ignore the last two terms in (3), so that

$$m = \pm\sqrt{\frac{N^2 - \hat{\omega}^2}{\hat{\omega}^2 - f^2}}k_h. \tag{4}$$

Downward phase velocity ($m < 0$) is associated with upward group velocity and hence upward energy propagation.

The perturbation wind vector is elliptically polarized. In the limit as $\hat{\omega} \to f$ the polarization becomes circular and as $\hat{\omega} \to N$ it becomes linear. For waves propagating energy upward the vector rotates in an anticyclonic manner with increasing height i.e. clockwise in the northern hemisphere ($f > 0$) and anticlockwise in the southern hemisphere. Figure 5a illustrates the hodograph of a wave with $\hat{\omega} = 2f$.

3.1 Vertical energy propagation

A simple way to determine the mean direction of propagation is to examine the sense of rotation of hodographs after the mean winds have been removed. Where many observations are available the "rotary decomposition" technique provides a more statistical approach (e.g. Vincent, 1984). Defining the complex wind vector as

$$V'(z) = u'(z) + iv'(z) = ae^{-imz} \tag{5}$$

and Fourier transforming $V'(z)$ gives $\mathcal{A}(m)$ in vertical wavenumber space. $\mathcal{A}(m)$ can be decomposed into the clockwise (C) and anticlockwise (A) components by forming:

$$C(m) = \frac{1}{2}\overline{\mathcal{A}(-m)\mathcal{A}^*(-m)} \tag{6}$$

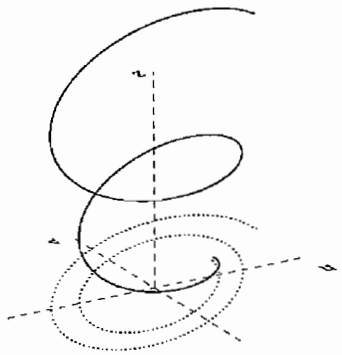


Figure 5: Perturbation wind profile as a function of height (solid line) for an idealized gravity wave with $\hat{\omega} = 2f$. The dotted line is the projection of the hodograph onto the horizontal plane.

and

$$A(m) = \frac{1}{2}\overline{\mathcal{A}(m)\mathcal{A}^*(m)} \tag{7}$$

where the overbar means an average over all spectra and * is the complex conjugate. The total energy is

$$E(m) = A(m) + C(m) \tag{8}$$

Winds associated with upward propagating waves rotate in an anticyclonic sense with increasing height, that is clockwise (anticlockwise) in the northern (southern) hemisphere. Hence the ratio $r_n(m) = C(m)/E(m)$ ($r_s = A(m)/E(m)$) gives the sense of vertical propagation at a particular wavenumber in the northern (southern) hemisphere. A value greater than 0.5 indicates more energy is going up than down.

Because m depends on the intrinsic frequency, and so can change in the presence of vertical wind shear, the best quantity to use is the ratio averaged over all wavenumbers, i.e. $r = \overline{C(m)}/\overline{E(m)}$. Radiosonde and rocket measurements all show that energy is primarily upward going in the stratosphere with r ~70-80%, but equally partitioned ~50% in the troposphere, which shows that the main wave sources are in the troposphere.

3.2 Horizontal propagation directions

Polarization relation (2) is simplified by rotating the coordinated system to align the axes with the major and minor axes of the perturbation ellipse. If $u'_{||}$ and $v'_{\perp}$ are the

horizontal perturbation velocities parallel and perpendicular to the wave vector then (2) becomes

$$v'_{\perp} = -i\frac{f}{\hat{\omega}}u'_{||} \tag{9}$$

This expression shows that the intrinsic frequency $\hat{\omega}$ can be determined in terms of f from the axial ratio of the hodograph.

Furthermore, the horizontal direction of propagation of the wave is aligned along the major axis. In Figure 5, for example, the major axis is aligned in the NE/SW quadrants, so inspection of the hodograph gives the propagation direction, but with a 180° ambiguity.

In practice, it is not quite so simple to deduce the intrinsic frequency and direction of propagation from the hodograph. A spectrum of waves will generally be present, each with a with different azimuth of propagation, and the background wind will change with height, changing the intrinsic frequency. A more quantitative, statistical, description of gravity wave motions is obtained by making an analogy between partially polarized gravity wave and electromagnetic wave fields. Vincent and Fritts (1987) introduced the concept of Stokes parameters, which allow the degree of polarization and alignment of the polarized component to be determined. Initially the technique was applied to time series of wind motions. Eckermann and Vincent (1989) extended the idea to vertical wind profiles where in vertical wavenumber space. The Stokes parameters are defined as:

$$I = \overline{u'^2} + \overline{v'^2} \tag{10}$$

$$D = \overline{u'^2} - \overline{v'^2} \tag{11}$$

$$P = 2\overline{u'v'cos\delta} \tag{12}$$

$$Q = 2\overline{u'v'sin\delta} \tag{13}$$

The parameters I, D, P, and Q have simple interpretations. I is the total energy in the spectrum, while D and P are measures of the anisotropy of the field. The parameter Q gives the sense of rotation of the polarization ellipse (positive is clockwise and negative is anticlockwise). With these parameters it is easy to calculate the phase

difference, δ, major axis orientation, ϕ, and ellipse axial ratio, AR, of the elliptically polarized wavefield:

$$\delta = tan^{-1}\left(\frac{Q}{P}\right) \tag{14}$$

$$2\phi = tan^{-1}\left(\frac{P}{D}\right) \tag{15}$$

$$AR = cot\xi \quad \text{where} \quad 2\xi = sin^{-1}\left(\frac{Q}{dI}\right) \tag{16}$$

$$d = \frac{(D^2 + P^2 + Q^2)^{1/2}}{I}. \tag{17}$$

The quantity d is the degree of polarization, which indicates the ratio of polarized to irregular motions. Note, sufficient averaging in either frequency or wavenumber space is necessary in order to ensure that estimates of d are statistically reliable (Eckermann and Vincent, 1989). Values of d range between 0 (unpolarized) to 1 (polarized); intermediate values imply a partially polarized wave field. Radiosonde soundings in the stratosphere show values of d ~0.5 or larger (Vincent and Alexander, 2000).

Additional information is required to resolve the 180° directional ambiguity. This can be achieved if w' is measured. The polarization relation linking w' to $u'_{||}$ is

$$w' = -\frac{k_h}{m}u'_{||} \tag{18}$$

The upward flux of horizontal momentum $\overline{u'_{||}w'} = -(k_h/m)\overline{u'^2_{||}}$ has the same sign as k_h for upward propagating waves ($m < 0$). In component terms, the relevant fluxes are $\overline{u'w'}$ and $\overline{v'w'}$.

Specialized radar techniques allow momentum fluxes to be measured directly so that directions can be ascertained. Normally, however, it is neither possible to measure momentum fluxes directly, nor to compute them indirectly due to difficulties in measuring the small values of w'. An alternative approach is to use temperature as a surrogate for w'. Kitamura and Hirota (1989), Hamilton (1991), and Vincent et al (1997) showed how the relative phase of temperature and horizontal wind perturbations can be used to determine the sense of horizontal phase progression. $\hat{T}'$ is related to $u'_{||}$ via

$$\hat{T}' = \mp i\frac{N^2}{g\hat{\omega}}\frac{k_h}{m}u'_{||}, \tag{19}$$

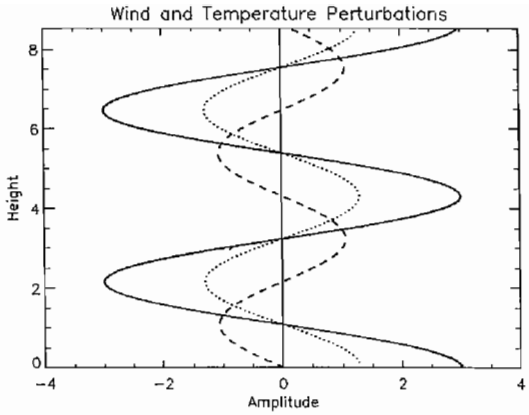


Figure 6: Vertical profiles of $u'_{||}$ (solid), w' (dotted), and T' (dashed) for a wave propagating in the postive x direction and upward.

where the negative sign is chosen when $km > 0$ and the positive sign when $km < 0$. The $i = \sqrt{-1}$ factor means that the horizontal motions are 90° out of phase with the temperature perturbations. Figure 6 shows the relationships between the three quantities for the case where $k_h > 0$ so that u' and w' are in phase (i.e. $\overline{u'_{||}w'} > 0$). Coldest temperatures lag a quarter cycle after maximum upward (positive) velocities.

Rearranging (18) and (19) gives

$$\overline{u'w'} = \frac{\hat{\omega}g}{N^2}\overline{u'_{||}\hat{T}'_{+90}} \tag{20}$$

where $\hat{T}'_{+90} = i\hat{T}'$ is obtained by first Hilbert transforming $\hat{T}'$ in the appropriate sense and then reverse transforming to produce a 90° phase shift in $\hat{T}'$ at all spatial frequencies without changing the amplitude (Vincent et al., 1997). Alternatively, the quadrature spectrum between $u'_{||}$ and $\hat{T}'$ can be used; where velocity components are available then $u'\hat{T}'_{+90}$ and $v'\hat{T}'_{+90}$ are the relevant quantities. Examples of the distributions of horizontal propagation directions (angular spectra) determined by this technique from radiosonde observations in the tropical lower stratosphere are shown in Figure 7. The directions are predominantly eastward, especially during the wet season (strongest convection) from December to July.

While it is useful to use the sign of (20) to obtain the direction of propagation it is less useful for determining momentum fluxes, since $\hat{\omega}$ is required. There is no unambiguous way of determining this quantity from observations alone because it is not always clear whether the fluctuations measured by radiosondes and rockets, for example, are due to a wide spectrum of waves or are dominated by a quasi-monochromatic waves. This issue is explored further by Vincent and Alexander

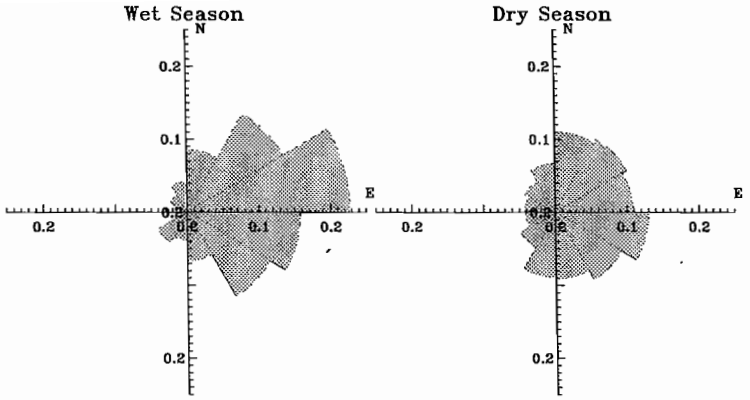


Figure 7: Angular spectrum of gravity waves observed in the 18-25 km height range over Cocos Island (after Vincent and Alexander, 2000)

(2000) and Alexander and Vincent (2000).

4 Momentum Fluxes

As gravity waves propagate into the atmosphere their amplitudes grow as $\sim exp(-z/2H)$ in order to conserve energy. For conservative motions the momentum flux $\rho\overline{u'w'}$ remains constant. Where the flux changes there is an associated body force, which in equatorial regions where the Coriolis force is low, leads to an acceleration of the zonal flow. In order to understand and quantify the role of gravity waves it is important to measure momentum fluxes and their change with height. The techniques using temperature described above are not very satisfactory. First, there is the requirement to know the value of $\hat{\omega}$ used in (20). Secondly, there is the issue of observational selection or filtering in which techniques can only view or select parts of the wave spectrum. This issue is discussed further in section 6.

A direct way of measuring momentum fluxes is to use Doppler radars. Figure 8 shows the configuration which uses two narrow beams. One beam points at an angle θ to the zenith and the other points in the complementary direction, $-\theta$. The radial velocity along the first beam at range R is

$$v_r(\theta, R) = u'sin\theta + w'cos\theta \tag{21}$$

and for the complementary beam

$$v_r(-\theta, R) = u'sin\theta - w'cos\theta \tag{22}$$

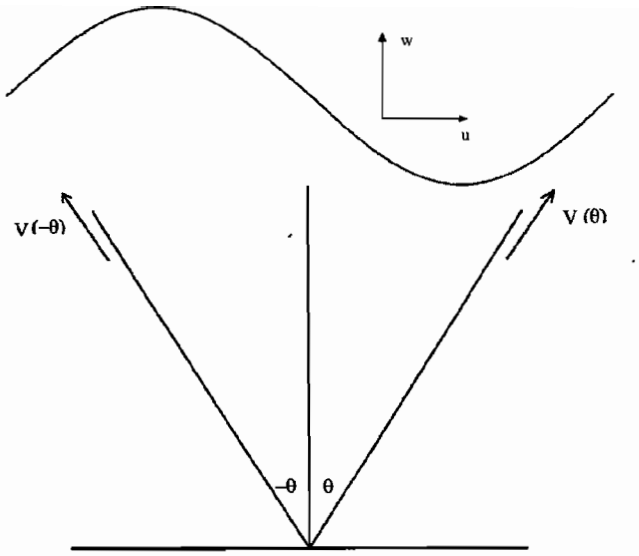


Figure 8: Configuration of the complementary dual-beam radar technique for measuring momentum fluxes (after Vincent and Reid, 1983).

Squaring and adding equations (21) and (22) and rearranging gives (Vincent and Reid, 1983)

$$\overline{u'w'}(z) = \frac{\overline{v_r^2(\theta, R)} - \overline{v_r^2(-\theta, R)}}{2sin2\theta} \tag{23}$$

so that the momentum flux can be measured at different heights, z. By steering beams in the EW and NS plans it is possible to measure both $\overline{u'w'}$ and $\overline{v'w'}$. This has proved a robust technique, but there some requirements that restrict its use. One problem is that the fluxes are derived from small differences in radial velocities, which requires very accurate measurements.

Another important factor is the need to know the pointing angles accurately, so narrow beams are required. This means large radars. At MF (frequency $\sim$ 2 MHz) the only radar with the capability to generate narrow beams is the 1-km diameter array at Buckland Park near Adelaide. At VHF the largest and most powerful MST radars can also exploit the technique. These include the MU radar in Japan, the Jicamarca radar in Peru and the MST radar in Tirupati, India.

Despite these limitations the technique has been used to test and verify predictions about wave mean-flow interactions, including interaction between gravity waves and tides (Fritts and Vincent, 1987) and wave-driven circulations. So far the technique has had limited use for studying wave mean flow interactions at the equator, where gravity waves are thought to play a major role in forcing the QBO and mesospheric

SAO. In order to explore the role of gravity waves in driving circulations at the equator a new MF radar is under construction at Pontianak in Indonesian West Kalimantan (0°, 109°E). This is a region of strong convection, probably the major source of gravity waves in the tropics.

5 Gravity Wave Spectra

A convenient way to summarize information about gravity waves is by computing spectra of their observed parameters. Spectra in frequency and wavenumber space show approximately invariant forms, despite the exponential decrease of density with height and regardless of season or geographic location. It was originally argued by VanZandt (1982) that there was a "universal" spectrum of atmospheric gravity waves, in analogy with internal wave spectra in the oceans. There has been vigorous debate about the mechanisms that constrain the wave motions and several "saturation" theories have been proposed. There is also debate as to whether the spectra have "universal" amplitudes (to within a factor of two) at all locations, although observations do show some degree of universality of slope.

5.1 Frequency spectra

Long time sequences of winds measured by radars lend themselves to ease of computation of frequency frequency. Irrespective of whether observations are made in the lower or upper atmosphere the spectra have a common form $S(\omega) \sim \omega^{-p}$, where ω is the observed frequency measured in a ground-based frame. The spectral index is usually $p \sim 3/2 - 5/3$. Spectra derived from winds measured at the mesospause ($\sim$ 85 km) with MF radars situated at the equator show a characteristic $\omega^{-5/3}$ shape, at least at periods shorter than about 1 day (Figure 9). At longer periods the motions are dominated by the effects of planetary-scale waves (Kovalam et al, 1999).

A complication with frequency spectra derived from ground-based observations is that they are determined in terms of ω and not intrinsic fequency, $\hat{\omega}$. Recent observations made with constant pressure balloons floating at almost constant heights near 20 km do not suffer from this limitation, however. The balloons, drifting under the influence of the mean winds and waves, act as Lagrangian tracers so that the horizontal and vertical motions are measured in intrinsic coordinates. Spectra computed from long-term (several day) observations at both the equator and Arctic show $S(\hat{\omega}) \sim \hat{\omega}^{-2}$ (Hertzog and Vial, 2001).

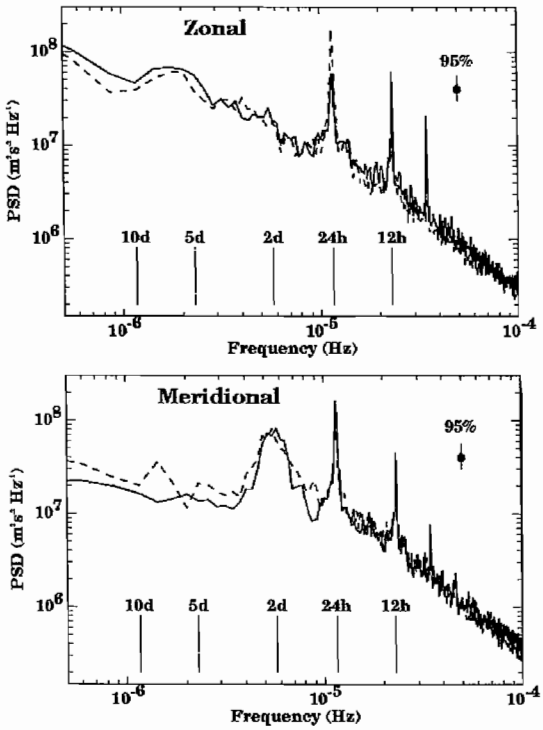


Figure 9: Frequency spectra of hourly averaged zonal (top) and meridional winds (bottom) averaged between 86 and 92 km at Christmas Island (2°N, 157°W) (solid line) and Pontianak (0°N, 109°E) (dashed line) for the period November 1995 to August 1997. Confidence limits at the 95% level are indicated (After Kovalam et al., 1999).

5.2 Vertical wavenumber spectra

Studies made with many different techniques over heights ranging from the troposphere to the lower thermosphere give broadly similar descriptions of vertical wavenumber spectra. At high wavenumbers they show spectra of the form $F(m) \sim m^{-t}$ with t in the range 2 to 3 (e.g. Smith et al., 1987). There have been many different attempts to provide a theoretical basis for this functional form based on a saturation process that limits wave growth. For example, Dewan and Good (1986) proposed a saturation limit imposed by convective or dynamical instability and argued that saturated waves should have spectra with an N^2/m^3 dependence.

At low wavenumbers (long vertical wavelengths) the waves need not be saturated and can grow with height. The transition between between saturated and unsaturated waves occurs at some characteristic wavenumber m_*. Smith et al (1987) gave a simple function form for the m spectrum as

$$F(\mu) = \frac{\mu^2}{1 + \mu^{s+t}} \tag{24}$$

where $\mu = m/m_*$ is a constant and $t = 3$. The spectral index for the low frequency part of the spectrum is difficult to measure, but is often taken as $s = 0$. This kind of spectrum is often referred to as the modified-Desaubies spectrum. A diagram encapsulating many different data sets taken in different regions of the atmosphere is shown in Figure 10. Based on radar observations in the troposphere and lower stratosphere, radiosonde, rocketsonde, GPS/MET and lidar observations in the stratosphere, and radar and lidar observations in the mesosphere the characteristic vertical wavelength $\lambda_* = 2\pi/m_*$ is found to increase from about 2 km in the troposphere to about 16-20 km at the mesopause. Some observations, notably lidar observations in the middle stratosphere, show amplitudes that are below saturation limits. Eckermann (1995) discuss the physical processes that can lead to these attenuated spectra.

6 Observational Selection

As waves propagate in a height-varying background wind the intrinsic frequency $\hat{\omega} = \omega - k_h \overline{u}$ will change, which impacts on which waves can be observed. To better understand the process it is useful to use a simplified form of the dispersion relation

$$|m| \sim \frac{N}{\hat{\omega}} k_h \sim \frac{N}{|\overline{u} - c|} \tag{25}$$

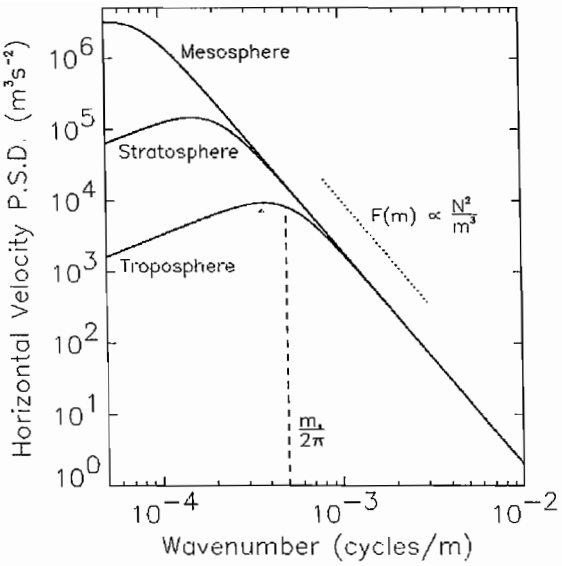


Figure 10: The modified-Desaubies spectrum of horizontal wind fluctuations in different atmopsheric regions (after Smith et al., 1987).

where c is the ground-based wave frequency. Figure 11 taken from Alexander (1998) helps with visualizing how Doppler shifting can effect m, depending on whether a wave is propagating against or with the mean flow. Eckermann (1995) and Alexander (1998) discuss the importance of the Doppler-shifting effect when an instrument or technique can only respond to waves in a certain range of m or vertical wavelength. This is equivalent to an observational filter, as exemplified in Figure 12 for radiosondes, rockets and the microwave limb sounder (MLS) instrument on UARS. At one extreme radiosondes "see" short wavelength waves, while MLS is sensitive to waves with long vertical wavelengths. Waves will move in or out of the observational window in height regions where there are strong shears in background wind, so observed changes in wave amplitude may be more apparent than real. Ground-based instruments, such as radars and lidars, will suffer a similar effect as changing background winds Doppler-shift wave energy in frequency (Fritts and VanZandt, 1987). Understanding how observational selection affects gravity wave studies requires careful analysis and modeling; all techniques are affected.

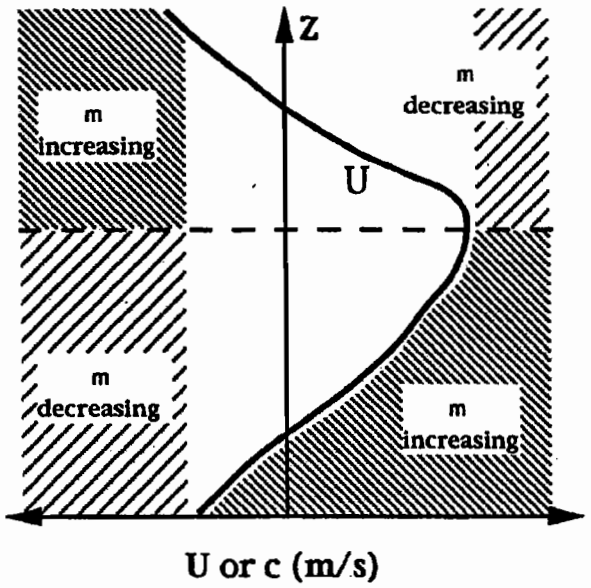


Figure 11: Schematic diagram showing the Doppler shifting effects of the mean wind on vertical wavenumber m (after Alexander, 1998).

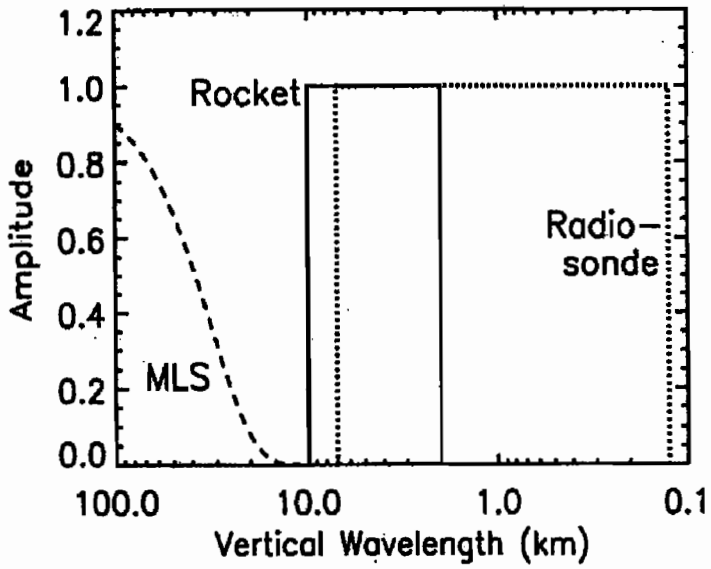


Figure 12: Observational filter functions showing fractional amplitude as a function of vertical wavelength for rocket sounding (solid), radiosondes (dotted) and Microwave Limb Sounder (MLS) (after Alexander, 1998).

7 References

Alexander, M.J., 1998: Interpretations of observed climatalogical patterns in stratospheric gravity wave variance, *J. Geophys. Res.*, **103**, 8627-8640.

Alexander, M.J., and R.A. Vincent, 2000: Gravity waves in the tropical lower stratosphere: A model study of seasonal and interannual variability, *J. Geophys. Res.*, **105**, 17,971-17,982.

Allen, S.J., and R.A. Vincent, 1995: Gravity wave activity in the lower atmosphere: Seasonal and latitudinal variations, *J. Geophys. Res.*, **100**, 1327-1350.

Dewan, E.M., and R.E. Good, 1986: Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere, *J. Geophys. Res.*, **91**, 2742-2748.

Eckermann, S.D., 1995: Effect of background winds on vertical wavenumber spectra of atmospheric gravity waves, *J. Geophys. Res.*, **100**, 14,097-14,112.

Eckermann, S.D., and R.A. Vincent, 1989: Falling sphere observations of anisotropic gravity wave motions in the upper stratosphere over Australia, *Pure and Applied Geophys.*, **130**, 510-532.

Fritts, D.C., and R.A. Vincent, 1987: Mesospheric momentum flux studies at Adelaide, Australia: Observations and a gravity wave/tidal interaction model, *J. Atmos. Sci.*, **44**, 605-619.

Gardner, C.S., and W. Yang, 1998: Measurements of the dynamical cooling rate associated with the vertical transport of heat by dissipating gravity waves in the mesopause region at the Starfire Optical Range, New Mexico, *J. Geophys. Res.*, **103**, 16,909-16,926.

Hamilton, K., 1991: Climatological statistics of stratospheric inertia-gravity waves deduced from historical rocketsonde wind and temperature data, *J. Geophys. Res.*, **96**, 20,831-20,839.

Hertzog, A. and F. Vial, 2001: A study of the dynamics of he equatorial lower atmosphere by use of ultra-long-duration balloons 2. Gravity waves, *J. Geophys.*

Res., **106**, 22,745-22,761.

Kitamura, Y., and I. Hirota, 1989: Small-scale disturbances in the lower stratosphere revealed by daily rawin sonde observations, *J. Meteorol. Soc. Japan*, **67**, 817-831.

Kovalam, S., R.A. Vincent, I.M. Reid, T. Tsuda, T. Nakamura, A. Nuryanto, and H. Wiryosumarto, 1999: Longitudinal variations in planetary wave activity in the equatorial mesosphere, *Earth, Planet. Sci.*, **51**, 665-674.

Smith, S.A., D.C. Fritts, and T.E. VanZandt, 1987: Evidence for a saturated spectrum of gravity waves, *J. Atmos. Sci.*, **44**, 1404-1410.

Tsuda, T., M. Nishida, C. Rocken, and R. Ware, 2000: Global Morphology of Gravity Wave Activity in the Stratosphere Revealed by the GPS Occultation Data (GPS/MET), *J. Geophys. Res.*, **105**, 7257-7274.

VanZandt, T.E., 1982: A universal spectrum of buoyancy waves in the atmosphere, *Geophys. Res. Lett.*, **9**, 575-578.

Vincent, R.A., 1984: Gravity wave motions in the mesosphere, *J. Atmos. Terr. Phys.*, **46**, 119-128.

Vincent, R.A., and I.M. Reid, 1983: HF Doppler measurements of mesospheric gravity wave momentum fluxes, *J. Atmos. Sci.*, **40**, 1321-1333.

Vincent, R.A., and D.C. Fritts, 1987: A climatology of gravity wave motions in the mesopause region at Adelaide, Australia, *J. Atmos. Sci.*, **44**, 1321-1333.

Vincent, R.A., S.J. Allen, and S.D. Eckermann, 1997: Gravity-wave parameters in the lower stratosphere, in *Gravity wave processes: Their parameterization in global climate models*, (ed. K. Hamilton), pp. 7-25, Springer-Verlag, Berlin.

Vincent, R.A., and M.J. Alexander, 2000: Gravity waves in the tropical lower stratosphere: An observational study of seasonal and interannual variability, *J. Geophys. Res.*, **105**, 17,971-17,982.

Chapter 5

Mean Meridional Circulation and Transport

Lesley Gray,

Rutherford Appleton Laboratory, U.K.

Mean Meridional Circulation and Transport

1. Background

We have seen that the semi annual oscillation (SAO) and the quasi-biennial oscillation (QBO) in zonal winds are the result of the interaction of vertically propagating waves with the background flow. There is also a clear SAO / QBO in temperature. The positive correlation between temperature anomalies averaged over the 30-50 hPa layer and the vertical shear in zonal wind over the same layer, taken from Singapore balloon observations (figure 1) shows that the tropical temperature QBO is in thermal wind balance with the vertical shears of the zonal winds i.e. $\partial u/\partial z$ is proportional to $-\partial T/\partial y$

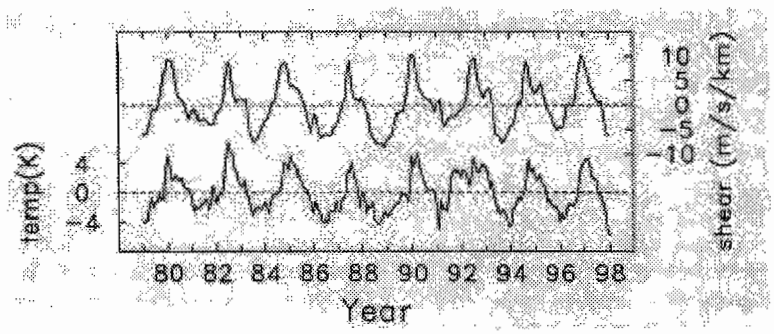


Figure 1 Equatorial temperature anomalies associated with QBO in the 30 – 50 – hPa layer (bottom curve) and vertical wind shear (top curve). *Reproduced with permission from Baldwin et al. (2001).*

A westerly vertical shear zone is therefore associated with a positive temperature anomaly at the equator and an easterly vertical shear is associated with a negative temperature anomaly. In westerly shear zones the maximum in temperature at the equator is maintained against thermal damping by adiabatic warming due to sinking motion. The opposite holds in easterly shear zones. The pattern of shear zones and meridional circulations is shown schematically in figure 2.

The presence of a warm temperature anomaly at the equator, associated with westerly vertical shear, also gives rise to anomalous radiative cooling, leading to mean sinking (or at least anomalously weak tropical upwelling, since the background flow is strongly upwelling). There may also be a significant radiative feedback of the QBO anomaly in ozone on the induced meridional circulation, since changes in ozone have a direct effect on short-wave heating.

This circulation with rising or sinking at the equator is compensated by an opposing circulation off the equator, which gives rise to (much smaller) anomalies in the subtropics and midlatitudes of opposite sign to that at the equator (see fig. 2).

Figures 3-5 show the latitude time-series at 32 hPa (approx. 24 km) of QBO anomalies in zonal wind, temperature and (residual mean) vertical velocities from meteorological analyses (Randel et al. 1999). The equatorial zonal wind anomalies (figure 3) extend

between 20°S – 20°N. Figures 3-5 show the latitude time-series at 32 hPa (approx. 24 km) of QBO anomalies in zonal wind, temperature and (residual mean) vertical velocities from meteorological analyses (Randel et al. 1999). The equatorial zonal wind anomalies (figure 3) extend between 20°S – 20°N. The temperature anomalies (figure 4) are in approximate thermal wind balance with the wind anomalies and the timing is determined by the vertical wind shear. There is also evidence of a phase change at around 15° latitude and a weak subtropical anomaly. The north polar signal, which is large (but not so significant statistically due to the large variability in this region) is associated with a QBO modulation of the amplitude and timing of NH stratospheric sudden warmings. The vertical velocities (figure 5) are consistent with the meridional circulation of figure 2, showing a phase change in the subtropics and a smaller return arm to the circulation in the subtropics. Note that the timing of the anomalies is different at 3 hPa (around 40 km) and 32 hPa (around 24 km), reflecting the downward propagation of the westerly and easterly shear zones.

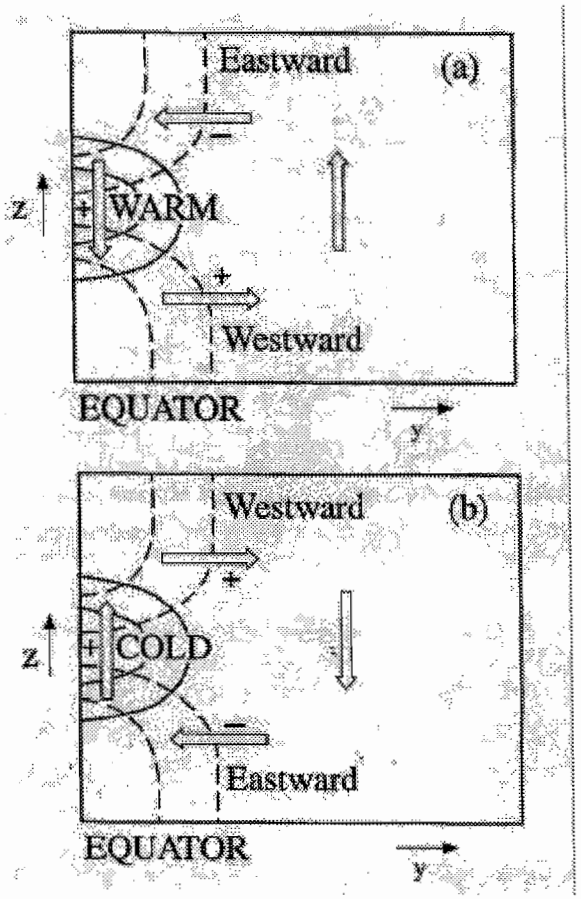


Figure 2 Schematic latitude-height sections showing the mean meridional circulation associated with the equatorial temperature anomaly of the QBO. Solid contours show temperature anomaly isotherms, and dashed contours are zonal wind isopleths. Plus and minus signs designate signs of zonal wind accelerations driven by the mean meridional circulation. (a) Westerly shear zone. (b) Easterly shear zone. After *Plumb and Bell* [1982b]. *Reproduced with permission from Baldwin et al. (2001).*

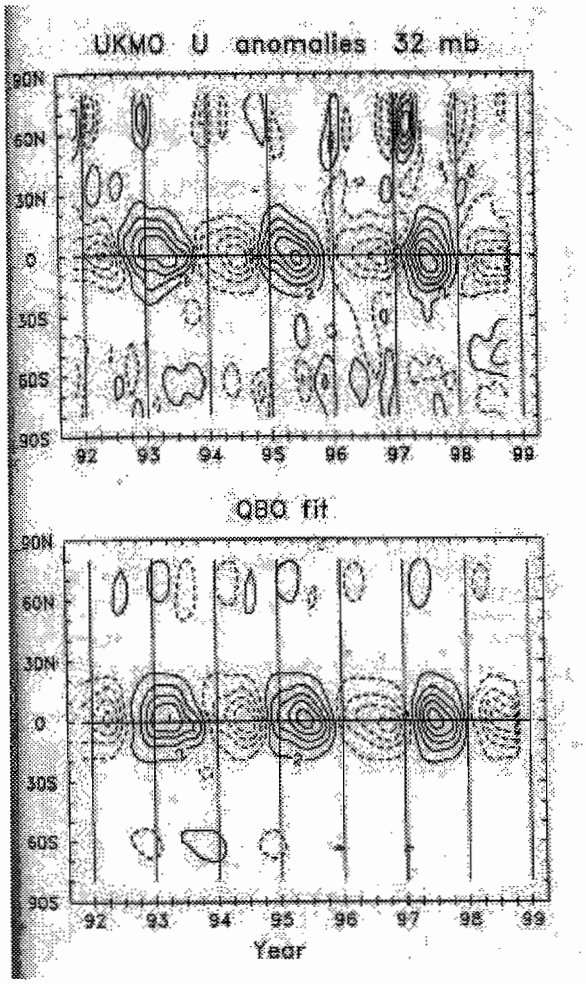


Figure 3 Latitude-time sections of UKMO zonal wind anomalies at 32mb. Top panel shows the full anomalies and bottom panel is the statistical QBO fit. Contours are ± 2, 6, 10 .... m s^{-1}
Reproduced with permission from Randel et al. (1999).

The QBO-induced circulation advects relatively long-lived tracers such methane (CH_4), nitrous oxide (N_2O), water vapour (H_2O) and lower stratospheric ozone (O_3), giving rise to distinctive patterns of anomalous distributions of tracers associated both with the SAO (section 3) and the QBO (sections 4,6).

In addition to the influence on longer-lived tracer chemicals, there is also an influence on shorter-lived chemicals via their temperature dependence (section 7). This can result in substantial percentage variations from year to year e.g. in nitric acid (HNO_3), nitrogen dioxide (NO_2) and dinitrogen pentoxide (N_2O_5), associated with the QBO.

A significant inter-hemispheric asymmetry in the timing and amplitude of the subtropical anomaly in temperature and trace gas distributions is also evident. This is due to the interaction of the QBO / SAO induced circulation with the annual cycle (section 6.2).

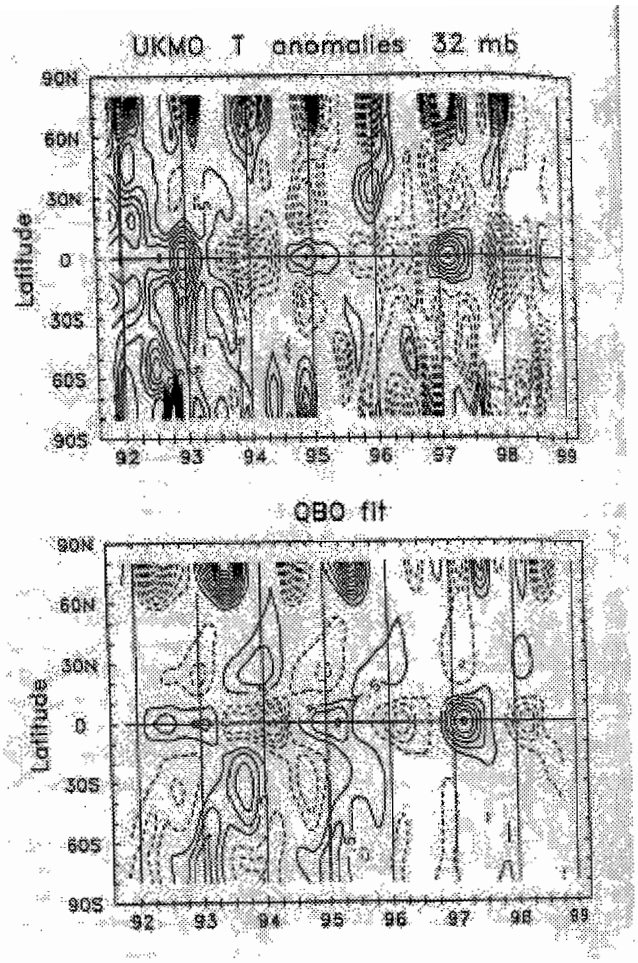


Figure 4 Altitude-time section of UKMO zonal-mean equatorial temperature anomalies. Shown are the full anomalies (top) and the QBO statistical fit (bottom). Contours are +- 0.5, 1.0, 1.5 .... K.
Reproduced with permission from Randel et al. (1999).

Plumb and Bell (1982) also noted that the advective effects of the meridional circulation can account for the observed asymmetry in the descent of the equatorial QBO easterlies and westerlies. Downward advection of momentum associated with the westerly shear zone enhances the descent of the westerlies, while the upward advection of momentum associated with easterly shear inhibits the descent of easterlies.

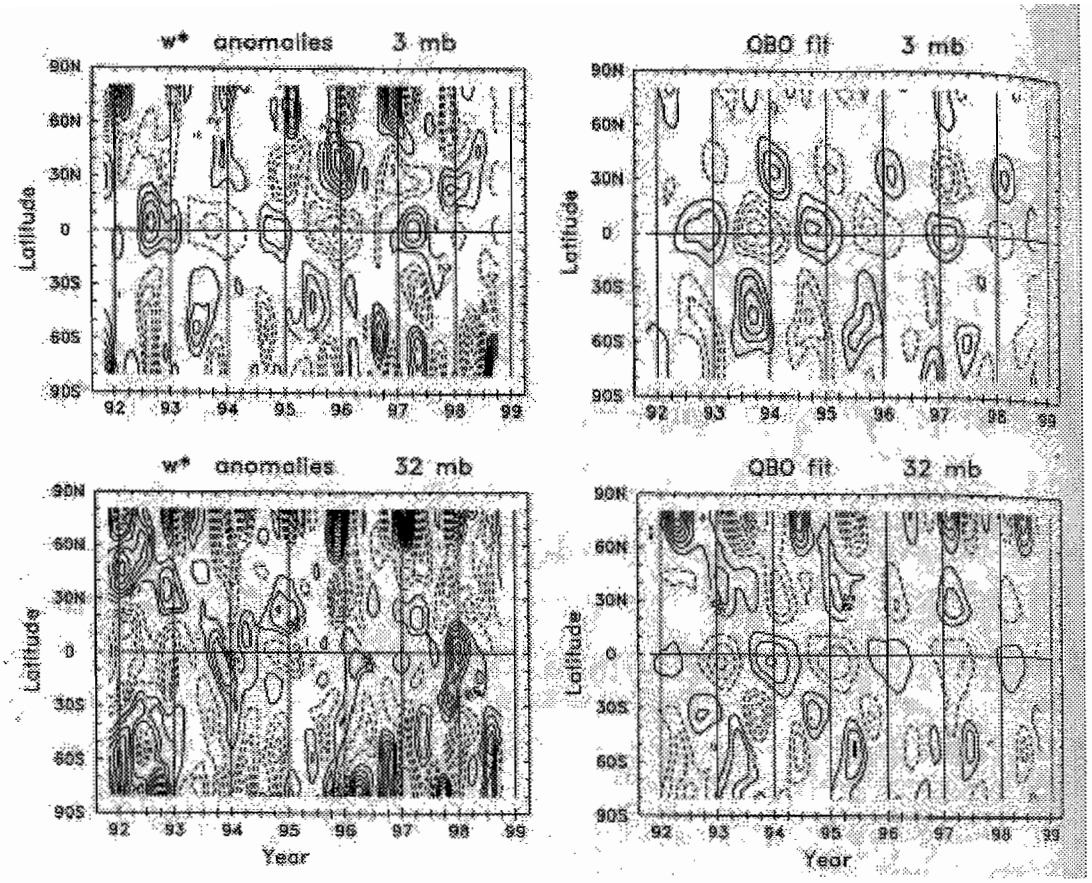


Figure 5 Latitude-time sections of residual mean vertical velocity (w*) interannual anomalies at 3.2 mb (top) and 32 mb (bottom). Left: the full anomalies. Right: the QBO fit. Contour interval is 2.0 (top) and 0.05 km/month (bottom), with zero contours omitted.

Reproduced with permission from Randel et al. (1999).

2. Long Lived Chemical Tracers

We can learn a lot about the mass flow in the middle atmosphere from distributions of quasi-conservative chemical 'tracers'. The usefulness of a chemical as a transport tracer depends crucially on its 'chemical lifetime' i.e. the characteristic timescale for replacement or destruction by local sources or sinks compared with the 'dynamical timescale' i.e. the time for advective processes to transport the tracer (e.g. through a scale height in the vertical or from pole to equator meridionally). If the chemical lifetime is very long, the chemical will be well-mixed and of little use as a tracer. If the chemical lifetime is very short, it's distribution will be entirely determined by local photochemical processes so that if the chemical is advected it will quickly adjust to its new surroundings and there will be no 'memory' of where it came from. If, on the other hand, the chemical and dynamical timescales are similar, the distribution will be determined by both transport and chemistry. If there are suitable vertical or meridional gradients (or both) in tracer distribution, then evidence of advection by the meridional circulation will be present in the tracer distribution.

In order to examine meridional circulations associated with the SAO and QBO we use a variety of tracers depending on which height region they have a good vertical gradient. CH_4, N_2O and H_2O have been especially useful in helping to understand meridional circulations associated with the SAO. Volcanic aerosol and O_3 have been useful for diagnosing the QBO circulation.

N_2O and CH_4 both have sources in the troposphere, are slowly transported into the stratosphere where they are destroyed by photolysis or oxidation, with a chemical lifetime of the order of months (see figures 6a and 6b). Their distributions therefore decrease with height, with useful vertical gradients in the mid to upper stratosphere. On the other hand, H_2O increases with height in the stratosphere, since it's prime stratospheric source is via methane oxidation. It has useful vertical gradients in the mid to upper stratosphere and also near the tropical tropopause region, although interpretation of the latter is complicated by the temperature dependence of water vapour in this region, as it enters the stratosphere from the troposphere via tropical upwelling. Volcanic aerosol injected into the lower stratosphere after major eruptions e.g. El-Chichon and Pinatubo also has useful vertical gradients in the lower stratosphere. Finally, and importantly, there is also a very marked QBO signature in ozone, although this is complicated to interpret since there are wide variations in its chemical lifetime with height.

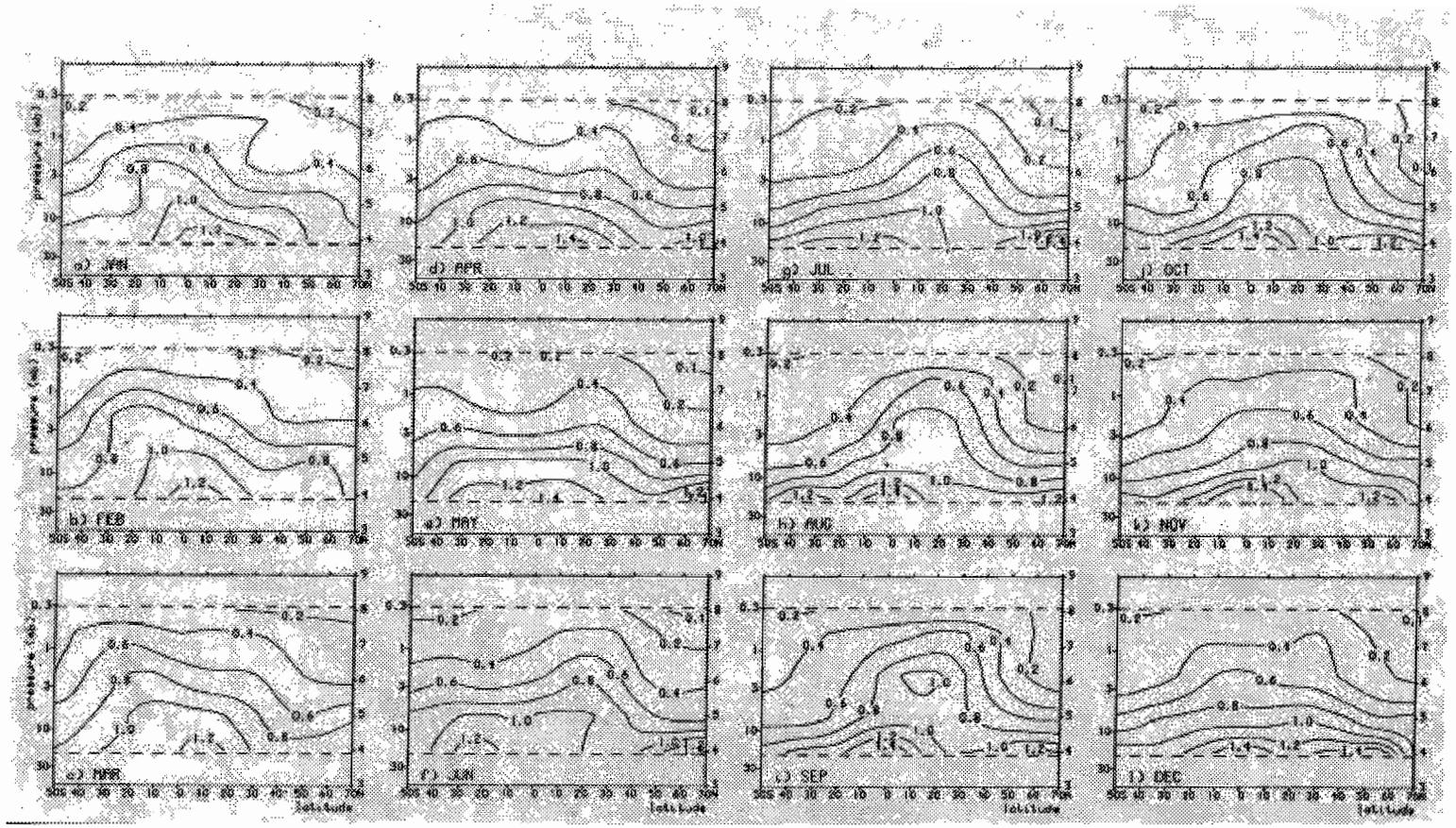


Figure 6a Monthly mean cross section of CH_4 (ppmv) between 50^0 S and 70^0 N for 1979 measured by the NIMBUS 7 SAMS.

Reproduced with permission from Jones and Pyle (1984)

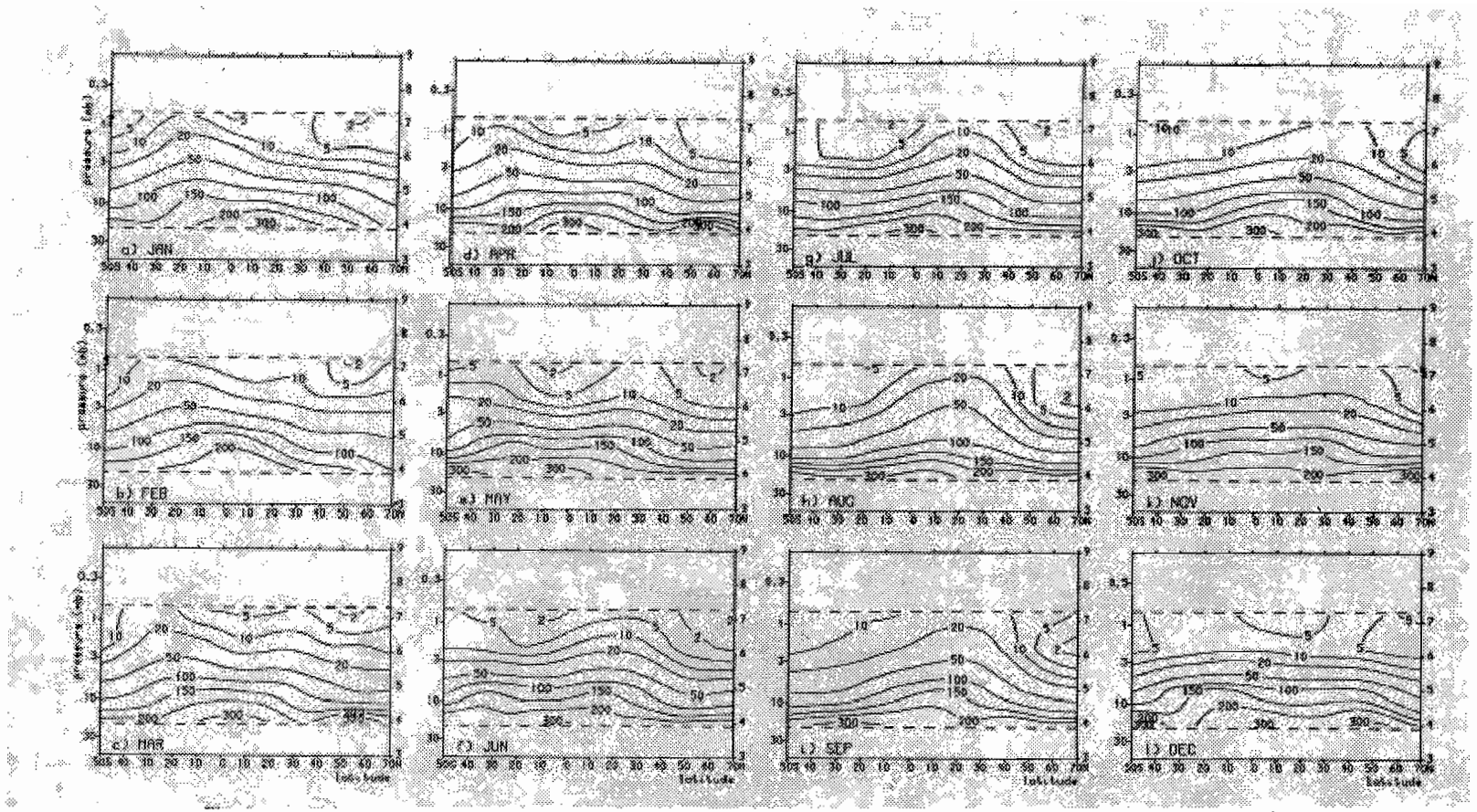


Figure 6a Monthly mean cross section of N_20 (ppmv) between 50^0S and 70 ^{0}N for 1979 measured by the NIMBUS 7 SAMS.
Reproduced with permission from Jones and Pyle (1984).

3. The Semi Annual Oscillation: Influence on Chemical Tracer Distributions

The time-series of equatorial zonal wind shows that there is a descending region of westerly shear in the upper stratosphere and lower mesosphere (35-60 km) in Feb – April and August – October (figure 7). This implies a region of anomalous equatorial descent propagating downward from around 60 km in February and August to around 35 km in April and October.

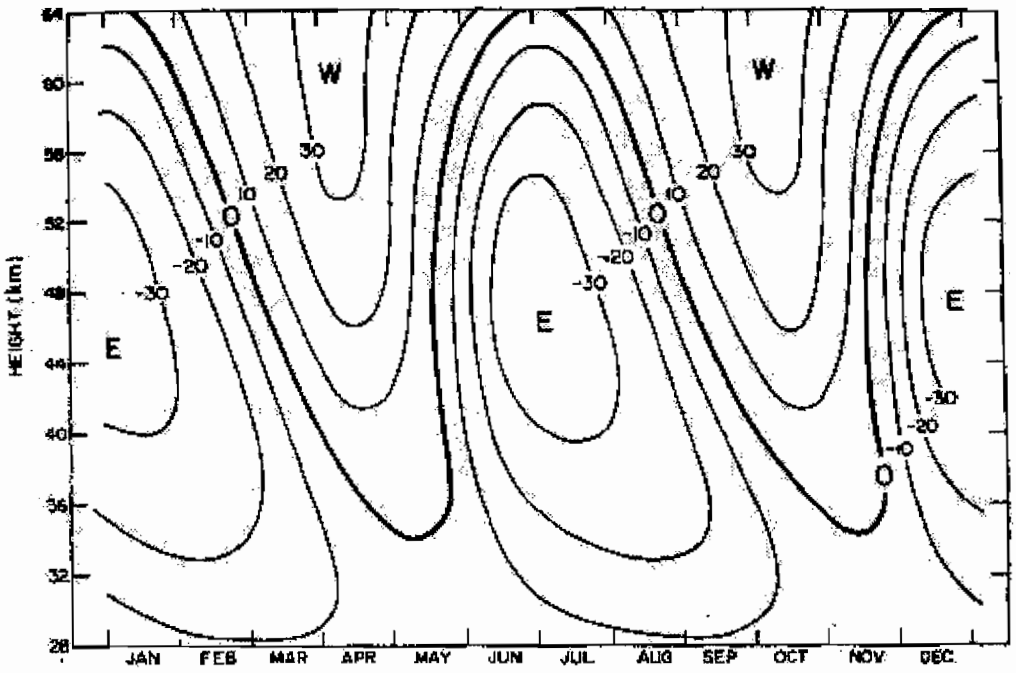


Figure 7 Time – height cross-section at the equator showing the estimated semi-annual mean zonal wind oscillation (ms^{-1}). (After Reed 1966)
Reproduced with permission from Gray and Pyle (1986).

Early satellite observations of N_2O and CH_4 have helped in the study of the SAO influence. Figure 6 shows the solstice and equinox distributions in 1979. Around January and July, during the solstice periods, there is a single maximum in trace gas centred approximately over the equator, consistent with tropical upwelling. The maximum is slightly off the equator into the summer hemisphere, consistent with advection by the mean summer-to-winter circulation during these times. Around March-May and November, however, this distribution is replaced by a 'double peak', with two maxima at around 30° S and 30°N above about 40 km. 2-dimensional modelling experiments have demonstrated that this is due to the induced meridional circulation associated with the descending westerly shear zones of the SAO (e.g. Gray and Pyle 1987). The maximum amplitude of the double peak in trace gases is found to slightly lag the maximum in westerly shear by about a month because the trace gases take a while to respond to the anomalous circulations. Also, the amplitude of the double peak is rather smaller in November and less prolonged than in March-April. This is believed to be due to an asymmetry in strength of the westerly shear zone around October compared with March.

More recent observations in figure 8 have confirmed the double peak feature. In fact, these more recent observations have revealed a noticeable QBO influence on the strength of the double peaks. In April 1993 and 1995 the double peaks are strong and well-formed but in 1994 and 1996 they are much smaller and less well-formed (Ruth et al., 1997, Randel et al. 1998). There is, of course, an expectation that the zonal wind QBO might influence the strength of the SAO westerly forcing and hence the strength and timing of the westerly shear. The waves responsible for the forcing of the SAO westerly phase propagate upwards through the equatorial lower stratosphere, which is dominated by the QBO. During the easterly phase of the QBO the westerly phase-speed waves propagate more effectively and the westerly forcing in the SAO height region is greater than during the opposite (westerly) phase of the QBO. However, figure 8 shows that the double peak structures are stronger during the *westerly* QBO phase, and not the easterly phase as one might expect. In a modelling study, Kennaugh et al (1987) showed that this is because in easterly QBO phase years the westerly shear and induced circulation associated with the SAO westerly phase is indeed stronger (as expected) but it descends through the atmosphere so quickly that the tracers do not have sufficient time in which to respond! In westerly QBO phase years the SAO-induced circulation descends more slowly and thus remains at any one level for long enough that the tracer distributions are able to respond.

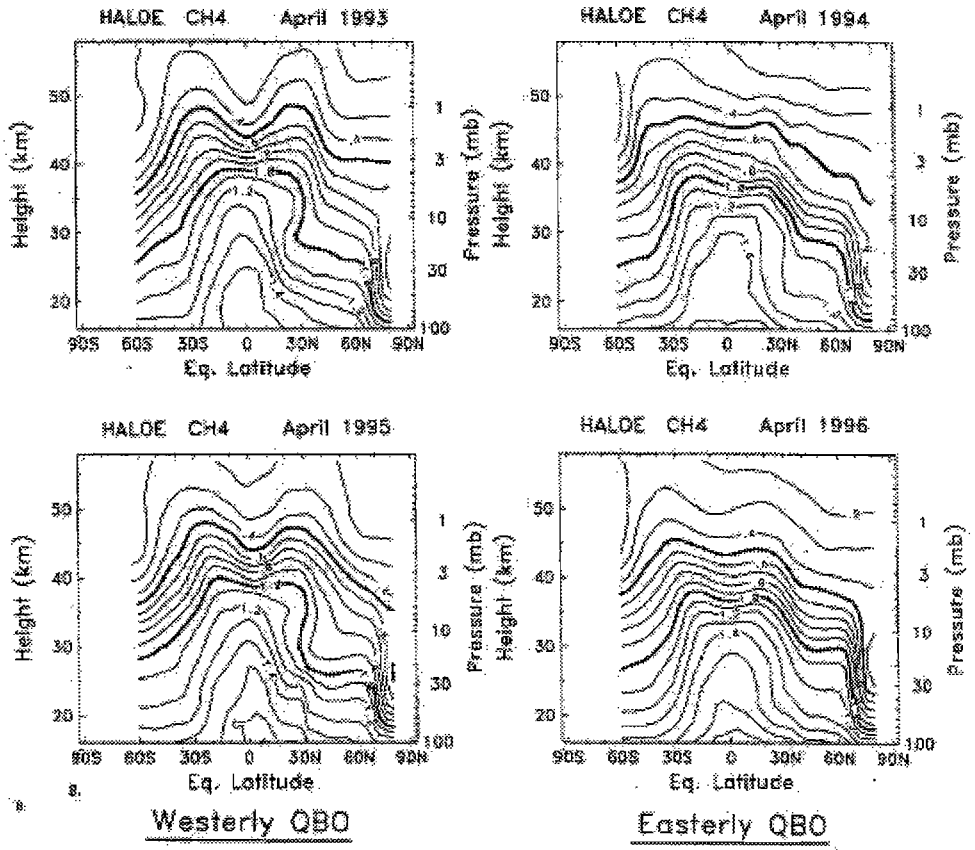


Figure 8 Meridional cross sections of CH, During April 1993, 1994, 1995, and 1996. Note the double-peaked upper stratosphere patterns in 1993 and 1995.

Reproduced with permission from Randel et al. (1988).

4. Influence of the QBO on Aerosol Distributions

As already mentioned, major volcanoes that inject aerosol into the lower stratosphere provide useful tracers of anomalous meridional circulations in the height region of the QBO. Figure 9 shows a classic study of aerosol observations (Trepte and Hitchman 1992). The top panel shows the distribution soon after the eruption of Mount Ruiz in 1984 during a descending westerly phase of the QBO. The bottom panel shows the distribution soon after the Mount Pinatubo eruption, during a descending easterly QBO phase. The direction of the implied QBO-induced circulations are marked by arrows. The aerosol distributions mirror these extremely well, with a well-formed double peak structure during the descending west QBO phase (top panel) and evidence of enhanced ascent during the descending east QBO phase (bottom panel). Note that the double peak in this figure is at 25 km, well below the level of the SAO double peak structures (40-50 km) described in the previous section.

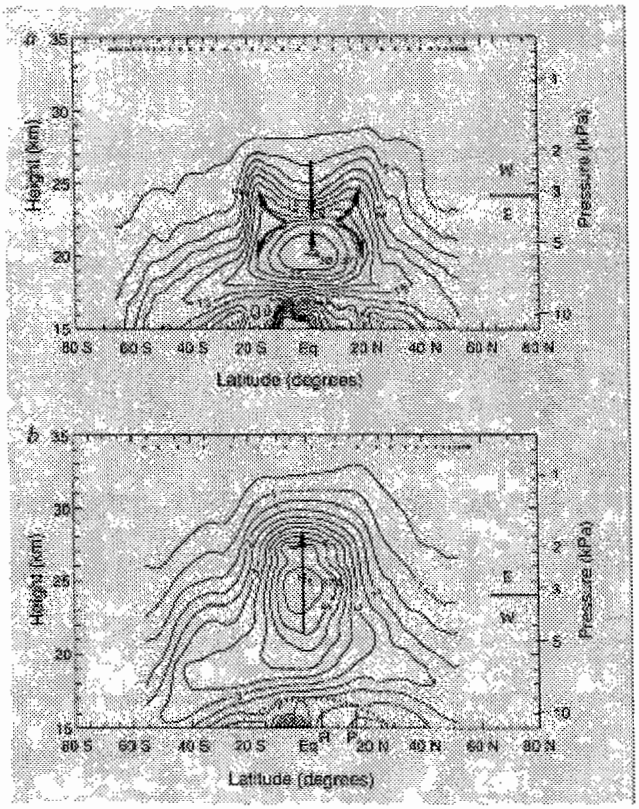


Figure 9 Latitude-height cross sections of observed aerosol extinction ratio for two 40-day periods representative of the two phases of the QBO. (a) Westerly shear phase centred on November 11 1984 (contour interval is 2.5). (b) Easterly shear phase centred on October 4th 1988 (contour interval is 0.5). After Trepte and Hitchman [1992].
Reproduced with permission from Baldwin et al. (2001).

5. Ozone: Background

Ozone has a central role in the Earth's atmosphere. Fig. 10 shows the typical height – latitude distribution of ozone in 'parts per million by volume' (ppmv). The maximum is only of the order 10 ppmv, situated over the equator at around 10 hPa (approx. 30-35 km). The total column abundance of ozone determines the amount of radiation that can reach the surface of the Earth in the near-ultraviolet region. 'Total' or 'column' ozone is defined as the equivalent thickness of the ozone layer at standard temperature and pressure. The global mean value is approximately 3mm. Maps of column ozone are usually expressed in terms of Dobson Units (DU) where 1 DU = 10^{-5}m, so that the global mean column ozone is approximately 300 DU. Figure 11 shows the time evolution of column ozone at all latitudes throughout the year. There is an equatorial minimum of around 240 DU present at all times of the year. At high latitudes there are maxima in springtime of each hemisphere, reaching around 450 DU in the NH in April and around 380 DU in the SH in November. (Note the local minimum near the South Pole in October, coinciding with the Antarctic polar vortex. After 1985, when this plot was produced, the Antarctic Ozone Hole becomes much more evident).

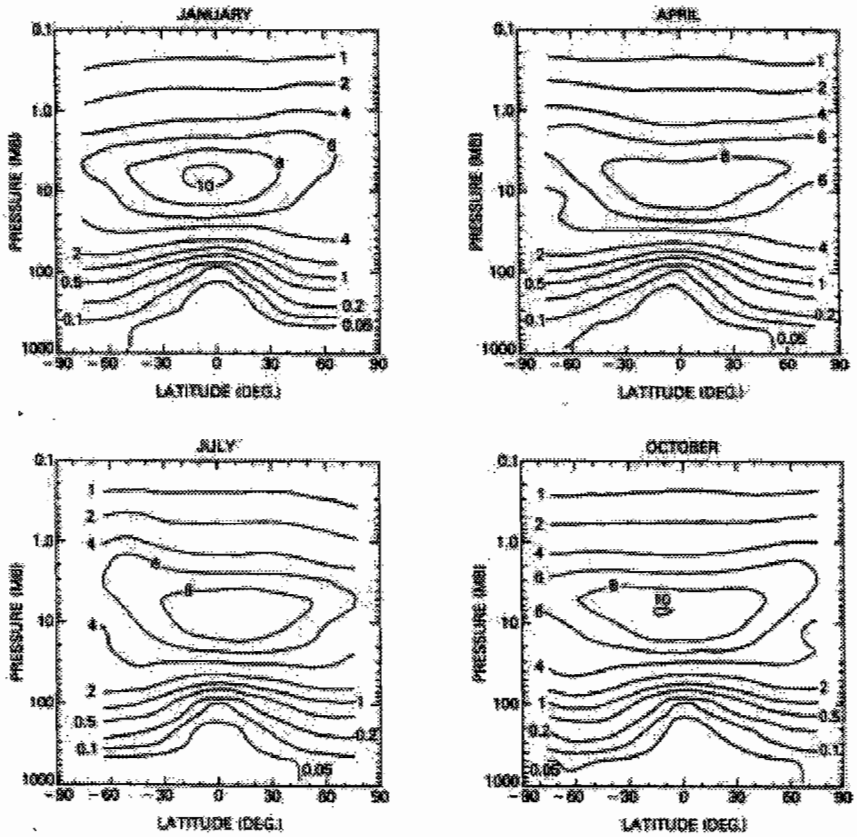


Figure 10 Latitude-height sections of the ozone ratio (ppmv) for January, April, July and October 1979, as observed by the *Nimbus 7* SBUV experiment. [After McPeters *et al* (1984).]
Reproduced with permission from Andrews et al. (1987).

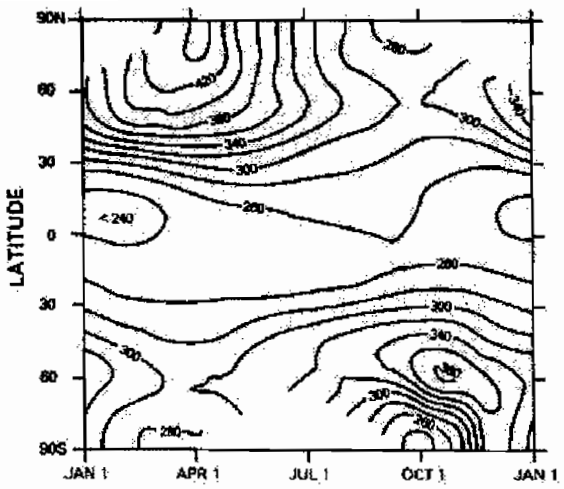


Figure 11 Time-latitude section showing the seasonal variations of total ozone (Dobson units) based on TOMS data. Note the springtime maxima near 90^0N and 60^0S. [After Bowman and Krueger (1985)].
Reproduced with permission from D.G. Andrews et al. (1987).

Figure 12 shows results of an analysis to determine trends in NH mid-latitude ozone amounts. Percentage deviations in total column ozone averaged over the latitude band 25^o-60^oN are plotted from ground-based and satellite observations. Note the substantial inter-annual variations in these measurements, a large part of which is associated with the QBO. In order for an accurate assessment to be made of the ozone trends over the past few decades, and in order to be able to model the atmosphere sufficiently well to make accurate predictions of future ozone trends, it is essential that the variability associated with the ozone QBO is understood, accounted for and removed from these trend analysis.

The photochemical lifetime of ozone varies substantially with height in the atmosphere, as illustrated by fig 13. The lifetime is of the order of hundreds of days at 20 km but reduces rapidly with height to only half a day at 45 km. As discussed earlier, by examining variations in tracer distributions, we can learn most about advection processes in the region where the chemical lifetime is of the same order as the dynamics (months). In the case of column ozone, this is the region around 25-30 km. The major features in the time-latitude section of column ozone (fig 11) can thus be understood in terms of the mean circulation across this 25-30 km region together with a knowledge of the background distribution of ozone (fig 10). The equatorial minimum in column ozone (fig 11) at all times of the year is due to tropical upwelling in this region, bringing up air with lower ozone amounts. The two maxima in spring of each of the hemispheres (April in the NH, October in the SH) can be explained as the accumulated result of transport by the mean summer-to-winter meridional circulation, with descent throughout the winter months transporting high ozone amounts from the equatorial source region down into the lower stratosphere.

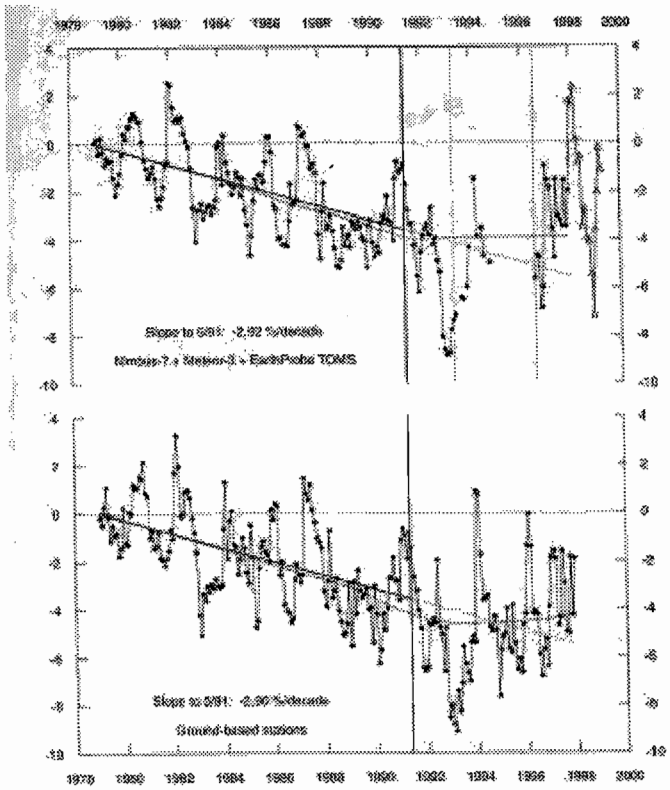


Figure 12 Deviations in total ozone, area weighted over 25^0-60^0N. A seasonal trend model (not including the effects of the solar cycle or the quasi-biennial oscillation) was fitted to the ozone data over the period from January 1979 to May 1991. The upper panel shows the TOMS record (1/79-4/99) and the lower panel shows a ground based series constructed from average ozone at stations in 5^0 latitude bands (1/79-12/97). Trend models are fit independently to the two data sets. The solid straight line represents the least squares fit to the deviations up to May 1991 and is extended as a dotted line through December 1997. The TOMS record is updated from WMO (1999) courtesy of L. Bishop, Allied Signal.)

Reproduced from DETR (1999).

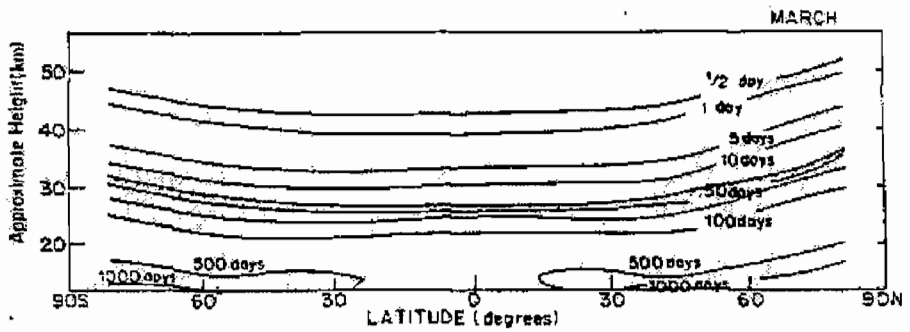


Figure 13 Latitude-height section of the photochemical lifetime of odd oxygen in the model for the month of March. The shaded area indicates the approximate position of the transition region, that is, the region in which the photochemical lifetime is of the same order as the time scale for meridional transport.

Reproduced with permission from Gray and Pyle (1989).

6. The Quasi Biennial Oscillation: Influence on Ozone

The QBO in column ozone is best illustrated by satellite observations e.g. from the SBUV (Solar Backscattered Ultra Violet) instruments in figure 14, which shows the latitude time-series of column ozone anomalies. The main features are an alternating equatorial anomaly of around 6-10 DU (approx. 4% of the background total column amount), a phase change in the subtropics and a subtropical / midlatitude anomaly of opposite sign of around the same magnitude.

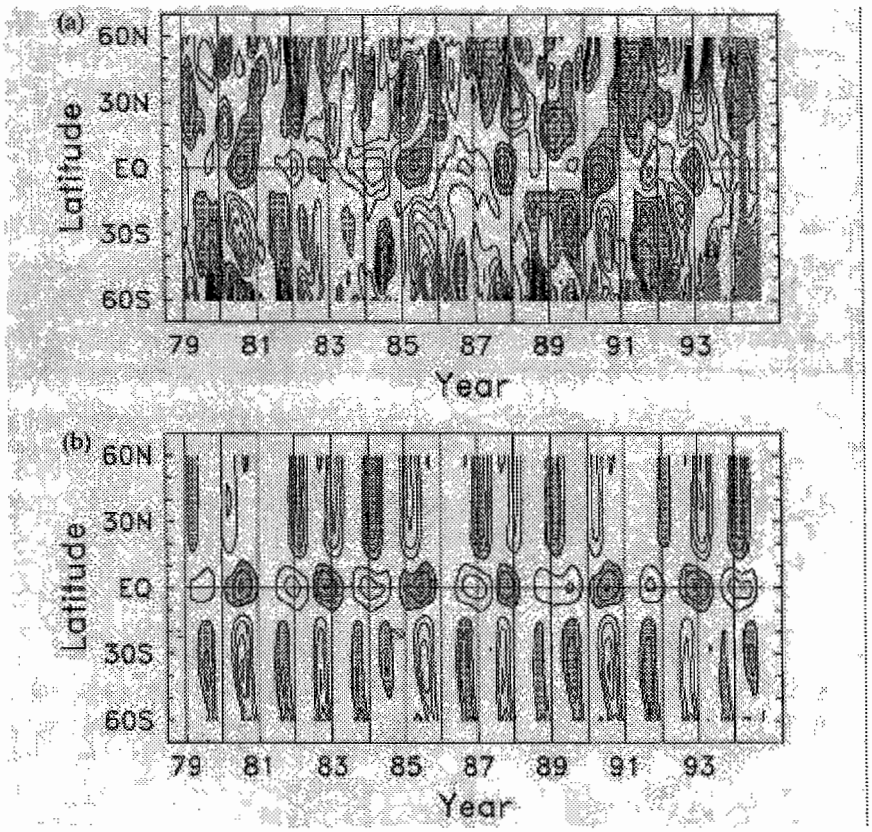


Figure 14 Latitude-time sections of column ozone anomalies from combined SBUV-SBUV/2 data: (a) full anomalies defined as deseasonalized and detrended over 1979-1994, and (b) the QBO component derived by seasonally varying regression analysis. Data in all panelled were multiplied by cos(latitude) to account for area weighting. Contour interval is 3 DU with 0 contours omitted and positive values shaded. Diagonal hatching denotes unreliable data. Vertical lines denote January of each year. After *Randel and Wu* (1996).
Reproduced with permission from Baldwin et al. (2001).

6.1 Equatorial Ozone QBO

The equatorial column ozone anomaly is closely associated with the equatorial zonal wind QBO and figure 14 shows clearly the irregular timing of the equatorial QBO, with anomalies that sometimes last only 9-12 months (e.g. 1982, 1983) and sometimes 12-18 months (e.g. 1989/90). Note, therefore, that the equatorial ozone QBO signal, like the zonal wind QBO, is not synchronized with the annual cycle.

Fig 15 shows the time-series of SBUV equatorial column ozone anomaly together with a reference QBO wind time-series (Randel and Wu 1996). There is excellent correspondence between the observed ozone anomaly and the reference wind series. Positive (negative) anomalies are present when the zonal winds in the lower stratosphere are westerly (easterly). To understand this relationship between zonal wind and ozone QBO signal, we consider the westerly QBO phase as an example. The time of the maximum westerly vertical wind shear at a particular level corresponds to the time of the warmest phase of the equatorial temperature QBO and hence to the time of maximum relative descent. (Remember that this does mean that there is actual descent, since the zonal mean circulation is always upwards at the equator in the lower stratosphere; relative descent simply means that this upwelling is weaker than the average). This anomalous descent occurs in the height region where the background ozone amounts increase with height (see fig 10). Thus, anomalous descent through the 25-30 km height region produces an anomalous increase in column ozone. Above about 30 km the ozone lifetime is relatively short (figure 13) and ozone is replaced by chemical production relatively quickly. Fig 16 shows a schematic of this process. The maximum total column ozone anomaly is the accumulated response to the descent of the westerly shear zone and will therefore occur when the column has been displaced farthest downward into the lower stratosphere. This will occur after the descent of the westerly shear to the lowermost stratosphere i.e. around the time of the maximum westerly anomaly at 20-25 km.

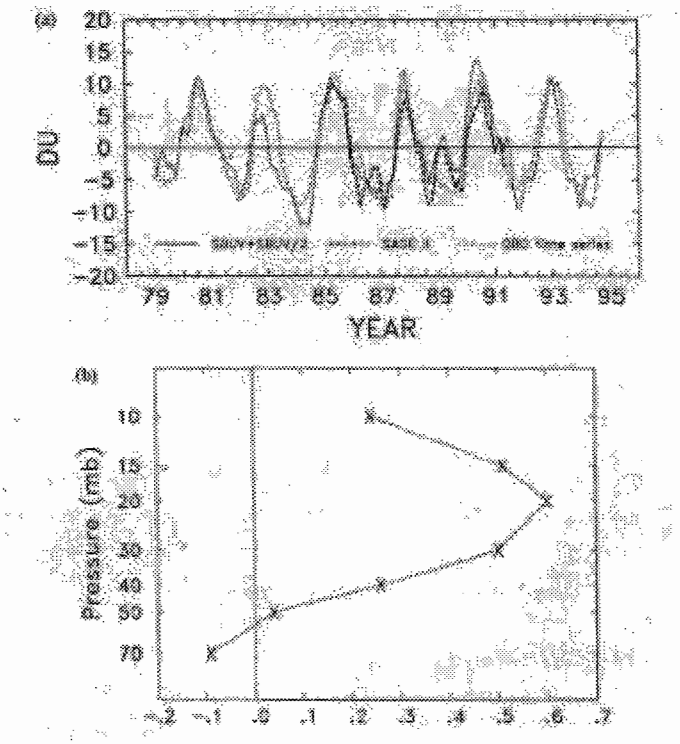


Figure 15 (a) Time series of equatorial ozone anomaly in Dobson units (DU, solid line) from solar backscattered ultraviolet (SBUV) and SBUV/2 together with a reference QBO wind time series (dotted curve compiled by multiplying the observed winds at Singapore by the weighting profile shown in figure b). After Randel and Wu (1996). *Reproduced with permission from Baldwin et al. (2001).*

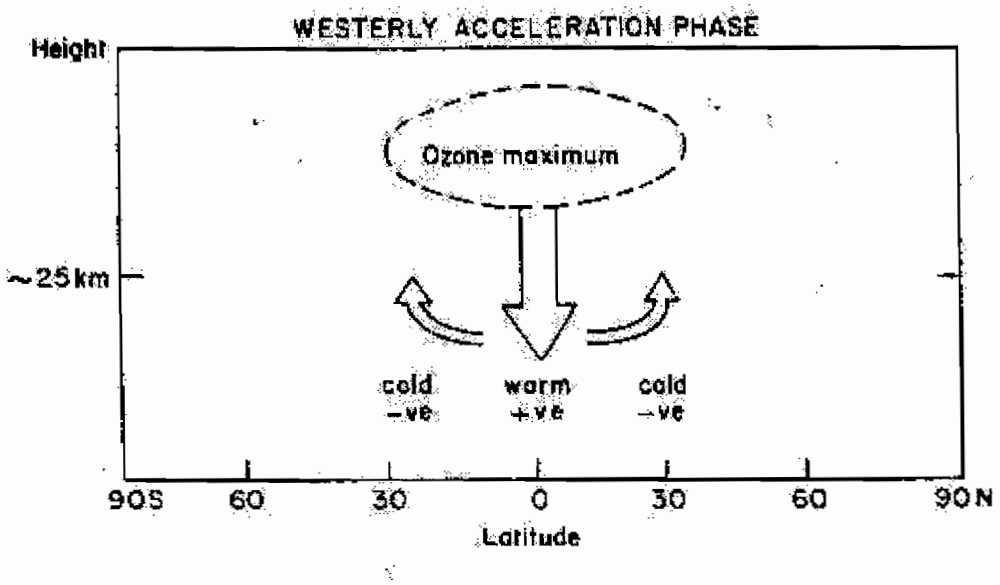


Figure 16 A schematic diagram showing the induced meridional circulation associated with a westerly phase of the QBO. The positive and negative signs indicate the expected column ozone anomaly.
Reproduced with permission from Gray and Chipperfield,1990.

Although the above mechanism accounts for a large component of the QBO in column ozone, there are nevertheless a number of additional factors that influence the column anomaly. Figure 17 shows the time height cross section of the QBO ozone density anomaly (DU km^{-1}). This can be used to visually determine the contribution from each height range to the total column anomaly (which is simply a sum in the vertical of the ozone density anomaly). Below about 28 km i.e. where the ozone lifetime is of order months, ozone is dynamically controlled and shows a gradual downward propagation of the anomaly with time, corresponding to the downward propagation of the shear zones with time. However, there is also a small contribution from above 28 km, where the chemical lifetime of ozone is shorter and is controlled by QBO changes in its photochemical sources and sinks (see section 7 for further details). The anomalies at the two levels are approximately a quarter cycle out of phase. There is therefore a small contribution to the column from above 28 km that also influences the timing of the total column anomaly.

As mentioned earlier, changes in ozone distributions in the lower stratosphere have radiative implications and, in particular, have a direct effect on short-wave heating. The effect of including the coupling between the ozone QBO anomalies and heating rates in computer models of the QBO tends to reduce the heating rate that would otherwise be calculated from a given temperature anomaly in the lower stratosphere. This is therefore an additional factor that needs to be taken into account when trying to understand the phase relationship between the QBO signals in ozone and zonal wind.

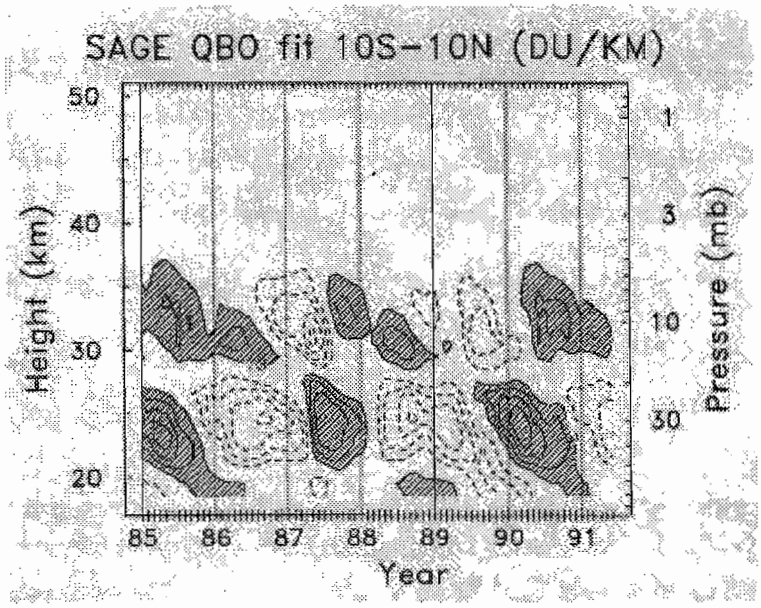


Figure 17 Height-time series of interannual anomalies in ozone density (DU km^{-1}) derived using a regression analysis to isolate the QBO variation. Contour intervals are 0.3 DU km^{-1}, with 0 contours omitted and positive values shaded. After Randel and Wu (1996).
Reproduced with permission from Baldwin et al. (2001).

6.2 Subtropical and Mid-latitude Ozone QBO

A QBO signal in column ozone in the subtropics that extends to middle and high latitudes is clearly evident in fig 14. There is a 180° phase change at around 15° in each hemisphere with the higher-latitude anomaly extending to at least 60° but with its maximum at approximately 30°- 40° latitude. The subtropical anomaly corresponds broadly to the return arm of the local equatorial QBO circulation, with ascent (descent) in the subtropics associated with westerly (easterly) equatorial shear.

However, there are several significant departures from the expected behaviour. Firstly, the induced equatorial QBO circulation is confined to low latitudes and hence cannot explain the signal poleward of about 30°. Secondly, the timing and the amplitude of the subtropical anomalies are not symmetric about the equator; the subtropical minima and maxima are approximately 6 months apart and coincide with the local late winter / spring. This latter feature is highlighted in fig 18 which shows a regression fit of Total Ozone Mapping Spectrometer (TOMS) column ozone amounts to the 30 hPa Singapore winds (Randel and Cobb 1994). On average, the subtropical regression anomalies reach a maximum in March and August in the NH and SH respectively. There is therefore a change in the period of the anomaly as one moves to higher latitudes.

Jones et al. (1998) used a 2-d model to show that this hemispheric asymmetry in the amplitude and timing of the subtropical and higher latitude ozone QBO is caused by an asymmetry in the QBO-induced meridional circulation. Rather than being symmetric about the equator, as figure 2 implies, their modelled circulation was strongest in the winter hemisphere. This is confirmed by figure 19 that shows a cross-section of the QBO

temperature anomaly and (residual mean) circulation in February 1994 derived from assimilated meteorological analyses (Randel et al., 1999). There is induced ascent coinciding with the minimum temperature anomaly at 25 km but a descending return arm of the circulation only in the NH subtropics and not the SH subtropics. Similarly, the temperature anomaly is highly symmetric.

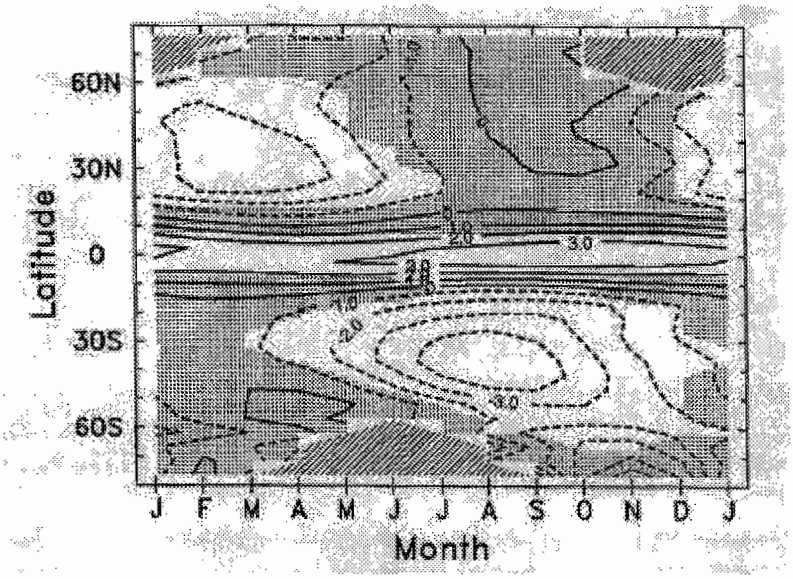


Figure 18 Latitude-time section of QBO-associated regression fit of zonal-mean Total Ozone Mapping Spectrometer (TOMS) column ozone (DU) to the 30-hPa Singapore winds for the period 1979-1994. Shading denotes regions where the statistical fits are not different from zero at the 2 sigma level. Hatched regions denote the polar night, where no ozone data are available.

Reproduced with permission from Baldwin et al. (2001).

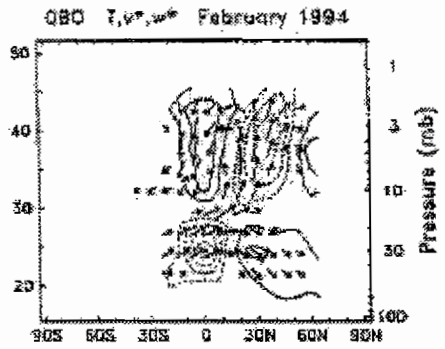


Figure 19 Cross section of QBO anomalies in February 1994. Temperature anomalies are contoured (± 0.5, 1.0, 1.5, K etc) and components of the residual mean circulation (v*, w*) are as vectors (scaled by an arbitrary function of altitude).

Reproduced with permission from Baldwin et al. (2001).

Jones et al. (1998) noted that because the QBO zonal wind jets are symmetric about the equator, the horizontal gradient $\partial u/\partial y$ is asymmetric (where u is the QBO-induced component of the zonal wind. Additionally, under solstice conditions, the background (QBO-independent) zonal wind gradient $\partial u/\partial y$ is also highly asymmetric. Hence, advection terms, particularly $v(\partial u/\partial y)$ and $v(\partial u/\partial y)$ introduced strong hemispheric asymmetry into the QBO-induced meridional circulation.

Figure 14 shows that occasionally there is a 'missed' subtropical anomaly, for example, in 1981, 1986 and 1991 in the NH and in 1993 in the SH. These missed anomalies are probably due to the timing of the equatorial QBO relative to the annual cycle. The formation of a significant winter subtropical anomaly requires not only a strong QBO-induced circulation but also a strong background (summer-to-winter) horizontal advection to help strengthen the winter side of the circulation. These conditions need to last for a month or two to allow the ozone distribution to respond to the induced circulation. If either of these requirements is not present for a sufficiently long time, the subtropical anomaly is unlikely to form in that year. Similarly, if the timing and duration of the equatorial wind QBO is such that it persists in the same phase for two successive winters, then anomalies of the same sign will occur in successive winters in the subtropics of that hemisphere. The latter is evident in the SH in 1983-1984 and 1988-89.

The mechanism for the poleward extension of the subtropical ozone QBO signal to mid-latitudes is not well understood. It is likely to involve an interaction between planetary waves and the equatorial QBO. The modulation of planetary wave forcing by the equatorial wind QBO results in a stronger large-scale (summer-to-winter) circulation in easterly phase years. Stronger downwelling in the winter midlatitudes will produce a positive (negative) column ozone anomaly in easterly (westerly) years, as observed (Tung and Yang, 1994). On the other, the extension of the ozone anomalies to middle and higher latitudes may be due to mixing processes associated with planetary scale waves. At midlatitudes, therefore, a number of factors and feedback processes may contribute to the ozone anomaly.

There is also an indication of a QBO influence in the winter polar regions. The polar ozone QBO is approximately in phase with the midlatitude ozone QBO and is seasonally synchronised in the same way, with maximum amplitude in springtime (Garcia and Solomon 1987, Randel and Cobb, 1994). Observational evidence for the polar ozone QBO is statistically less significant, partly due to the high level of interannual variability in the springtime vortex. It is possible that there may be a feedback loop between the modulation of the extratropical temperature by the QBO (due to a QBO in planetary wave transport), the formation of polar stratospheric clouds (which is highly temperature dependent) and hence with the underlying chemical destruction that gives rise to the ozone hole (Mancini et al. 1991, Butchart and Austin, 1996).

7. The Quasi Biennial Oscillation: Influence on Other Trace Gases

Model studies of the QBO that include a treatment of chemical processes in the stratosphere in addition to the dynamical and radiative processes predict a strong QBO in

many of the trace gases that have a relatively short chemical lifetime. This includes e.g. members of the nitrogen family and chlorine family. A short chemical lifetime means that the trace gas distribution will not show any response to advection by the mean circulation (since it will very quickly re-establish photochemical equilibrium to its new surroundings). However, it may still show a response to the QBO in temperature or to a QBO signal in any other trace gas that influences its chemistry. These model predictions have subsequently been confirmed in some of the trace species for which there are adequately long time-series of measurements.

7.1 The Nitrogen family

Figure 20 shows the latitude time-section of the QBO signal in a variety of trace gases at 22 km, taken from an idealised 2-dimensional model study (Gray and Chipperfield 1990). Although the lifetime of, for example, NO_2 is relatively short, it may be considered as part of a 'family' of nitrogen-containing trace gases, general referred to as NOy, the odd nitrogen family. While there may be chemical interactions and hence relatively short timescale variations in concentration of the individual members of the family, the sum of the individual members

$$NOy = NO_2+NO+HNO_3+2N_2O_5+HO_2NO_2+ClONO_2$$

has a relatively long chemical timescale. The distribution of NOy in the stratosphere increases with height in the lower stratosphere, so the QBO signal in NOy is similar to the column ozone QBO anomaly. It has a negative (positive) anomaly in equatorial latitudes during the easterly (westerly) phase of the zonal wind QBO, a phase change at around 20° and a subtropical anomaly of opposite sign, with an asymmetry in the amplitude of the subtropical anomaly. There is a peak-to-peak variation of approximately 25% at the equator. This variation is much larger than the 10% percentage variation in N_2O at this level and reflects the steeper vertical gradients in NOy compared with N_2O. As already mentioned, the chemical lifetime of NO_2 is relatively short and the QBO signal is therefore not a direct result of the QBO-induced circulation. Instead, it is a result of the temperature QBO through the temperature dependence of the following reaction:

$$NO + O_3 \rightarrow NO_2 + O_2$$

A positive temperature anomaly at the equator, associated with a westerly shear zone, results in the $NO:NO_2$ partitioning being shifted in favour of NO2, resulting in a positive NO_2 anomaly. The converse is true in the case of easterly shear. The resulting percentage variation at this level is 30-40 %.

The modelled percentage variation in N_2O_5 is extremely large, of the order of 60-70 %. Again, this is due to the temperature dependence of the key chemical reaction that influences the abundance of N_2O_5. At present, there are no measurements of N_2O_5 with which to confirm this prediction.

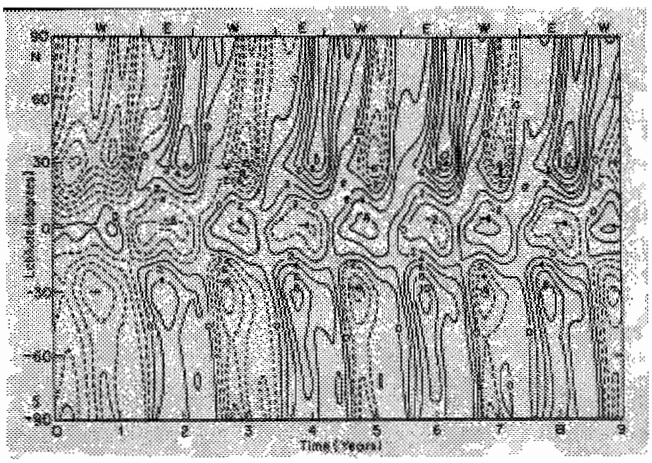


Fig 20a Latitude time-series of the modelled column ozone anomaly (Dobson Units). Contour interval 2 DU. Dashed lines denote negative values. The direction of the equatorial zonal winds at approximately 50 mb is indicated at the top of the figure.

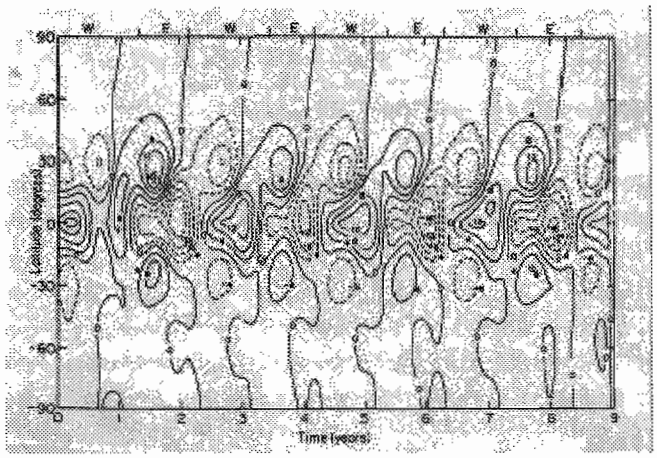


Fig 20b As figure 20a but for total odd nitrogen ($NO + NO_2 + HNO_3 + 2 * N_2O_2 + HO_2NO_2 + ClONO_2$). Contour interval 3%.

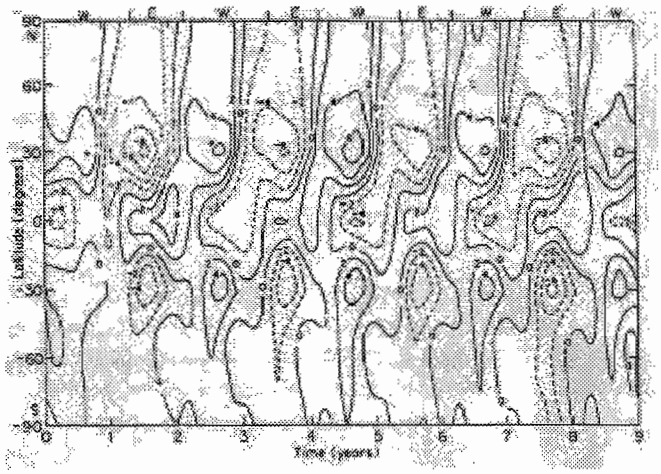


Fig 20c Latitude time-series of the modelled QBO anomaly in N_2O mixing ratio at 22km expressed as a percentage of the

ambient N_2O at that level. Contour interval 2%.

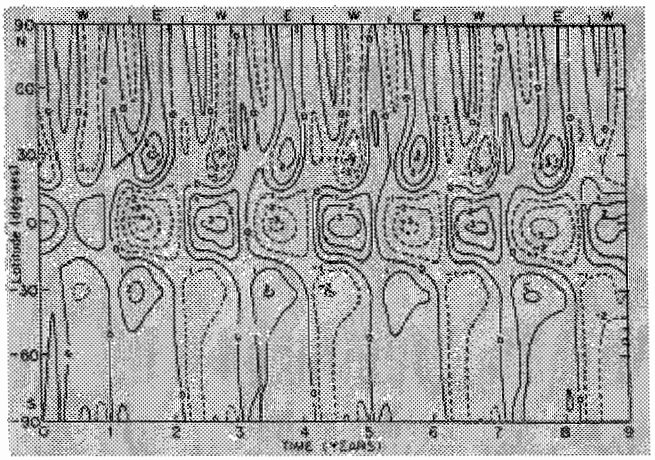


Fig 20d Latitude time-series of the modelled QBO anomaly in temperature in degrees Kelvin. Contour interval is 1K.

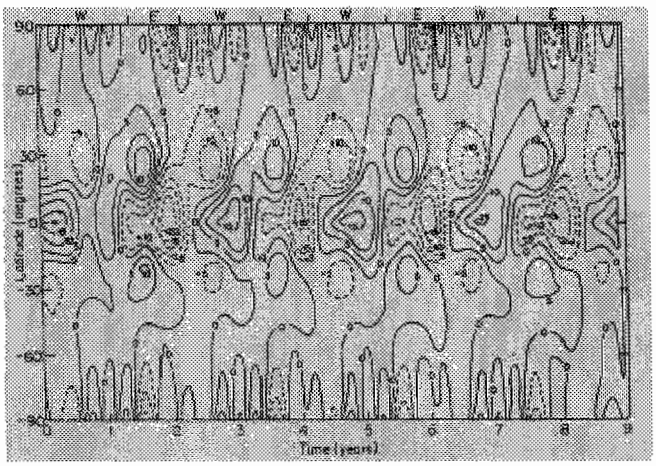


Fig 20e As figure 20c but for NO_2. Contour interval 5%.

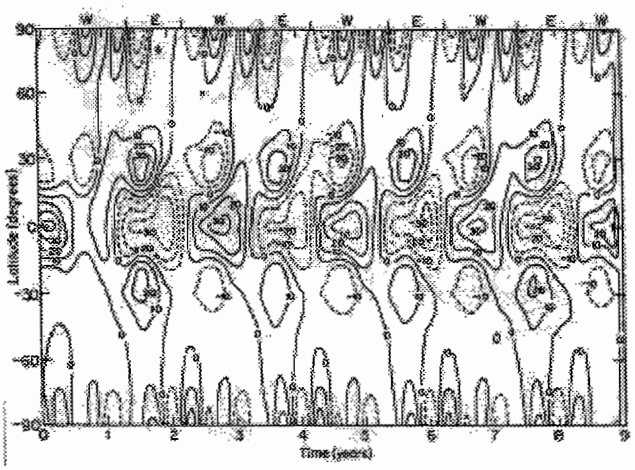


Fig 20f As figure 20c but for N_2O_5. Contour interval 8%.

Figs 20a-f reproduced with permission from Gray and Chipperfield (1990).

The 'chemically-controlled' ozone QBO anomaly above about 28 km in figure 17 (as opposed to the 'transport-controlled' anomaly below 28 km) can be explained by changes in the photochemical sources and sinks of ozone, primarily via transport-induced variations of NOy (Chipperfield et al. 1994, Jones et al. 1998).

7.2 Water vapour

Figure 21 shows the distribution of water vapour in the stratosphere (ppmv) in January 1993 and January 1994 measured by satellite. Unlike the troposphere, the stratosphere is very dry, the main source being via methane oxidation. For this reason, water vapour mixing ratios increase with height (remember that methane decreases with height). Note the minimum value situated at the tropical tropopause and a single-peaked structure centred over the equatorial latitudes that is almost a mirror-image of the methane distribution. This single-peak structure is consistent with the mean tropical upwelling advecting relatively dry air from the region of the minimum at the tropical tropopause higher into the stratosphere. There is some indication of the influence by the induced QBO-circulation, indicated by the arrows.

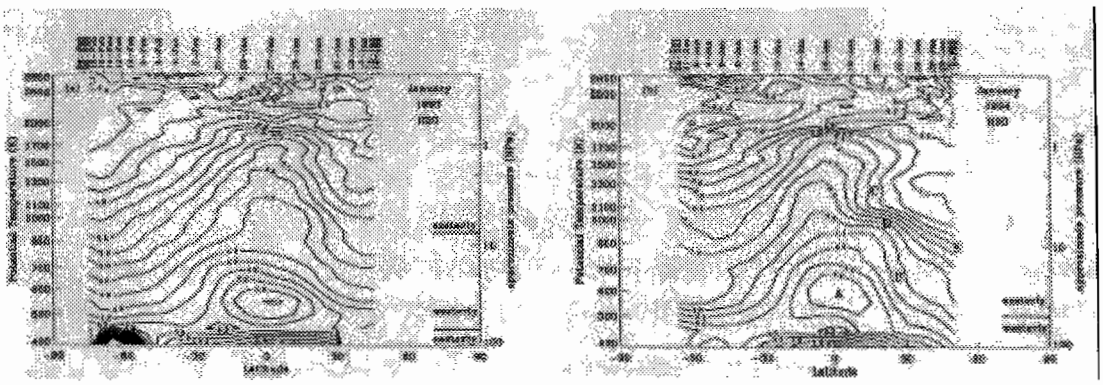


Figure 21 Latitude-theta cross sections of zonally averaged HALOE water vapor (ppmv). (a) sunrise observations: 10 Jan-25 1993, (b) sunset observations: 26 Dec 1993-4 Feb 1994. Contour interval 0.2 ppmv. Negative latitudes are south. The direction of the Singapore equatorial winds is also marked.
Reproduced with permission from Gray and Russell (1999).

Figure 22 shows the time-height section of the QBO anomalies in H_2O over the equator for the period 1992-99. The QBO anomaly ascends slowly with time, in contrast to the slow descent of the ozone QBO. This upward propagation is a reflection of the dynamics affecting the equatorial tropopause minimum. There is a well-defined annual signature in the tropopause water vapour minimum, closely tied to the annual temperature variation. This annual minimum in H_2O is advected upwards by the background tropical upwelling. (This is sometimes called the 'tape recorder' effect, since the annual 'imprint' of cold temperatures is imposed on the water vapour distribution and this signal is then advected upwards and away from the tropopause region). Figure 21(a) shows that the minimum is

centred around 40 hPa in January 1993 and has a rather flat, pancake shape. InJanuary 1994 however, it is slightly higher up and has a more rounded, circular shape (figure 21(b)). One possible mechanism for this QBO signal in tropical H_2O is a QBO-modulation of the tropical ascent rate, which is expected to be stronger in an easterly QBO wind shear zone than in the westerly shear zone. However, this does not appear to be the whole explanation, since there are some discrepancies between the observed behaviour of water vapour and methane. Another possibility is the QBO in tropopause temperatures since the water vapour abundance at tropopause level is highly dependent on the temperature as it enters the lower stratosphere through the tropopause region. More research is required to fully understand this QBO signal.

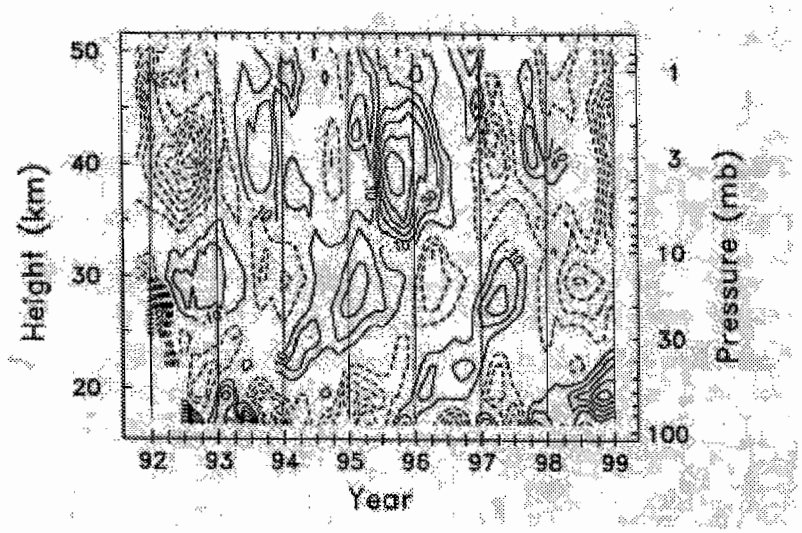


Figure 22 Time-height cross section of interannual anomalies in H_2O over the equator from the Halogen Occultation Experiment (HALOE) instrument. The contour interval is 0.1 ppmv, with 0 contour omitted.
Reproduced with permission from Baldwin et al. (2001).

Figure 21 also indicates a fairly substantial QBO influence on the H_2O distribution in the subtropics and mid latitudes. For example there is a so-called 'staircase' effect in January 1994 (figure 21(a)) with isolines of H_2O nearly vertical at points B and C, but nearly horizontal at point D (Dunkerton and O'Sullivan, 1996). This is an example of the competing influences of vertical advection by the mean circulation that will tend to steepen the isolines versus the influence of planetary-scale mixing processes that will tend to flatten the isoline gradients (Gray, 2000).

References:

Butchart, N. and J. Austin, On the relationship between the quasi biennial oscillation, total chlorine and the severity of the Antarctic ozone hole. *Q.J.Roy. Meteorol.* Soc., 122, 183-217, 1996.

Chipperfield, M.P., L. J. Gray. J.S. Kinnersley, and J. Zawodny, A two dimensional model study of the QBO signal in SAGE II NO_2 and O_3, Geophys. Res. Lett., 21, 589-592, 1994.

DETR, Stratospheric Ozone 1999. 99EP0458. H.M.S.O., 1999.

Dunkerton, T.J. and D.J. O'Sullivan, Mixing zone in the tropical stratosphere above 10 mbar, *Geophys. Res. Lett.*, 23, 2497-2500, 1996.

Garcia, R. R., and S. Solomon, A possible relationship between interannual variability in Antarctic ozone and the quasi-biennial oscillation, *Geophys. Res. Lett.,* 14, 848-851, 1987.

Gray, L.J., A model study of the influence of the quasi biennial oscillation on trace gas distributions in the middle and upper stratosphere. *J. Geophys. Res.*, 105, 4539-4551, 2000.

Gray, L.J. and M.P. Chipperfield, On the interannual variability of trace gases in the middle atmosphere, *Geophys. Res. Lett.*, 17, 933-936, 1990.

Gray, L.J. and J.A. Pyle, The semi annual oscillation and equatorial tracer distributions, *Q.J.R. Meteorol. Soc.*, 112, 387-407, 1986.

Gray, L.J. and J.M. Russell III. Interannual variability of trace gases in the subtropical winter stratosphere. J. Atmos. Sci., 1999.

Jones, D. B. A., H. R. Schneider, and M.B. McElroy, Effects of the quasi-biennial oscillation on the zonally averaged transport of tracers. *J. Geophys. Res.* 103, 11,235-11,249. 1998.

Kennaugh, R., S. Ruth, and L.J. Gray, Modeling quasi-biennial variability in the semiannual double peak, *J. Geophys. Res.,* 102, 16,169-16,187, 1997.

Mancini, E., G. Visconti, G. Pitari and M. Verdecchia, an estimate of the Antarctic ozone modulation by the QBO, *Geophys. Res. Lett.,* 18, 175-178, 1991.

Plumb, R.A. and R.C. Bell, Equatorial waves in steady zonal shear flow, *Q.J. Roy. Meteorol. Soc.*, 108, 313-334, 1982a.

Plumb, R.A. and R.C. Bell, A model of the quasi biennial oscillation on an equatorial beta plan, *Q.J. Roy. Meteorol. Soc.*, 108, 335-352, 1982b.

Randel, W. J. and J.B. Cobb, Coherent variations of monthly mean column ozone and lower stratospheric temperature, *J. Geophys. Res.,* 99, 5433-5447, 1994. and Cobb 1994

Randel, W.J., and F. Wu, Isolation of the ozone QBO in SAGE II data by singular decomposition, *J. Atmos. Sc.,* 53, 2546-2559, 1996.

Randel, W.J., F. Wu, J.M. Russell III, A. Roche and J. Waters, Seasonal cycles and QBO variations in stratospheric CH4 and H2O observed in UARS HALOE data. *J. Atmos. Sci.*, 55, 163-185, 1998.

Randel, W.J.F. Wu, R. Swinbank, J. Nash, and A. O'Neill, Global QBO circulation derived from UKMO stratospheric analyses, *J. Atmos. Sci.,* 56, 457-474, 1999.

Ruth , S., R. Kennaugh, L.J. Gray, and J.M. Russell III, Seasonal semiannual and interannual variability seen in measurements of methane made by the UARS Halogen Occultation Experiment, *J. Geophys. Res.,* 102, 16189-16,199, 1997.
Trepte , C.R., and M.H. Hitchman, Tropical Stratospheric circulation deduced from satellite aerosol data, *Nature,* 355, 626-628. 1992.
Tung, K., and H. Yang, Global QBO in circulation and ozone, part I, Re-examination of observational evidence, *J. Atmos. Sci.,* 51,2699-2707, 1994.

Chapter 6

Modelling the Equatorial Atmosphere

Elisa Manzini,

Max-Planck Institut, Hamburg, Germany

1. Introduction

General circulation models (hereafter GCMs) of the atmosphere have been developed in the last 50 years or so from relatively simple and idealized mathematical representation of the atmospheric flow, as the pioneering work of Charney et al. (1950), to the highly complex and technical global models now used in most of the world's operational weather forecast centres (e.g. European Centre for Medium-Range Weather Forecast, National Center for Environmental Prediction, Meteo-France, etc.). At every stage of their evolution, GCMs have synthesized our understanding of the physical processes responsible for the weather and climate of the Earth. They can be considered our most comprehensive tools to test our knowledge on the general circulation of the atmosphere (hence their name), to investigate how the weather systems and the climate may have changed in the past (e.g., during the ice ages) and how they may change in the future due to the anthropogenic perturbations of the atmospheric composition (as the emissions of carbon dioxide and methane thought to be responsible of global warming and the chlorofluorocarbons involved in the ozone hole). What allowed (and is allowing) the tremendous development of the current GCMs from the early versions has been the availability of more and more powerful (high speed) computers and data storage facilities as well as the realization of global scale monitoring systems (satellites, radiosonde networks, surface observation networks) and their world wide organization and management. Nowadays, the development of a GCM usually requires a team of scientists and is commonly the result of research activities going on at a number of research centres.

Given the complexity of the physical processes included in a GCM, there are no analytical solutions to its equations. Therefore the problem is solved by means of highly specialized numerical methods by discretizing the model equations on a 3 dimensional grid covering the whole globe and in time. Consequently, only part of the physics represented in the model is explicitly resolved, while the rest is approximated by means of parameterization methods. The basic requirements for a model to be a modern GCM are to be global, to have sufficient resolution to explicitly resolve the meteorological synoptic scale weather disturbances, and to have state-of-the-art parameterizations of sub-grid physical processes. Modern GCMs have specialized into three major classes: (i) numerical weather prediction models, (ii) climate models, and (iii) Earth system models.

Numerical weather prediction models deal with the simulation of the weather patterns and require a detailed description of the evolution of the atmosphere over a few days. This is principally an initial condition problem that has necessitated the advancement of such technics as data assimilation. Weather prediction makes use of models aimed at short range time integration and high spatial resolution, and performance depends crucially on the sophistication of the assimilation scheme. These models usually do not extend above the stratosphere.

The simulation of climate instead requires long term atmospheric statistics, measures of the variability of meteorological parameters such as temperature and precipitation. The focus being the prediction of the statistical properties of the climate system, an example being the seasonal forecast associated with the ENSO (El Nino Southern Oscillation) phenomenon. Models aimed at long range time integration and low spatial resolution are therefore used in this case. In addition, care must be taken in order to specify adequate boundary conditions, to which the climate can be quite sensitive. On the contrary, initial conditions are usually not relevant. Most of the atmospheric models that are used so far in climate research have the top boundary in the

lower stratosphere. Following the pioneering work of Fels et al. (1980), in the last two decades several efforts have been made to extend existing climate models to the mesosphere (the so called middle atmosphere models) and recently there are a few efforts to develop models from the Earth surface to the thermosphere (Fomichev et al. 2002).

In addition to models of the atmosphere, the simulation of climate requires the coupling of the atmospheric GCM to an ocean GCM and to take into account at the least as boundary conditions other climate components, as the cryosphere (snow and ice), the lithosphere (land), and the biosphere (natural and anthropogenic). While it is nowadays common practice to use coupled atmosphere-ocean models in climate simulations that focus on the troposphere, middle atmospheric models coupled to ocean models are still somewhat too demanding to the computing resources currently available at climate research centres.

GCMs that also include a representation of biogeochemical processes are the so called Earth system models. Interactive Earth system models are those that allow for feedback between the components (e.g., the atmosphere, the ocean, the biosphere, etc.) of the climate system.

2. Basic equations

A simulation with a GCM requires to compute the numerical solutions of a set of nonlinear partial differential equations based on the general principles of conservation of momentum, energy and mass for the atmospheric flow.

These equations describe the physical processes that relate the temporal evolution of thermo-hydrodynamical quantities and are specialized to the problem of a thin layer of fluid on a rotating sphere and are commonly referred to as the primitive equations. A major simplification used being the hydrostatic approximation.

The primitive equations are known, and they can be found in textbooks for climate studies. Here the primitive equations are reported following Peixoto and Oort (1992).

Horizontal equations of motions in spherical coordinates:

$$\frac{du}{dt} = \left(f + u\frac{\tan\varphi}{a}\right)v - \frac{1}{\rho a\cos\varphi}\frac{\partial p}{\partial\lambda} + F_\lambda$$

$$\frac{dv}{dt} = -\left(f + u\frac{\tan\varphi}{a}\right)u - \frac{1}{\rho a}\frac{\partial p}{\partial\varphi} + F_\varphi$$

where:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{u}{a\cos\varphi}\frac{\partial}{\partial\lambda} + \frac{v}{a}\frac{\partial}{\partial\varphi} + w\frac{\partial}{\partial z}$$

Hydrostatic equation:

$$\frac{\partial p}{\partial z} = -\rho g$$

Continuity equation:

$$\frac{\partial \rho}{\partial t} = -\frac{1}{a\cos\varphi}\left[\frac{\partial}{\partial \lambda}(\rho u) + \frac{\partial}{\partial \varphi}(\rho v \cos\varphi)\right] - \frac{\partial}{\partial z}(\rho w)$$

First law of thermodynamics (energy conservation):

$$C_p \frac{d}{dt}(\ln T) - R_d \frac{d}{dt}(\ln p) = \frac{Q}{T}$$

Equation of state for air (perfect gas):

$$p = \rho R_d T_v$$

Balance equation for water vapor:

$$\frac{dq}{dt} = (e - p) + D$$

Where: λ is the longitude, φ the latitude, and z the geometrical height of the coordinate system.

The dependent variables are: u zonal wind, v meridional wind, w vertical velocity, ρ density of air, p pressure, T temperature, and Tv virtual temperature, $Tv = T[1+(Rv/Rd -1)q]$

As usual, f defines the Coriolis parameter, a the radius of the Earth, C_p the atmospheric specific heat at constant pressure, R_d is the gas constant for dry air, and R_v the gas constant for water vapor.

The other terms represent forcing and dissipative processes: F_λ and F_φ include the contribution of the surface stress, the wave stress, and the eddy and molecular diffusion. Q includes the net heating/cooling rate from the radiative flux (solar and terrestrial), the latent and convective heat, and the dissipation of kinetic energy. In the balance equation for water vapor, e is the evaporation, p the precipitation and D the eddy and molecular diffusion.

To solve the primitive equations it is necessary to specify external forcing and boundary and initial conditions. For a model of the atmosphere, the external forcing to be considered is the solar radiation, that varies according to the orbit and inclination of the rotation axis of the Earth, defining the annual cycle.

Boundary conditions include the specification of the distributions of land and sea, of land surface types (such as forest, grass, and deserts), and of the sea surface temperature. Boundary

conditions are important because they affect the heat, moisture and momentum exchanges between the surface and the atmosphere.

The initial atmospheric state to be specified is usually not so crucial for climate studies that usually employ very long (multi-years) integrations where the so-called 'spin up' time, one year or so, is neglected. Generally, to avoid technical problems some care has anyway to be used in the choice of the initial conditions. Currently, atmospheric states from meteorological analysis are used as initial conditions.

In practice, the continuity equation is re-written in terms of the surface pressure (by vertical integration and by using the hydrostatic equation) that therefore substitute the density as a prognostic variable (see for instance Williamson et al. 1987). Thereafter, a vertical coordinate that is function of pressure is generally used.

Given that there are no analytical solutions, the primitive equations have to be approximated to a discrete set of equations and solved numerically on a computer. Therefore, inaccuracies are introduced due to the discrete resolution in space and time.

A most successful approach to discretize the primitive equations, widely used in the last 30 years, consists in using a finite series of spherical harmonic basis functions for the horizontal part and finite difference for the vertical part of the coordinate system. Thereafter, nonlinear terms and parameterized physics are calculated on an associated spacial grid. The convenience of this approach (named the spectral transform method) is that it is fast and accurate at relatively low horizontal resolutions and it facilitates the use of the implicit time scheme, a method to reduce noise and sources of numerical instability during the integration. Early works that developed the basis of the spectral models in use today include: Robert (1969), Orzag (1970), Bourke (1972, 1974), Hoskins and Simmons (1975), Simmons et al. (1978), Machenhauer (1979) and Robert et al. (1985). Major problems of this formulation are encountered when it is applied to the distribution of clouds and tracers as well as in the treatment of mountains at the bottom boundary. This is simply because the transform method produces ripples close to sharp changes in distributions and positive definite quantities may become negative.

Concerning the vertical coordinate, most models use a hybrid coordinate system, terrain following at the bottom and pure pressure above the tropopause (Phillips 1957; Kasahara 1974; Fels et al. 1980; and Simmons and Strüfing 1981).

Another widely used approach is to discretize the primitive equations on a longitude latitude grid (e.g. the so called grid-point models). Thereafter, finite differences or finite volumes are at the base of the numerical algorithms used. A shortcoming of this approach is the need to use a so-called polar filter, in order to overcome problems related to the convergence of the meridians at the poles. Literature on the first grid-point models includes the works of Smagorinsky (1963), Smagorinsky et al. (1965) and Arakawa (1966).

An alternative approach is to use a geodesic grid, such as the icosahedral - hexagonal grid, in order to avoid the previously mentioned polar problem. Although this formulation initially lost with respect to the spectral approach, it is receiving renewed attention because at high horizontal resolutions it may perform more efficiently than the transform method and it is also very appealing in climate-chemistry models because of the role of transport in determining the distributions of the chemical compounds (Ringler et al. 2000).

3. Physical Parameterizations

The forcing and the dissipative terms describe physical processes that are usually at smaller spacial and temporal scales than the synoptic meteorological motions and cannot be explicitly represented in GCMs. Therefore, they are parameterized: their most relevant effects are approximated in term of the large scale resolved variables. To parameterize a physical process is usually not straightforward: It is required a combination of first principles as well as empirical and pragmatic approaches. Most current climate research deals with the design, improvement and implementation of physical parameterizations. Concerning middle atmosphere GCMs, an important aspect is how to parameterize gravity wave effects.

Processes that are usually parameterized include:

- Shortwave and longwave radiative transfer. This is a basic parameterization, at the core of any GCM. Given the difference in the spectral range covered by the solar and terrestrial fluxes, usually independent schemes are used for the solar (shortwave) and terrestrial (longwave) part. The numerical schemes used generally differ in the spectral resolution considered. The so called line-by-line models are too computationally expensive for GCMs. Other source of differences arising among the schemes are the specification and treatment of the distributions of gases, aerosols, and clouds. Current topics of research focus on the radiative properties of clouds and aerosols and the coupling of radiative schemes to parameterizations of the microphysics of clouds and aerosols. Some physical processes and feedback involved in cloud physics are still poorly known.

- Cloud and convective processes. These parameterizations take into account condensation, evaporation, the distribution of clouds and the vertical eddy transport of momentum, heat, moisture, cloud water, ice, and tracers. Differences among convective schemes involve the degree of complexity at which the physical processes are considered, from simple convective adjustment to mass flux schemes. More comprehensive parameterizations usually increase the number of prognostic variables in the model, as the inclusion of continuity equations for cloud water and ice in the treatment of stratiform clouds.

- Turbulence transfer in the planetary layer and in the free atmosphere, kinetic energy dissipation. Usually diffusion operators are employed, the major differences among the schemes being how the diffusion coefficients are calculated.

- Land-surface processes such as surface - atmosphere exchanges (e.g. heat, momentum and moisture fluxes). Coupling to more or less complex parameterizations of the soil (water capacity, drainage, vegetation, snow cover, albedo, etc.) and the consideration of sub-grid scale inhomogeneities give rise to differences among the schemes used.

- Momentum flux deposition due to gravity waves. Differences in the schemes arise from the processes considered to be responsible for the dissipation and breaking of the gravity waves and from the gravity wave sources that are included. Most parameterizations consider orographic gravity waves and follow the Lindzen (1981) approach for the treatment of propagation, dissipation and breaking. More recently, schemes that allow the consideration of a broad spectrum of waves from a variety of sources have been developed (Medvedev and Klassen 1995; Hines 1997ab; Warner and McIntyre 2001). These latter schemes are usually only applied

in models that include the middle atmosphere, because they represent the effects of gravity waves in the mesosphere. For a controlled comparison of the Hines and Warner-McIntyre parameterizations see Charron et al. (2002). A challenge in parameterizing the effects of gravity waves is how to specify the gravity wave spectrum in the troposphere, related (in no straightforward way) to the gravity wave sources. See for instance Manzini and McFarlane (1998) and Medvedev et al. (1998) for the impact of using different gravity wave source spectra but the same parameterization for propagation, dissipation and breaking of the gravity waves. Limits to the developments of a gravity wave parameterization have been insufficient information from observations that may help to constrain the parameterizations, such as the availability of a climatology and variability on the gravity wave parameters as typical wind speeds, phase speeds, horizontal and vertical wavelengths, etc. Currently, this is a quite active subject of research.

4. Current Research in Middle Atmosphere Modelling

An important aspect of the modelling of the middle atmosphere circulation is the consideration of the effects of gravity waves, in particular the momentum that is transported by upward propagating gravity waves and deposited in the large scale flow at high altitudes (e.g. mesosphere). Gravity waves are predominantly generated in the troposphere by a variety of meteorological phenomena, such as convection, frontal systems, shear instability, flow over orography. Gravity waves have small amplitudes in the troposphere, therefore their role in tropospheric meteorology is at first order negligible. However, because of the air compressibility, their amplitude increases with height so that at some elevation they may become nonlinear and break, interacting with the large scale flow. Linear gravity waves with phase speeds close to that of the large scale flow also affect the large scale flow itself, by being easily dissipated.

In the extra-tropics, breaking and dissipation of gravity waves in the upper mesosphere are responsible for the local reversal of the temperature gradient and in addition influence the temperature lower down, namely in the polar winter stratosphere (Haynes et al. 1991). A global effect of the gravity waves is their contribution to the dynamical forcing of the upper branches of the Brewer - Dobson circulation, the large scale circulation that transports chemical constituents and aerosols (McIntyre 1992; Garcia and Boville 1994).

The status of middle atmosphere modelling for a number of GCMs with tops at 1 hPa or lower pressure has been recently summarized by Pawson et al. (2000). In Figure 1 (from Pawson et al. 2000), the January zonal mean zonal wind is shown for observations and a number of GCMs that participated in the Pawson et al model intercomparison. Comparison with the observations (top left) indicates that most models are characterized by strong westerlies (by thermal wind balance associated with a cold bias in the temperature field) in the winter stratosphere (northern hemisphere). The excessive strength of the westerlies is believed to be related to a poor representation of the effects of gravity waves, e.g. limitations in the gravity wave parameterizations.

In the tropics, gravity waves play a role in forcing zonal wind changes by momentum flux deposition. In addition, convection is important because of its role in generating both the gravity waves as well as a realistic spectrum of equatorial planetary waves (see for instance Ricciardulli and Garcia, 2000 and Garcia this volume). Generally, horizontal resolution in GCMs is too low for a convective parameterization to generate gravity waves. However, vertically propagating

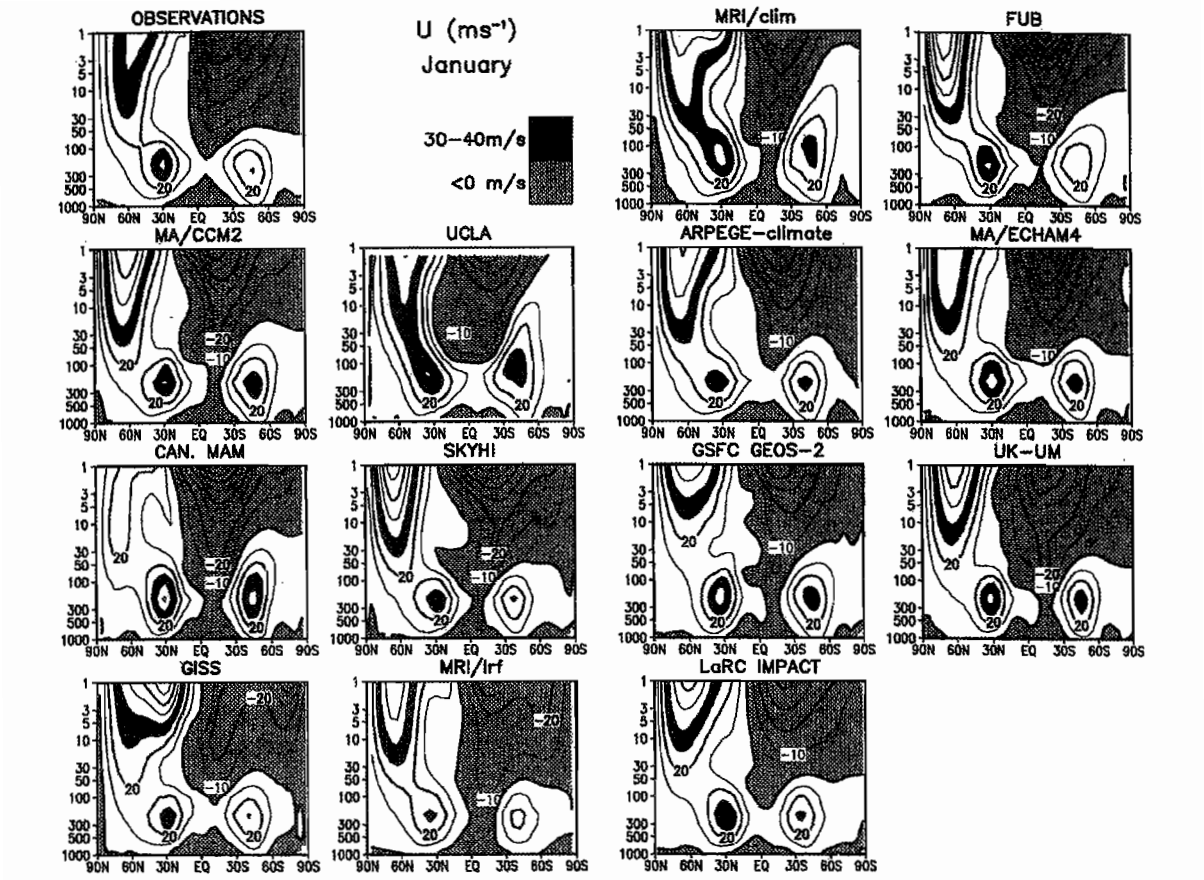


Figure 1:
January multi year average, zonal mean zonal wind (ms^{-1}) from observations (NCEP CPC analysis) and 13 models. Contour interval: 10 ms^{-1}. Easterlies (negative values) are light shaded and westerlies (between 30-40 ms^{-1}) are dark shaded (from Pawson et al., 2000).

inertia gravity waves may be forced by at least some convective schemes, such as the convective adjustment in the SKYHI model (Manzini and Hamilton 1993).

The zonal mean zonal wind at the Equator for a number of models is shown in Figure 2 (from Amodei et al. 2001). In general the models show some success in the simulation of the semiannual oscillation in zonal wind at the stratopause, mostly as a modulation of the strength of the easterlies, while only one model (EUGCM) has alternating westerlies at 1 hPa or lower pressure. However, all of them do not simulate the quasi-biennial oscillation (QBO) in zonal winds in the lower stratosphere (e.g. the low frequency variation from -30 to 10 ms^{-1} seen in the analysis, top panel, between 10 and 100 hPa). The simulation of westerlies in the equatorial upper stratosphere is sensitive to both the equatorial Kelvin waves (planetary scale) and the gravity waves. The sensitivity of the simulation of the semiannual oscillation in zonal wind at

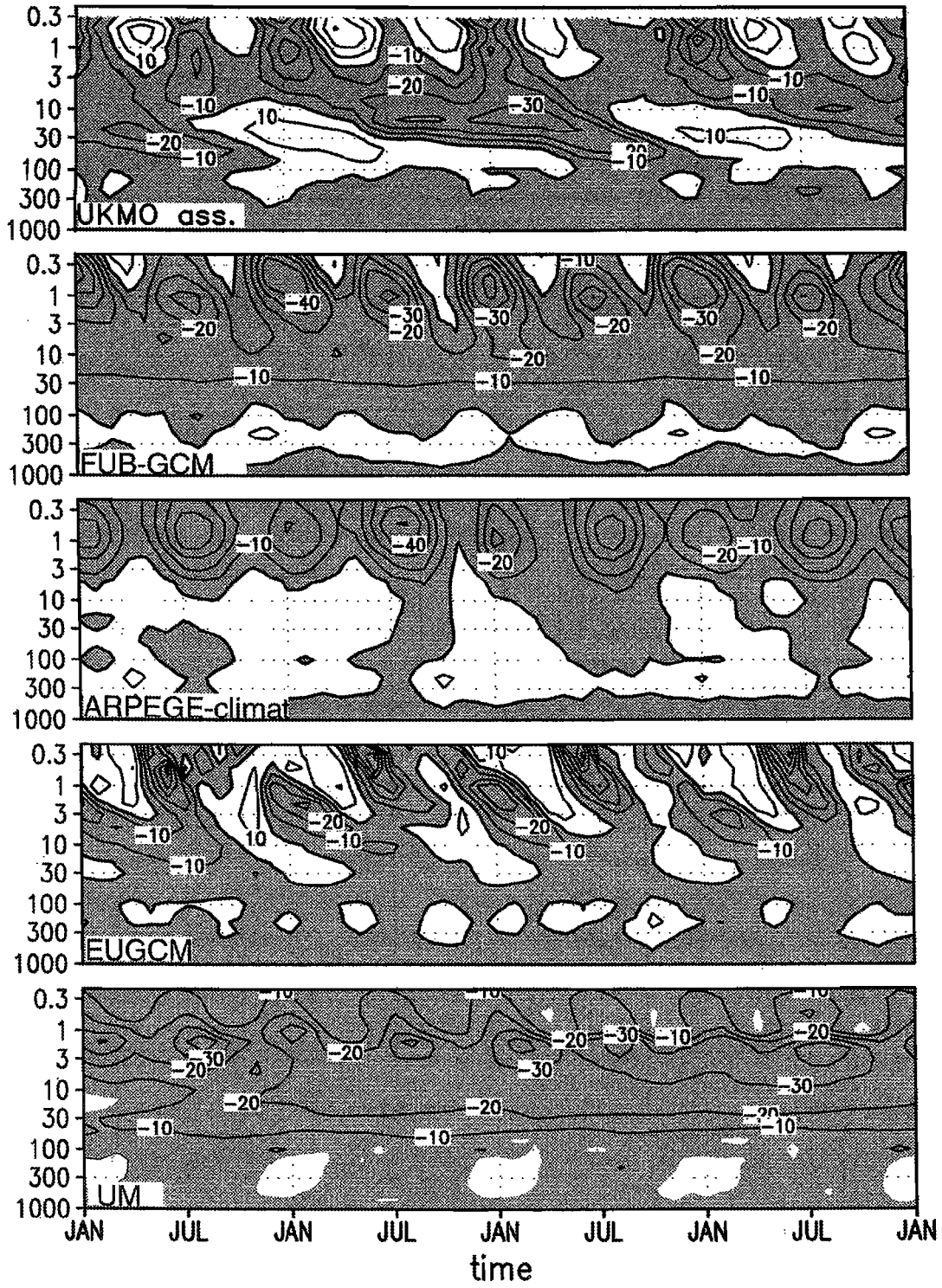


Figure 2:

Equatorial zonal mean zonal wind (ms^{-1}) form observations (UKMO assimilation from 1992 to 1996, top panel) and 4 models. The vertical axis is in hPa. Contour interval: 10 ms^{-1}. Easterlies (negative values) are shaded (from Amodei et al. 2001).

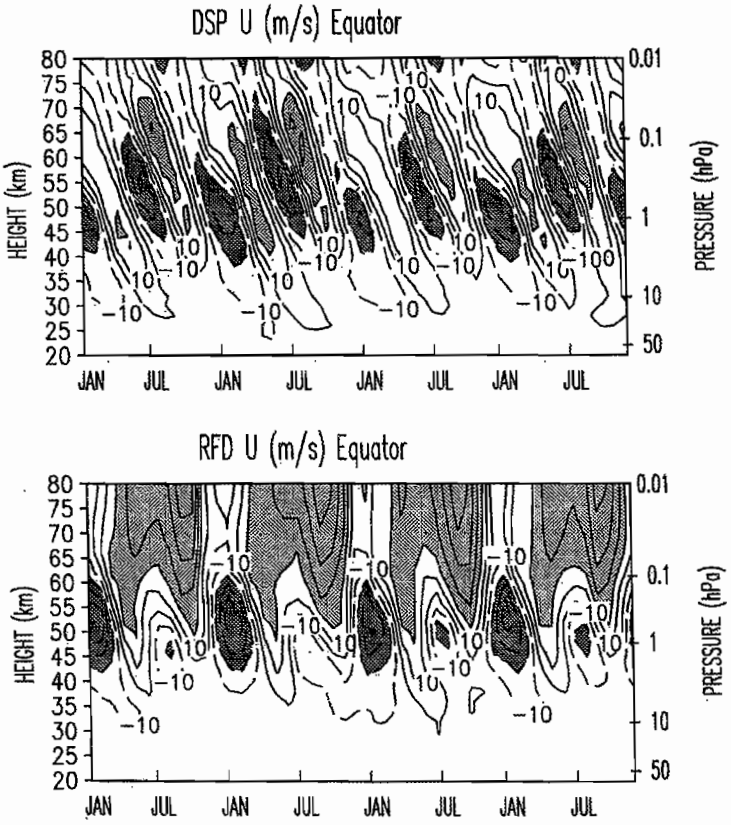


Figure 3:
Equatorial zonal mean zonal wind (ms^{-1}) from the MAECHAM4 model with the Hines gravity wave parameterization and no Rayleigh drag (top) and with Rayleigh drag only (bottom). Contours interval: 10 ms^{-1}. (light shading: > 20 ms^{-1}; dark shading: < -20 ms^{-1}) (from Manzini et al. 1997).

the stratopause to the parameterization of the momentum flux deposition from a gravity wave spectrum is shown for instance in Figure 3 (from Manzini et al. 1997). The zonal wind in the bottom panel is from a model version with Rayleigh friction (linear damping of the simulated winds in the upper mesosphere), while a parameterization of the momentum flux deposition from a broad spectrum of waves is included in the simulation shown in the upper panel. Alternating easterlies and westerlies are present in both simulations, however the westerlies and their downward propagation are enhanced in the simulation with the gravity wave parameterization. Note also that the only model that has westerlies in Figure 2 also includes a gravity wave parametrization with waves with positive phase speed.

Instead of parameterizing gravity wave effects, an alternative approach is to explicitly resolve the waves and their effects on the large scale flow by integrating the primitive equations at

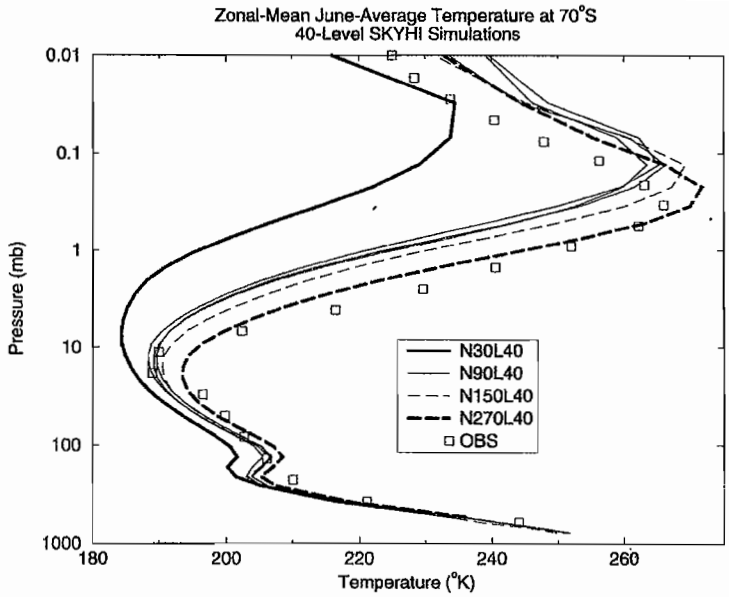


Figure 4:
June zonal mean temperature at 70°S from observations (climatology of Fleming et al. 1988, squares) and from SKYHI L40 model at four different horizontal resolutions. Results from N30 (heavy solid curve) are for a 10 year average, while the other curves are individual monthly means (for N90, 3 individual Junes are shown) (from Hamilton et al. 1999).

extremely high horizontal and vertical resolutions. A systematic analysis of the sensitivity of the SKYHI model to further increases in both the horizontal and vertical resolutions is reported by Hamilton et al. (1999). They have shown that the extratropical circulation is sensitive to horizontal resolution, while the representation of low frequency zonal wind variations at the Equator is sensitive to vertical resolution. Figure 4 (from Hamilton et al. 1999) shows a noticeable improvement with horizontal resolution in the SKYHI simulation of the temperature structure at high latitudes, with virtually no cold bias in the temperature field for the highest resolution employed (N270, where 270 is the number of grid rows between the pole and the Equator, about a third of a degree of horizontal resolution).

Figure 5 (from Hamilton et al. 1999) shows the equatorial zonal winds, as the SKYHI vertical resolutions is increased from 40 to 80 levels (between the ground and ~80 km). As to some extent for the models previously shown (e.g. Figures 2 and 3) the simulation at low vertical resolution (upper panel) presents a semiannual oscillation in the upper stratosphere (~1 hPa) and weak uniform winds lower down. For L80 (middle panel), large zonal wind variations (-30 to 30 ms^{-1}) spontaneously appear. The results in the bottom panel are for the high resolution model, with the annual cycle turned off (e.g., constant equinoctial condition), to verify that the zonal wind variations are not locked to the annual cycle, given that their typical period is close to one year, instead of quasi-biennial. To run a GCM at very high resolution is however

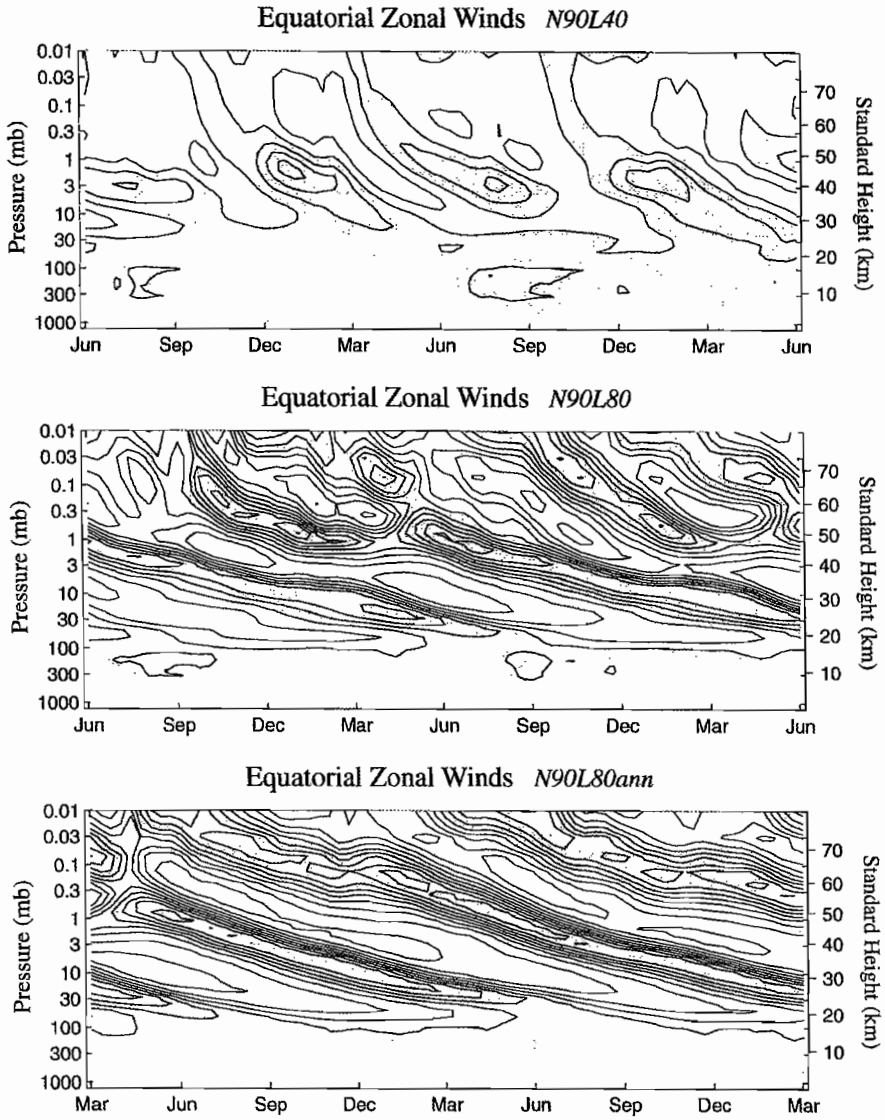


Figure 5:

Equatorial zonal mean zonal wind (ms^{-1}) from the SKYHI N90 model at different vertical resolution: L40 (top) and L80 (middle). Bottom panel is again at L80, but from a simulation without the annual cycle (e.g., perpetual equinox). Contour interval: 10 ms^{-1}. Easterlies (negative values) are light shaded (from Hamilton et al. 1999).

computationally quite expensive and it is not yet possible to run such models for the time scale of interest to climate studies. An additional limitation is that, although such simulations have shown the effects of a broader spectrum of gravity waves on the middle atmosphere, they still may not represent gravity wave processes (such as their generation) in an adequate way. Eventually, to explicitly resolve the gravity waves, GCMs will need to be based on non hydrostatic equations.

As the reported results have shown, the QBO in zonal wind in the lower stratosphere is proving a hard task for standard GCMs. In addition to problems related to the representation of gravity waves and the large scale wave spectrum generated by the convection (Ricciardulli and Garcia 2000), the simulation of the QBO in models may be affected by lack of horizontal and vertical resolutions (see Figure 5), the amount of dissipation present in the lower stratosphere and by the strength of the residual circulation in the lower tropical stratosphere (Dunkerton 1997). To date, there has been some success in the simulation of QBO like oscillations by modifying the GCMs so that to increase the resolved wave spectrum (Takahashi 1996) or by using a gravity wave parameterization properly tuned (Scaife et al. 2000). Recent results (Giorgetta et al. 2002) indicate that a combination of such factors as fine spacial resolution, well behaved convection and some amount of gravity waves are all essential factors for obtaining a realistic QBO.

With the further upward extension of GCMs to the lower thermosphere, a new challenge to the modelling of tropical dynamics will be the simulation of the full range of zonal wind oscillations, from the QBO in the lower stratosphere, to the semiannual oscillations at the stratopause and mesopause, to the mesospheric low frequency (also quasi biennial) oscillation detected by satellite observations (for a review see Baldwin et al. 2001). The realization of such a range of oscillations could be an ultimate test on the wave field simulated by GCMs.

Appendix 1: Links related to GCMs and climate research:

- On line documentation of the NCAR CCM3 model (Boulder, CO, USA):
http://www.cgd.ucar.edu/cms/ccm3/ccm3lsm_doc/

- Climate research at GFDL (Princeton, NJ, USA):
http://www.gfdl.noaa.gov/~kd/ClimateDynamics/climate.html

- General circulation model development at GISS (New York, NY, USA):
http://www.giss.nasa.gov/research/modeling/gcms.htm

- Model developments at Max Planck Institute (Hamburg, Germany):
http://www.mpimet.mpg.de/working_groups/wg2/index.html

- Hadley Centre for Climate Prediction and Research (Bracknell, UK):
http://www.meto.govt.uk/research/hadleycentre/

- Climate research at Laboratoire de Meteorologie Dynamique (Paris, France):
http://www.lmd.jussieu.fr/en/Climat/LMD_Climat_frame.html

- Atmospheric Model Intercomparison Project:
http://www-pcmdi.llnl.gov/amip/amiphome.html

Acknowledgements

The author would like to thank M. A. Giorgetta and H. Schmidt for reading the manuscript and for their constructive comments.

The following publishers and authors are gratefully acknowledged for permission to reprint the previously published figures:

American Meteorological Society:

Figures 5 and 14 of Hamilton, K., R.J. Wilson, and R.S. Hemler, 1999: Middle Atmosphere simulations with high vertical and horizontal resolution version of a GCM: Improvements in the cold pole bias and generation of a QBO-like oscillation in the tropics. *J. Atoms. Sci.*, **56**, 3829-3856.

Figure 3 of Pawson, S., K. Kodera, K. Hamilton, T. G. Shepherd, S. R. Beagley, B. A. Boville, J. D. Farrara, T. D.A. Fairlie, A. Kitoh, W. A. Lahoz, U. Langematz, E. Manzini, D. H. Rind, A. A. Scaife, K. Shibata, P. Simon, R. Swinbank, L. Takacs, R. J. Wilson, J. A. Al-Saadi, M. Amodei, M. Chiba, L. Coy, J. de Grandpre, R. S. Eckman, M. Fiorino, W. L. Grose, H. Koide, J.N. Koshyk, D. Li, J. Lerner, J. D. Mahlman, N. A. McFarlane, C. R. Mechoso, A. Molod, A. O'Neill, R. B. Pierce, W. J. Randel, R. B. Rood, and F. Wu, 2000: The GCM-Reality Intercomparison Project for SPARC: Scientific Issues and Initial Results, *Bull. Am. Meteorol. Soc.*, **81**, 781-796.

American Geophysical Union:

Figure 5 of Manzini, E., N.A. McFarlane, and C. McLandress, 1997: Impact of the Doppler Spread Parameterization on the simulation of the middle atmosphere circulation using the MA/ECHAM4 general circulation model. *J.Geophys. Res.*, **102**, 25751-25762.

European Geophysical Society:

Figure 2 of Amodei, M., S. Pawson, A.A. Scaife, U. Langematz, W. Lahoz, D.M. Li, and P. Simon, 2001: The SAO and Kelvin waves in the EuroGRIPS GCMs and the UK Met. Office analyses. *Ann. Geophys.*, **19**, 99-114.

References

Amodei, M., S. Pawson, A.A. Scaife, U. Langematz, W. Lahoz, D.M. Li, and P. Simon, 2001: The SAO and Kelvin waves in the EuroGRIPS GCMs and the UK Met. Office analyses. *Ann. Geophys.*, **19**, 99-114.

Arakawa, A., 1966: Computational design for long term numerical integration of the equation of fluid motion: Two-dimensional incompressible flow. Part 1. *J. Comp. Phys.*, **1**, 119-143.

Baldwin, M.P., L.J. Gray, T.J. Dunkerton, K. Hamilton, P.H. Haynes, W.J. Randel, J.R. Holton, M.J. Alexander, I. Hirota, T. Horinouchi, D.B.A. Jones, J.S. Kinnersley, C. Marquardt, K. Sato, and M. Takahashi, 2001: The Quasi-biennial oscillation. *Rev. Geophys.*, **39**, 179-229.

Bourke, W., 1972: An efficient, one level primitive equation spectral model. *Mon. Wea. Rev.*, **100**, 683-689.

Bourke, W., 1974: A multi level spectral model. 1. Formulation and hemispheric integrations. *Mon. Wea. Rev.*, **102**, 687-701.

Charney, J.G., R. Fjortoft, and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237-254.

Charron, M., E. Manzini, and C.D. Warner, 2002: Intercomparison of gravity wave

parameterizations: Hines Doppler-Spread and Warner and McIntyre Ultra-Simple Schemes. *J. Meteorol. Soc. Japan*, (accepted).

Dunkerton, T.J., 1997: The role of gravity waves in the Quasi-biennial oscillation. *J. Geophys. Res.*, **102**, 26053-26076.

Fels, S.B., J.D. Mahlman, M.D. Schwarzkopf, and R.W. Sinclair, 1980: Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamical responses. *J. Atmos. Sci.*, **37**, 2265-2297.

Fleming, E.L., S. Chandra, M.R. Schoeberl, and J.J. Barnett, 1988: Monthly mean global climatology of temperature, wind, geopotential height, and pressure for 0-120 km. NASA Tech. Memo. 100697, 85 pp, Greenbelt, MD, USA.

Fomichev, V.I., W.E. Ward, S.R. Beagley, C. McLandress, J.C. McConnell, N.A. McFarlane, and T.G. Shepherd, 2002: The extended canadian middle atmosphere model: Zonal-mean climatology and physical parameterization. *J. Geophys. Res.* (in press).

Garcia, R.R., and B.A. Boville, 1994: 'Downward Control' of the mean meridional circulation and temperature distribution of the polar winter stratosphere, *J. Atoms. Sci.*, **51**, 2238-2245.

Giorgetta, M.A., E. Manzini, and E. Roeckner, 2002: Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves. *Geophys. Res. Lett.* (submitted).

Hamilton, K., R.J. Wilson, and R.S. Hemler, 1999: Middle Atmosphere simulations with high vertical and horizontal resolution version of a GCM: Improvements in the cold pole bias and generation of a QBO-like oscillation in the tropics. *J. Atoms. Sci.*, **56**, 3829-3856.

Haynes, P.H., C.J. Marks, M.E. McIntyre, T.G. Shepherd, and K.P. Shine, 1991: On the 'Downward Control' of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.*, **48**, 651-678.

Hines, C.O., 1997a: Doppler spread parameterization of gravity wave momentum deposition in the middle atmosphere. Part 1: Basic formulation, *J. Atmos. Solar Terr. Phys.*, **59**, 371-386.

Hines, C.O., 1997b: Doppler spread parameterization of gravity wave momentum deposition in the middle atmosphere. Part 2: Broad and quasi monochromatic spectra and implementation, *J. Atmos. Solar Terr. Phys.*, **59**, 387-400.

Hoskins, B.J., and A.J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.*, **101**, 637-655.

Kasahara, A., 1974: Various vertical coordinate system used for numerical weather prediction. *Mon. Wea. Rev.*, **102**, 509-522.

Lindzen, R.S., 1981: Turbulence and stress due to gravity wave breakdown. *J. Geophys. Res.*, **86**, 9707-9714.

Machenhauer, B., 1979: The spectral method. In *Numerical methods used in atmospheric models. GARP Publication series,* **17**, 121-275, WMO, Geneva, Switzerland.

Manzini, E., and K. Hamilton, 1993: Middle atmospheric travelling waves forced by latent and convective heating. *J. Atmos. Sci.*, **50**, 2180-2200.

Manzini, E., and N.A. McFarlane, 1998: The effect of varying the source spectrum of a gravity wave parameterization in a middle atmosphere general circulation model. *J. Geophys. Res.*, **103**, 31523-31539.

Manzini, E., N.A. McFarlane, and C. McLandress, 1997: Impact of the Doppler Spread Parameterization on the simulation of the middle atmosphere circulation using the MA/ECHAM4 general circulation model. *J.Geophys. Res.*, **102**, 25751-25762.

McIntyre, M.E., 1992: Atmospheric dynamics: Some fundamentals, with observational implication. *Int. School Phys 'Enrico Fermi' CXV Course. The use of EOS for studies of atmospheric physics* (Eds J.C. Gille and G. Visconti), North-Holland, pp 313-386.

Medvedev, A.S., and G.P Klaassen, 1995: Vertical evolution of gravity wave spectra and the parameterization of associated wave drag, *J. Geophys. Res.*, **100**, 25841-25853.

Medvedev, A.S., G.P Klaassen, and S.R. Beagley, 1998: On the role of an anisotropic gravity wave spectrum in maintaining the circulation of the middle atmosphere, *Geophys. Res. Lett.*, **25**, 509-512.

Orzag, S.A., 1970: Transform method for the Calculation of vector coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.*, **27**, 890-895.

Pawson, S., K. Kodera, K. Hamilton, T. G. Shepherd, S. R. Beagley, B. A. Boville, J. D. Farrara, T. D.A. Fairlie, A. Kitoh, W. A. Lahoz, U. Langematz, E. Manzini, D. H. Rind, A. A. Scaife, K. Shibata, P. Simon, R. Swinbank, L. Takacs, R. J. Wilson, J. A. Al-Saadi, M. Amodei, M. Chiba, L. Coy, J. de Grandpŕe, R. S. Eckman, M. Fiorino, W. L. Grose, H. Koide, J.N. Koshyk, D. Li, J. Lerner, J. D. Mahlman, N. A. McFarlane, C. R. Mechoso, A. Molod, A. O'Neill, R. B. Pierce, W. J. Randel, R. B. Rood, and F. Wu, 2000: The GCM-Reality Intercomparison Project for SPARC: Scientific Issues and Initial Results, *Bull. Am. Meteorol. Soc.*, **81**, 781-796.

Pexioto, J.P., and A.H. Oort, 1992: *Physics of climate*. American Institute of Physics, New York, New York, USA.

Phillips, N.A., 1957: A coordinate system having some special advantages for numerical forecasting. *J. Meteor.*, **14**, 184-185.

Ricciardulli, L., and R.R. Garcia, 2000: The excitation of equatorial waves by deep convection in the NCAR community climate model (CCM3). *J. Atmos. Sci.*, **57**, 3461-3487.

Ringler, T.D., R.P. Heikes, and D.A. Randall, 2000: Modelling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores. *Mon. Weather Rev.*, **128**, 2471-2490.

Robert, A.J., 1969: The integration of a spectral model of the atmosphere by the implicit method. Proceeding of *WMO/ICSU, Symposium on numerical weather prediction,* **S-VII**, 19-24, Tokyo, Japan.

Robert, A.J., T.L. Yee, and H. Richie, 1985: A semi-lagrangian and semi-implicit numerical integration scheme for multilevel atmospheric models. *Mon. Wea. Rev.*, **113**, 388-394.

Scaife, A.A., N. Butchart, C.D. Warner, D. Stainforth, W. Norton, and J. Austin, 2000: Realistic quasi-biennial oscillations in a simulation of the global climate. *Geophys. Res. Lett.*, **27**, 3481-3484.

Simmons, A.J., B.J. Hoskins, and D.M. Burridge, 1978: Stability of the semi-implicit method of time integration. *Mon. Wea. Rev.*, **106**, 405-412.

Simmons, A.J., and R. Strüfing, 1981: An energy and angular momentum conserving finite difference scheme, hybrid coordinates and medium range weather prediction. *ECMWF Technical Report*, **28**, Reading, England.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. 1. The basic experiment. *Mon. Wea. Rev.*, **91**, 99-104.

Smagorinsky, J., S. Manabe, and J.L. Halloway, 1965: Numerical results from a nine level circulation model of the atmosphere. *Mon. Wea. Rev.,* **93**, 727-768.

Takahashi, M., 1996: Simulation of the stratospheric quasi-biennial oscillation using a general circulation model. *Geophys. Res. Lett.*, **23**, 661-664.

Warner, C.D., and M.E. McIntryre, 2001: An ultra-simple spectral gravity wave parameterization for non-orographic gravity waves. *J. Atmos. Sci.*, **58**, 1837-1857.

Williamson, D.L., J.T. Kiehl, V. Ramanathn, R.E. Dickinson, and J.J. Hack, 1987: Description of the NCAR CCM1, *NCAR/TN-285+STR*, Boulder, Colorado, USA.