Circumpolar Vortex Studies Using MSU Temperature Data

Paul D. Reasor and Michael T. Montgomery

Research Supported by the Cooperative Institute for Research in the Atmosphere (CIRA).

Colorado State University

DEPARTMENT OF ATMOSPHERIC SCIENCE

PAPER NO. 619

CIRCUMPOLAR VORTEX STUDIES USING MSU TEMPERATURE DATA

Paul D. Reasor and Michael T. Montgomery

Research supported by the Cooperative Institute for Research in the Atmosphere (CIRA)

Principal Investigator: Michael T. Montgomery

Department of Atmospheric Science

Colorado State University

Fort Collins, CO 80523

October 1996

Atmospheric Science Paper No. 619

ABSTRACT

Irreversible mixing associated with breaking planetary waves at the edge of the circumpolar vortex exerts a westward force on the atmosphere. The tendency of this wave forcing is to decelerate the polar night jet and warm the polar regions of the stratosphere. While the tendency is opposed by an induced Eulerian-mean circulation, it is observed that substantial reductions in jet strength and departures from radiative equilibrium still occur. In the steady-state limit the departures from radiative balance at any level may be accounted for by the total wave forcing above that level. The wave forcing, and hence vortex strength, varies from year to year in a biennial fashion, suggesting a link with the equatorially-confined quasi-biennial oscillation (QBO) (Tung and Yang, 1994).

These ideas will be explored in the context of a seasonal and interannual study of the austral circumpolar vortex using brightness temperature data from the Microwave Sounding Unit (MSU). Observations of the seasonal evolution of the vortex are presented along with direct and indirect evidence of wave forcing. A 50-mb vortex circulation index derived from the MSU is shown to qualitatively correlate with the phase of the QBO in some years. It is concluded, however, that the effects of wave forcing may be too weak in the lower stratosphere to consistently observe a QBO signal at high latitudes.

ACKNOWLEDGEMENTS

This work was supported by the Cooperative Institute for Research in the Atmosphere (CIRA). Numerical computations were performed using Hewlett Packard (HP) Apollo series computers belonging to the Montgomery Research Project. Programs were written in HP FORTRAN 9000 and in the Interactive Data Language (IDL) of Research Systems, Inc.

We wish to thank Professor Graeme Stephens for encouraging us to learn more about stratospheric vortex dynamics and, in particular, to use the MSU data to explore the connection between the QBO and circumpolar vortex strength. We also would like to acknowledge Ian Wittmeyer for providing us with the MSU data and the vortex area calculations used in this thesis.

CONTENTS

1 Introduction 1

2 Circumpolar Vortex Dynamics 3
2.1 Introduction . . . 3
2.2 Circumpolar Vortex Formation . . . 4
2.3 Wave Forcing of the Circumpolar Vortex . . . 5
2.3.1 Wave-Mean-Flow Interaction . . . 7
2.3.2 Wave Breaking Dynamics . . . 14
2.4 The Stratospheric Wave Pump . . . 17
2.5 Circumpolar Vortex Interannual Variability . . . 20
2.5.1 The Quasi-biennial Oscillation . . . 25
2.5.2 Waveguides and Poleward Focusing . . . 25
2.5.3 Forcing of the Mean Flow . . . 26
2.6 Summary . . . 27

3 MSU Data 29
3.1 Introduction . . . 29
3.2 Data Source . . . 30
3.3 Comparison of MSU and ECMWF Data . . . 34
3.4 High-Latitude Data . . . 37
3.5 Summary . . . 42

4 Technique 43
4.1 Introduction . . . 43
4.2 Computation of Dynamical Quantities . . . 44
4.3 Summary . . . 48

5 Climatology of the Austral Polar Vortex 50
5.1 Introduction . . . 50
5.2 Seasonal Evolution . . . 51
5.3 Interannual Variability . . . 62
5.3.1 1985 . . . 63
5.3.2 1986 . . . 63
5.3.3 1987 . . . 64
5.3.4 1988 . . . 64
5.3.5 1989 . . . 65
5.3.6 1990 . . . 65
5.3.7 1991 . . . 66
5.3.8 1992 . . . 66

5.3.9 Polar Vortex Modulation 66
5.4 Summary 70

6 Conclusion **72**
6.1 Suggested Future Work 73

A Figures **75**

B The Spectral Method **93**

References **97**

Chapter 1

INTRODUCTION

The circumpolar vortices of the Northern and Southern Hemispheres are observed to vary interannually, sometimes with dramatic differences in strength over consecutive years. A variety of methods are used in an attempt to explain this phenomenon; some focus on the chemical transports (e.g., ozone) in the stratosphere and troposphere, others increase our understanding through climatological studies, while yet others approach the problem from a theoretical perspective, using the polar stratosphere as a "dynamical laboratory" for testing the latest theories. The different approaches appear to be converging on a common view of stratospheric variability.

The conceptual model of the stratospheric circulation as being maintained by a wave-driven extratropical "pump", recently detailed by Holton et al. (1995), provides an appropriate starting-point for a discussion of interannual variability. The evolution of the circumpolar vortex and the strength of the seasonal stratospheric circulation depend upon the intensity of wave forcing in the extratropics. Understanding variations in the vertical propagation of planetary waves is, then, the fundamental problem that must be addressed. Topographic and thermal forcing in the troposphere are believed to be the primary sources of waves in the stratosphere, suggesting that stratospheric variability may result from variability in the troposphere. Observations show, however, that a significant portion of the interannual variability in the polar stratosphere, at least for column ozone, can be tied to the tropical quasi-biennial oscillation (QBO) (Tung and Yang, 1994ab).

Observations of a QBO in the extratropics go back to work by Angell and Korshover (1963). Studies since have confirmed the existence of a QBO in both tracer fields and dynamical quantities in the extratropics of the Northern and Southern Hemispheres. Holton

and Tan (1980), for example, observed a QBO in the high-latitude geopotential heights of the Northern Hemisphere stratosphere and hypothesized that a QBO modulation of the zero zonal-mean wind line may be responsible for such an extratropical signal. The effect of this QBO modulation on the propagation of planetary waves in the midlatitudes has since been recognized as a key concept in understanding the low-frequency (quasi-biennial) variability of the extratropical stratosphere. Dunkerton and Baldwin (1991) demonstrated that the Eliassen-Palm fluxes in the Northern Hemisphere high-latitude stratosphere are more convergent during the easterly phase of the QBO, implying a more disturbed circumpolar vortex. Studies of the correlation between vortex strength and QBO phase are less numerous in the Southern Hemisphere due to the shorter record of data there. Shiotani and Hirota (1985) found, using 15 years of data, that the timing of the winter-spring jet shift in the Southern Hemisphere stratosphere is correlated with the QBO. In an 11-year climatology of the austral winter-spring stratosphere, Atkinson (1993) noted that the intensity of wave activity and the phase of the QBO qualitatively correlated for most years. We are not aware of any study that has focused on the interannual variability of the austral circumpolar vortex using primarily Microwave Sounding Unit data.

We begin this thesis with a review of the known dynamics pertinent to the stratosphere, focusing on the circumpolar vortex. The theoretical discussion is concluded with a summary of Tung and Yang's hypothesized mechanism for a QBO modulation of wave propagation in the extratropical stratosphere. Using brightness temperature data from the Microwave Sounding Unit, we develop a method for computing dynamical fields in the lower stratosphere. With eight years (1985–1992) of data we then explore the seasonal evolution and interannual variability of the austral circumpolar vortex. Observations of the seasonal evolution are shown to be consistent with the theoretical ideas presented. We conclude with observations of interannual variability, qualitatively correlating the variability in circumpolar vortex circulation with the phase of the QBO.

Chapter 2

CIRCUMPOLAR VORTEX DYNAMICS

2.1 Introduction

In recent years the circumpolar vortices of the Northern and Southern Hemispheres have undergone a great amount of scrutiny in an effort to understand not only their dynamics, but also the controls they place on the chemical transports and general circulation of the stratosphere. The role of wave dynamics, both in modulating the vortex evolution and in mixing fluid parcels near the vortex edge, has been highlighted, and continues to be an active area of research. While some progress has been made in determining the sources of the waves and describing the basic physics governing their interaction with the stratospheric flow (James, 1988; Mechoso and Hartmann, 1982; Shiotani et al., 1990), the dynamical picture is still incomplete. Interannual variations in the timing and intensity of wave propagation in the stratosphere are not yet fully understood.

This chapter reviews dynamical theory pertinent to the evolution of the circumpolar vortex, focusing on how waves drive the vortex away from a state of radiative balance. Although evidence of this forcing can be obtained directly from the vortex, the waves have a more far-reaching effect by inducing a mean meridional circulation within the stratosphere. In the steady-state limit this circulation is related to tracer transports, providing a unique means of inferring the intensity of wave forcing in the stratosphere. We conclude by summarizing Tung and Yang's (1994b) hypothesized mechanism of high-latitude interannual variability that links the phase of the quasi-biennial oscillation to the strength of wave activity in the polar stratosphere.

2.2 Circumpolar Vortex Formation

The circumpolar vortex forms as the maximum in incident solar radiation shifts from the polar regions to lower latitudes. As this transition proceeds through the fall, polar values of daily solar radiation at the top of the atmosphere decrease to zero, and a subsequent drop in temperatures takes place. In the absence of dynamical effects the temperature profile in the stratosphere is governed by a balance between solar and infrared (IR) heating by atmospheric constituents. Absorption by O_3 is responsible for a large part of the net solar heating in the stratosphere. The net longwave cooling is dominated by CO_2, and to a lesser extent, O_3 and H_2O. Where there is no solar radiation, as in the winter polar regions, the heating rate in the stratosphere depends only on emissions from the troposphere. The radiative equilibrium temperature (T_r) in this case is obtained by time-integrating the sum of the IR heating rates for each constituent until the temperature asymptotes to a steady state. Where the top of the atmosphere solar radiation is nonzero, solar heating by O_3 must also be included. Such computations have been carried out (e.g., Manabe and Moeller, 1961; Wehrbein and Leovy, 1982) and show a definite winter-to-summer hemispheric temperature gradient with a maximum gradient occurring near 60°. On the large scales under consideration a geostrophic adjustment process occurs in which the wind adjusts to the mass field. The resulting winds are in a thermal wind balance. According to the thermal wind equation (see Section 4.2) the strong gradient in temperature implies substantial westerly winds. Since the largest meridional gradients in T_r occur in the upper stratosphere, this is where the jet core should be located.

While the existence of the circumpolar vortex is accounted for by these simple considerations, the observed circumpolar vortex differs significantly from what would be predicted by radiative physics alone. The temperatures within the austral polar vortex in the lower stratosphere, for example, are approximately 10 K warmer than T_r. The discrepancy is accounted for when we include dynamical processes like planetary wave propagation into our theoretical framework. These waves, as discussed in Section 2.3, can decelerate the mean flow and produce the aforementioned warming. While both Northern and Southern Hemispheres experience some degree of warming, in the Northern Hemisphere, where

wave activity is vigorous and nearly saturates the stratosphere, the deceleration of the jet sometimes leads to a complete flow reversal (i.e., from westerlies to easterlies). Although wave activity in the Southern Hemisphere is considerably less intense and zonal flow reversals are extremely rare, the waves nevertheless play an important role in the spring-time breakdown of the circumpolar vortex. In the next section we discuss the origin of these waves and how they cause such remarkable disruptions of the radiatively driven vortex.

2.3 Wave Forcing of the Circumpolar Vortex

Observations show that planetary-scale Rossby waves play an important role in governing the evolution of the austral polar vortex from its formation in early winter to its gradual break-up in late spring. While there still exists some debate as to the origin of all the waves in the winter stratosphere, it is an established fact that the dominant quasi-stationary wavenumber one component is produced in the troposphere. Combined thermal and topographic forcing are believed to be the primary sources (Quintanar and Mechoso, 1995). The propagation of these large-scale forced waves can be modeled quite simply by the frictionless, adiabatic quasi-geostrophic (QG) equations on a β-plane in log-p coordinates

$$\frac{\partial u_g}{\partial t} + u_g\frac{\partial u_g}{\partial x} + v_g\frac{\partial u_g}{\partial y} - f_0 v_a - \beta y v_g = 0, \tag{2.1}$$

$$\frac{\partial v_g}{\partial t} + u_g\frac{\partial v_g}{\partial x} + v_g\frac{\partial v_g}{\partial y} + f_0 u_a + \beta y u_g = 0, \tag{2.2}$$

$$\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y} + \frac{1}{\rho_0}\frac{\partial(\rho_0 w)}{\partial z} = 0, \tag{2.3}$$

$$\frac{\partial \theta}{\partial t} - u_g\frac{\partial \theta}{\partial x} + v_g\frac{\partial \theta}{\partial y} + w\theta_{0z} = 0, \tag{2.4}$$

where θ_0 is the reference potential temperature, and the 'g' and 'a' subscripts denote geostrophic and ageostrophic wind components, respectively. Throughout the rest of this thesis, subscripts of x, y, and z will signify derivatives with respect to that coordinate, unless otherwise stated.

Equations (2.1)–(2.4) can be combined into the QG potential vorticity (QG PV) equation,

$$\frac{\partial q_g}{\partial t} + u_g\frac{\partial q_g}{\partial x} + v_g\frac{\partial q_g}{\partial y} = 0, \tag{2.5}$$

where the QG PV is given by

$$q_g = f_0 + \beta y + \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} + \frac{f_0}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{\theta_{0z}}\theta\right). \tag{2.6}$$

Equation (2.5) expresses the material conservation of q_g following the geostrophic wind in frictionless adiabatic flow. Upon defining a streamfunction such that

$$u_g = -\frac{\partial \psi}{\partial y}, \qquad v_g = \frac{\partial \psi}{\partial x} \tag{2.7}$$

and

$$\frac{\theta}{\theta_{0z}} = \frac{f_0}{N^2}\frac{\partial \psi}{\partial z} \qquad \text{(hydrostatic equation)}, \tag{2.8}$$

the QG PV can be written as

$$q_g = f_0 + \beta y + \nabla^2 \psi + \frac{f_0^2}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0}{N^2}\frac{\partial \psi}{\partial z}\right), \tag{2.9}$$

where N is the Brunt Väisällä frequency, and will be assumed constant in the discussion to follow.

Linearizing (2.5) about a zonal, time-independent basic state, and substituting in solutions of the form

$$\psi' = \psi_0 e^{z/2H} e^{i(kx+ly+mz-\omega t)} \tag{2.10}$$

results in the following dispersion relation for vertically propagating Rossby waves

$$\omega = k\bar{u}_g - k\bar{q}_y / \left[\frac{f_0^2}{N^2}\left(m^2 + \frac{1}{4H^2}\right) + k^2 + l^2\right], \tag{2.11}$$

where H is the vertical scale height and $\bar{q}$ is the basic state PV. $\bar{q}$ incorporates planetary vorticity as well as the vorticity of the background flow. Its meridional derivative is generally positive in the stratosphere due to the former contribution, although exceptions do exist, especially poleward of the polar night jet (PNJ).

We can use (2.11) to diagnose where vertical propagation is limited or entirely prohibited. If the PNJ is so strong that the waves are carried along with the westerly mean flow, or the wavelengths are small enough to preclude significant interaction of the PV anomalies with the surrounding fluid, then it is possible that vertical propagation of stationary waves (i.e., $\omega = 0$) of the type previously discussed might be inhibited at high latitudes.

Also notice that if the basic state flow is easterly, vertical propagation is prohibited, since Rossby waves can only propagate westward relative to the mean flow. During the austral winter this is not a major consideration in the polar regions, as flow reversals from westerly to easterly are rare, unlike in the Northern Hemisphere. We do, however, find easterlies in the tropical stratosphere during the easterly phase of the QBO. Vertically propagating Rossby waves will tend to avoid this region, as discussed in Section 2.5.2.

In general the basic state of the winter Southern Hemisphere is quite favorable for the propagation of forced planetary Rossby waves from the troposphere into the stratosphere. To investigate how these waves interact with the stratospheric flow to produce the observed deviations from radiative equilibrium, we first consider the classical approach of separating the flow into its zonal-mean and eddy (i.e., wave) components. Then, we briefly discuss an alternate method using the quasi-conservative PV to diagnose the influence of the waves on the circumpolar vortex. The latter method, while having the advantage of not making the sometimes inappropriate wave-mean-flow separation, is inherently the same as the former.

2.3.1 Wave-Mean-Flow Interaction

Upon zonally averaging (2.1)–(2.4), we find that both the momentum and thermodynamic equations contain eddy terms. Obviously a flux of momentum into or from the zonal-mean flow will have a tendency to accelerate or decelerate it. Additionally, the flux of heat associated with the waves will alter the thermal structure of the stratosphere, and, if we assume that the stratosphere is in a thermal wind balance, the zonal-mean flow. In this framework it is not straightforward to diagnose the effect of waves on the zonal-mean flow. To make explicit the eddy forcing of the zonal-mean flow, we transform the zonally averaged equations. Including a diabatic heating term in the thermodynamic equation and defining (Andrews et al., 1987)

$$\bar{v}^* \equiv \bar{v}_a - {\rho_0}^{-1}(\rho_0 \overline{v_g' \theta'}/\theta_{0z})_z \tag{2.12}$$

$$\bar{w}^* \equiv \bar{w}_a + (\overline{v_g' \theta'}/\theta_{0z})_y, \tag{2.13}$$

the following transformed Eulerian-mean (TEM) equations result:

$$\frac{\partial \bar{u}_g}{\partial t} - f_0 \bar{v}^* = \rho_0^{-1} \nabla \cdot \mathbf{F}, \tag{2.14}$$

$$\frac{\partial \bar{\theta}}{\partial t} + \bar{w}^* \theta_{oz} = \bar{J}/c_p, \tag{2.15}$$

$$\frac{\partial \bar{v}^*}{\partial y} + \rho_0{}^{-1} (\rho_0 \bar{w}^*)_z = 0, \tag{2.16}$$

$$f_0 \frac{\partial \bar{u}_g}{\partial z} + H^{-1} R e^{-\kappa z/H} \frac{\partial \bar{\theta}}{\partial y} = 0, \tag{2.17}$$

where the vector on the right-hand side of the zonal momentum equation, called the Eliassen-Palm (E-P) flux, is given by

$$\mathbf{F} \equiv \left(0, -\rho_0 \overline{v_g' u_g'}, \rho_0 f_0 \overline{v_g' \theta'} / \theta_{oz}\right). \tag{2.18}$$

This transformation has the benefit of combining the explicit eddy terms into a wave forcing in the momentum equation, demonstrating the fact that both the eddy momentum and heat fluxes contribute to the forcing of the zonal-mean flow in a prescribed manner. For example, one might naively assume from the non-transformed zonal momentum equation that a non-zero momentum flux divergence implies a change in the zonal-mean flow. According to the TEM equations, however, a heat flux divergence could, in fact, exactly cancel this change, resulting in zero net forcing.

The transformed vertical velocity, while mathematically convenient for the purposes of obtaining (2.14), also has a useful physical interpretation. As Rossby waves propagate vertically into the stratosphere, they transport heat poleward. To maintain thermal wind balance, an Eulerian-mean meridional circulation with a corresponding streamfunction, $\rho_0 \overline{v_g' \theta'} / \theta_{0z}$, is induced which tends to counteract this transport by fluxing heat downward and equatorward. In the stratosphere it is observed that these two heat transport mechanisms tend to cancel each other. In the idealized case of exact cancellation there is no net transport of heat into the polar stratosphere, resulting in a thermal field that is in a state of radiative equilibrium. The actual Eulerian-mean circulation is then equivalent to the wave-induced circulation, and from (2.13) $\bar{w}^*$ is zero. In certain situations, however, the wave-induced circulation does not offset the transport of heat by the waves, resulting

in a net transport of heat into the polar stratosphere. The residual circulation accounting for this contribution to adiabatic temperature change is $\bar{w}^*$. In the steady-state limit, this exactly equals the zonally-averaged diabatic circulation. Therefore, on seasonal timescales we may infer the cross-isentropic mass transport of various chemical species, like ozone, from computations of $\bar{w}^*$.

Returning to (2.14), we see that the effect of the waves is to either accelerate or decelerate the zonal-mean flow. If, however, the eddy forcing term on the right-hand side of the zonal momentum equation vanishes, then the waves will have no effect on the zonal mean (although they still can have an effect on the *time-mean* flow). This can be shown explicitly by constructing an elliptic equation for the residual streamfunction. The left-hand side takes on the appearance of a two-dimensional Laplacian operator acting on the residual streamfunction, while the right-hand side contains the wave and diabatic forcings. On the boundaries of the domain we assume the residual streamfunction to be zero, consistent with vanishing $\rho_0\bar{w}^*$ and $\bar{v}^*$ there. Under the assumption that $\nabla \cdot \mathbf{F}$ and $\bar{J}$ equal zero, the streamfunction equation reduces to Laplace's equation. By the uniqueness theorem, any solution satisfying the prescribed boundary condition is the only solution. It follows that a residual streamfunction of zero is the solution to our boundary-value problem. Furthermore, $\bar{w}^*$ and $\bar{v}^*$ must also be zero everywhere. The TEM momentum equation then reduces to the statement that $\bar{u}_g$ is time-invariant. A more elegant statement of this non-acceleration theorem, originally described by Charney and Drazin (1961), is given by the generalized Eliassen-Palm theorem (Andrews and McIntyre, 1976):

$$\frac{\partial A}{\partial t} + \nabla \cdot \mathbf{F} = D + O(\alpha^3), \tag{2.19}$$

where $A \equiv \rho_0 \overline{q'^2}/2\bar{q}_y$ is the wave-activity density, D the nonconservative forcing, and α the wave amplitude. The wave-activity density is a measure of wave amplitude (see Andrews et al., 1987 or James, 1994). The last term on the right-hand side of (2.19) represents the nonlinear effects. For linear, conservative, steady waves the E-P flux divergence is identically zero, and, as previously stated, the zonal-mean flow is unaffected.

This E-P flux diagnostic has a particularly interesting interpretation under the QG scaling or, more generally, for slowly varying basic states in the WKBJ sense. Assuming

wavelike solutions (2.10), the momentum and heat fluxes for the vertically propagating Rossby waves are given by

$$\overline{v_g' u_g'} = -\frac{kl}{2}|\psi_0|^2 e^{z/H}, \qquad \overline{v_g' \theta'} = \frac{km}{2}|\psi_0|^2 \frac{f_0 \theta_0}{g} e^{z/H}, \tag{2.20}$$

where ψ_0 is the wave amplitude. Also,

$$\overline{q'^2} = \frac{1}{2}\left[k^2 + l^2 + (f_0/N)^2(m^2 + 1/4H^2)\right]^2 |\psi_0|^2 e^{z/H}. \tag{2.21}$$

Using the dispersion relation, (2.11), to compute the y and z components of the group velocity, and combining these expressions with (2.20), we find

$$\mathbf{c_g} = \frac{4\bar{q}_y e^{-z/H}}{\rho_0 \left[k^2 + l^2 + (f_0/N)^2(m^2 + 1/4H^2)\right]^2 |\psi_0|^2} \left(0, -\rho_0 \overline{v_g' u_g'}, \rho_0 f_0 \overline{v_g' \theta'}/\theta_{0z}\right). \tag{2.22}$$

The group velocity, as it has been defined in (2.22), indicates the direction of wave energy propagation in the y-z plane.

Substituting (2.21) into the definition of wave-activity density, we discover that the factor multiplying the vector expression in (2.22) is just the inverse of the wave-activity density. Furthermore, according to (2.18), the vector expression is the E-P flux. Therefore, (2.22) can be written as

$$\mathbf{F} = A\mathbf{c_g}. \tag{2.23}$$

So, under the conditions stated, the E-P flux vectors point in the direction of wave energy propagation. This fact proves useful in the real atmosphere on planetary scales where the QG approximation is valid. Given maps of the E-P flux vector, one can determine those regions most favorable for propagation, as well as where wave energy is being deposited. An example is shown in Fig. 2.1. In the troposphere of the Southern Hemisphere the E-P flux vectors are vertically aligned, signifying a poleward transport of heat, and of largest magnitude between 40° and 50°S (Mechoso et al., 1985). The waves responsible for this transport of heat result from strong baroclinic instability. At higher levels in the troposphere the E-P flux vectors bend equatorward, signifying a poleward transport of zonal momentum. As McIntyre (1982) points out, this is a result of the spherical geometry, and is in some sense a "defocusing" of the waves. A similar equatorward bending of the

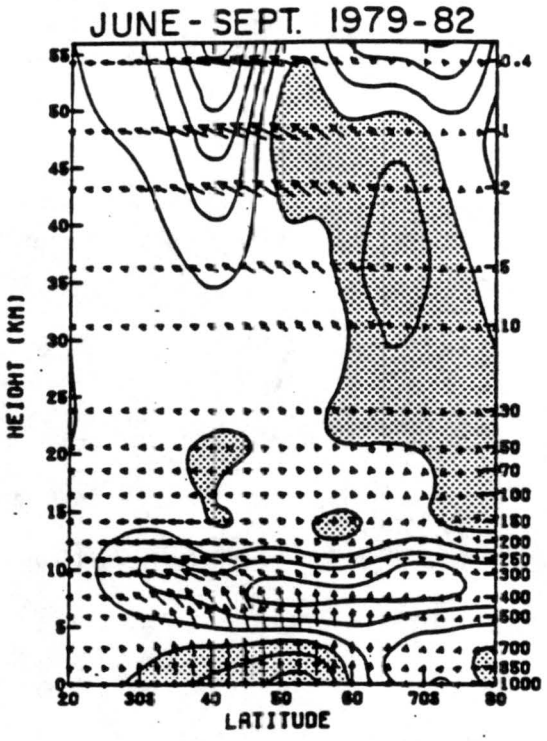


Figure 2.1: Eliassen-Palm cross-section in the Southern Hemisphere. Shaded regions depict positive E-P flux divergence. Contour interval 2 $ms^{-1}day^{-1}$ (From Mechoso et al. (1985)).

E-P flux vectors occurs in the middle and upper stratosphere. Around 40°S they strongly converge, providing evidence of negative wave forcing at the vortex edge. Note that this wave forcing decreases from the upper to lower stratosphere.

Substituting (2.23) into (2.19), we find that the amplitude of linear conservative waves is governed by:

$$\frac{\partial A}{\partial t} + \mathbf{c_g} \cdot \nabla A = -A \nabla \cdot \mathbf{c_g}. \tag{2.24}$$

Equation (2.24) shows that in regions of energy convergence (i.e., $\nabla \cdot \mathbf{c_g} < 0$), the wave amplitude following the group velocity will increase. One way of converging wave energy is by narrowing the space through which the waves can propagate. A real-atmosphere example of this phenomenon is discussed in Section 2.5 as it pertains to the interannual variability of the circumpolar vortex.

To complete our discussion of wave-mean-flow interaction, we would like to explicitly show the relationship between wave forcing and departures from radiative equilibrium in

the stratosphere. In terms of zonal-mean temperature (2.15) becomes:

$$\frac{\partial \overline{T}}{\partial t} + N^2 H R^{-1} \bar{w}^* = \bar{J}/c_p. \tag{2.25}$$

For seasonal time periods the transient terms can be neglected. Combining (2.14), (2.16), and (2.25) we obtain:

$$-\frac{\partial}{\partial y}\nabla \cdot \mathbf{F} + f_0 \frac{\partial}{\partial z}\left(\frac{\rho_0 \kappa \bar{J}}{HN^2}\right) = 0 \tag{2.26}$$

To simplify (2.26) we must investigate the physics behind the diabatic heating rate, J. For CO_2 and H_2O in the stratosphere, radiative exchange between layers is negligible, so IR cooling is well approximated by the "cooling-to-space" term in the radiative transfer equation (Rodgers and Walshaw, 1966). The longwave heating rate per unit mass for a particular spectral band, $\Delta\nu$, is then given by

$$J_{LW,\Delta\nu}(T(z)) \approx -\frac{\pi}{\rho_0(z)} B_{\Delta\nu}(T(z)) \frac{d\bar{T}_f}{dz}(z, \infty), \tag{2.27}$$

where $B_{\Delta\nu}$ is the Planck blackbody function,

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1}, \tag{2.28}$$

integrated over $\Delta\nu$ and $\bar{T}_f$ is the transmission from level z to the top of the atmosphere. In Eq. (2.28) h is Planck's constant, k is Boltzmann's constant, c is the speed of light, and T is the absolute temperature.

Now define a radiative-photochemical equilibrium temperature, T_r, such that

$$J_{SW} + J_{LW}(T_r) = 0, \tag{2.29}$$

where J_{SW} and J_{LW} denote the *total* solar and longwave heating, respectively. Also note that the temperature dependence of the solar heating has been neglected, excluding processes like photochemical acceleration[1]. Thus, the radiative equilibrium temperature

---

[1]The inclusion of the coupling between ozone concentration and temperature will decrease the ozone heating, resulting in an enhanced damping rate. This phenomenon is most noticeable above 35 km, where ozone heating is a maximum (Lindzen and Goody, 1965).

is defined such that shortwave heating and longwave cooling exactly cancel each other. The net radiative heating rate per unit mass is then given by

$$J \equiv J_{SW} + J_{LW}(T) = -J_{LW}(T_r) + J_{LW}(T). \tag{2.30}$$

Considering small departures, δT, from radiative equilibrium, the band-integrated Planck function can be Taylor expanded about T_r:

$$B_{\Delta\nu}(T) = B_{\Delta\nu}(T_r) + \left(\frac{dB_{\Delta\nu}}{dT}\right)_{T=T_r} \delta T + \cdots \tag{2.31}$$

Substituting (2.31) to first order into (2.27), summing over all significant spectral bands, and using this result in (2.30), the following expression for the diabatic heating rate results:

$$J/c_p \approx -\alpha_r(z, T_r)\delta T, \tag{2.32}$$

where

$$\alpha_r(z, T_r) = -\frac{\tau}{c_p \rho_0(z)} \left\{ \sum_{\Delta\nu} \left(\frac{dB_{\Delta\nu}}{dT}\right)_{T=T_r} \left(\frac{d\bar{T}_f}{dz}(z, \infty)\right) \right\}. \tag{2.33}$$

α_r is referred to as the Newtonian cooling coefficient. It gives the rate at which large-scale temperature disturbances, like those associated with forced planetary Rossby waves, are damped by radiation. Typical values in the upper stratosphere are on the order of 0.15/day. Thus, the radiative damping time, α_r^{-1}, is about a week. From the upper to lower stratosphere the rate decreases, and typical damping times can be on the order of a month or longer.

Horizontal variations of the damping time enter through the temperature dependence of the Planck function and transmission. First, consider only the former contribution. In the middle stratosphere radiative cooling by CO_2 dominates the net cooling rate. The damping time, based only on longwave emissions in the absorption band of CO_2 (around 670 cm^{-1}), is found to depend on temperature as (Kiehl and Solomon, 1986; hereafter KS)

$$\alpha_r^{-1} \propto T_r^2 e^{h\nu/kT_r} = T_r^2 e^{960/T_r}. \tag{2.34}$$

The resulting latitudinal profile of α_r^{-1} in the winter middle stratosphere shows a maximum at the pole, where T_r is a minimum, and a decrease out through the midlatitudes.

Thus, based on the Planck function contribution alone, temperature perturbations will be damped more quickly in the midlatitudes than in the polar regions. Horizontal variations of α_r^{-1} also enter through the temperature dependence of the absorption coefficient, and hence the transmission. The addition of this effect does not appear to change the *qualitative* results just discussed (KS; fig. 5). In the lower stratosphere the addition of longwave heating by ozone alters the latitudinal structure of the damping time, especially in the tropics, but there still exists a decrease from the pole to the midlatitudes (KS; fig. 5).

Using (2.32) and assuming that vertical variations in density dominate the vertical derivative, (2.26) can be solved for the departure from the radiative equilibrium temperature, $\delta\overline{T}$:

$$\delta\overline{T} = \alpha_r^{-1}\left(\frac{N^2H^2}{\rho_0 f_0 R}\right)\frac{\partial}{\partial y}\nabla\cdot\mathbf{F}. \tag{2.35}$$

This formula, while it disallows significant vertical derivatives of $\delta\overline{T}$ and large deviations from radiative equilibrium, may be a good "zeroth-order" diagnostic of actual seasonal departures from T_r. A more accurate means of approximating $\delta\overline{T}$ would be to vertically integrate (2.26), using the fact that density goes to zero with height. We then obtain

$$\delta\overline{T}(y,z) = \frac{N^2H}{\alpha_r(y,z)\rho_0(z)f_0R}\frac{\partial}{\partial y}\int_z^\infty(\nabla\cdot\mathbf{F})dz'. \tag{2.36}$$

Thus, not only does the wave forcing at the level of interest contribute to observed departures from radiative equilibrium, but also the wave forcing above that level.

2.3.2 Wave Breaking Dynamics

Recall the large region of E-P flux convergence in Fig. 2.1. Physical interpretation of this is aided by noticing that

$$\nabla\cdot\mathbf{F} = \rho_0\overline{v'q'}, \tag{2.37}$$

where q' is the eddy QG PV. Thus, a region of convergent E-P flux is associated with a meridional down-gradient flux of PV. This expression, however, is only valid for the quasi-linear problem of waves interacting with a zonal-mean flow. One means of inferring the transport of PV in the fully nonlinear problem is through consideration of the wave-breaking process. Wave breaking, as defined by McIntyre (1985), is the irreversible

deformation of material contours (e.g., PV contours) that otherwise would simply undulate reversibly. It is tempting to attach a particular phenomenon to the *concept* of wave breaking; the above definition, however, is completely general. While many phenomena may be responsible for the observed wave breaking, McIntyre suggests that they are all just subclasses of something more all-encompassing. To really understand why and how the mixing takes place, we must consider carefully the details of the subclass of interest.

Analytical details of wave breaking are not, in general, readily obtained, as the processes involved are quite complex. Models describing the most simplistic subclasses of breaking do, however, lend some insight into the qualitative aspects of atmospheric phenomena. The process of mixing at the vortex edge is analogous to that associated with a Rossby-wave nonlinear critical layer (Stewartson, 1978; Warn and Warn, 1978; Killworth and McIntyre, 1985). If one forces, for example, a horizontal shear flow with a sinusoidal oscillation, the perturbation vorticity will evolve at the critical line, $\bar{u} - c = 0$, in a manner quite unlike that of the adjacent fluid. In 1978 Stewartson and Warn & Warn found analytical solutions to the inviscid nonlinear Rossby-wave equation using the method of matched asymptotic expansions. An example of this solution (Haynes, 1989) is shown in Fig. 2.2b. The right sketch depicts the initial and final distributions of PV according to the solution of the Rossby-wave equation. It is clear that a down-gradient flux of PV has taken place.

In the polar stratosphere we have in mind the stripping of high PV air from the polar vortex and subsequent mixing into the middle latitudes. This happens as vertically propagating Rossby waves, with their associated PV anomalies, advect PV from the vortex edge. Figure 2.2a illustrates this process. The left sketch shows an idealized plot of the PV before and after wave breaking. Initially, the PV has a uniform meridional gradient. The PV stripped from the vortex by the breaking waves at high latitudes is mixed into the lower latitudes, creating a uniform, or well-mixed, region. Sharper gradients of PV at the vortex edge result, creating a stronger restoring force. The "inelasticity" of the vortex inhibits the formation of large-amplitude Rossby waves. Thus, large-scale mixing via the irreversible deformation of PV at the vortex edge is reduced.

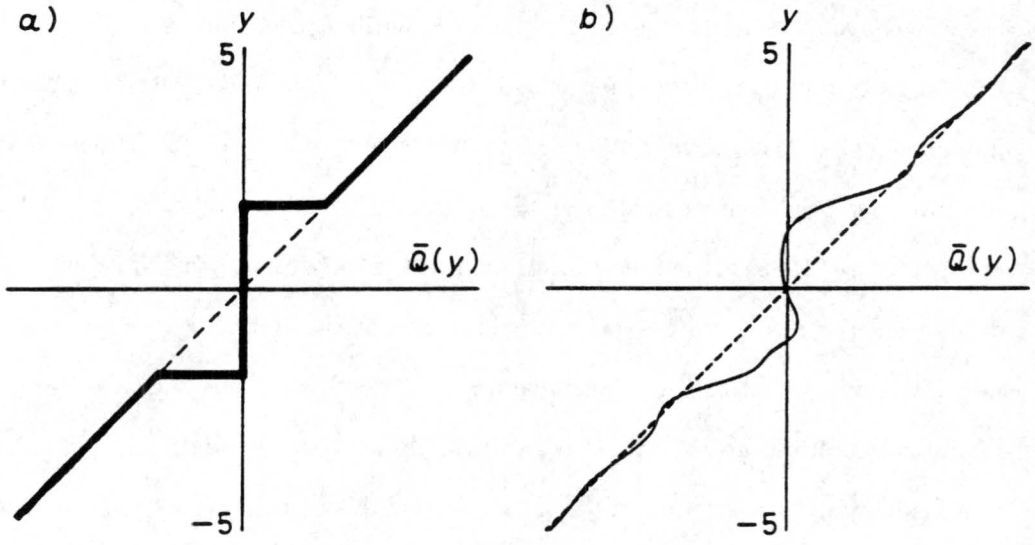


Figure 2.2: The left sketch depicts the idealized rearrangement of PV due to wave breaking. The right sketch shows an analytical solution to the Rossby-wave nonlinear critical-layer problem. The dashed(solid) line represents the initial(final) distribution of PV (From McIntyre (1992)).

2.4 The Stratospheric Wave Pump

The warming within the circumpolar vortex described in the previous sections is indicative of a large-scale subsidence in the polar regions. For timescales much greater than α_r^{-1} we can regard the temperature field to be in a steady-state. The thermodynamic equation, (2.25), reduces to

$$\bar{w}^* = \frac{R\bar{J}}{c_p N^2 H}. \tag{2.38}$$

Thus, at high latitudes, where $\bar{J} < 0$ (i.e., $T > T_r$), the residual vertical velocity is downward. Using observed heating rates in a radiative transfer model, Rosenlof (1995, hereafter R) made estimates of this vertical velocity in the stratosphere. For January of 1993 R found values of the 70-mb vertical velocity in the Northern Hemisphere on the order of mm s^{-1}, with a maximum of about -0.65 mm s^{-1} extending from the pole to around 70°N. The Southern Hemisphere values in July were less, consistent with a weaker wave forcing there. The maximum downward velocity of -0.5 mm s^{-1} peaked around 50°S. The position and narrowness of the peak are most likely due to the greater radius of maximum winds (i.e., focus of convergence) and lesser latitudinal extent of the residual circulation. Near the pole the velocity reached a local minimum of about -0.2 mm s^{-1}. While this velocity is small, it may still cause noticeable departures from radiative equilibrium over seasonal timescales, as illustrated in Section 5.2.

Since $\bar{w}^*$ is identified with the zonal-mean diabatic vertical velocity in this time mean, it is reasonable to infer that chemical tracers in the polar stratosphere will be transported from higher to lower levels. This has interesting implications for the column ozone at high latitudes; when $\bar{w}^*$ is particularly strong, ozone-rich air from the mid-stratosphere will be transported down into the lower stratosphere, increasing the column amount. In Section 2.5 we show how this phenomenon can be used as an indicator of enhanced wave activity in the stratosphere.

The residual vertical velocity in the polar stratosphere is only one part of the "big picture" in the stratosphere. Consider the steady-state version of the momentum equation, (2.14),

$$\bar{v}^* = -f_0^{-1}\rho_0^{-1}\nabla \cdot \mathbf{F}. \tag{2.39}$$

As found in Section 2.3.1, regions where the waves act to decelerate the flow and produce a subsequent warming are marked by large values of E-P flux convergence. From (2.39) we then expect to find a poleward meridional flow. By continuity the established circulation can be closed with an equatorward flow at lower levels and an upward vertical velocity in the tropics.

The control that the wave-induced force places on the mean meridional circulation in the stratosphere can be demonstrated explicitly by combining the steady-state versions of (2.14) and (2.16):

$$-f_0^{-1}\frac{\partial}{\partial y}\nabla\cdot\mathbf{F}+\frac{\partial}{\partial z}\left(\rho_0\bar{w}^*\right)=0. \tag{2.40}$$

In the winter hemisphere the largest wave-induced forcing occurs in the upper stratosphere at midlatitudes. Away from this region the forcing term in (2.40) drops out, and we find that the streamlines are lines of constant latitude. The resulting streamline pattern is aligned primarily in the vertical with minimal extent North and South of the forcing.

Equation (2.40) can be integrated vertically given the appropriate boundary conditions. Assuming $\rho_0\bar{w}^*$ is zero at the surface, integration of (2.40) from the surface to some level, z, results in an expression for $\bar{w}^*$ at z. This, however, is not the most advantageous approach. It requires an explicit knowledge of boundary-layer frictional forces, which may be sensitive to the wave-induced residual circulation extending down into the boundary layer (Haynes et al., 1991; hereafter HMMSS). If we instead assume that $\rho_0\bar{w}^*$ approaches zero as we go to infinity, we can integrate (2.40) from some level, z, to the top of the atmosphere and get $\bar{w}^*$ at z (HMMSS):

$$\bar{w}^*(y,z)=-\rho_0^{-1}f_0^{-1}\frac{\partial}{\partial y}\int_z^\infty(\nabla\cdot\mathbf{F})dz'. \tag{2.41}$$

Because the vertical velocity is determined by the forcing above, the waves are said to exert a downward control. In the thought-experiment introduced by HMMSS the wave forcing is prescribed within certain levels of the middle atmosphere. Not only must the effects of planetary wave activity be parameterized, but also the effects of the more complicated phenomenon of mesospheric gravity wave drag (Garcia, 1991). The contrast between the

complications of upward and downward control, then, do not seem as stark given the parameterization problems of both. Downward control, however, does have the advantage of separating the conceptual problems associated with boundary-layer frictional forces and stratospheric wave-induced forces, focusing on the latter (Haynes, et al., 1996; Egger, 1996). Comparing the residual circulation estimated by heating rates and by downward control, Rosenlof and Holton (1993) found reasonable agreement during the solstice seasons. They concluded that improved parameterization of gravity wave drag would need to be included for better estimates of the residual circulation.

In the preceding discussion it was found that the residual circulation is determined solely by the prescribed wave forcing. No mention of the shortwave forcing was made. To elucidate the physics of the steady-state limit, we return to (2.14)–(2.17). For simplicity we use the f-plane approximation and, following Holton et al. (1995), we parameterize only the longwave cooling by the Newtonian cooling approximation. Assuming the wave forcing, shortwave heating, and residual vertical velocity to be oscillatory functions of time with a characteristic timescale of σ^{-1}, an equation for the spatial portion of $\bar{w}^*$ is obtained:

$$\frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0\hat{w})\right)+\left(\frac{i\sigma}{i\sigma+\alpha}\right)\frac{N^2}{{f_0}^2}\frac{\partial^2\hat{w}}{\partial y^2}=\left(\frac{i\sigma}{i\sigma+\alpha}\right)\frac{R}{H{f_0}^2}\frac{\partial^2\hat{Q}}{\partial y^2}+\frac{1}{f_0}\frac{\partial^2\hat{G}}{\partial y\partial z}, \tag{2.42}$$

where α^{-1} is the radiative cooling timescale, $\hat{w}$ the residual vertical velocity, $\hat{Q}$ the shortwave heating rate, and $\hat{G}$ the wave forcing.

For timescales long compared to the radiative cooling timescale (i.e., $\sigma^{-1} \gg \alpha^{-1}$)

$$\frac{i\sigma}{i\sigma+\alpha} \longrightarrow 0. \tag{2.43}$$

Thus, the meridional derivative term on the left-hand side and the shortwave forcing term on the right-hand side will disappear in this limit. Even if the shortwave *heating* is nonzero, it will have no effect on the secondary circulation. As Holton et al. point out, the temperature-dependent longwave cooling will respond so as to balance the shortwave heating. There will, however, be a "dynamical" heating in the vicinity of the forcing leading to deviations from radiative equilibrium and a residual circulation, as discussed above.

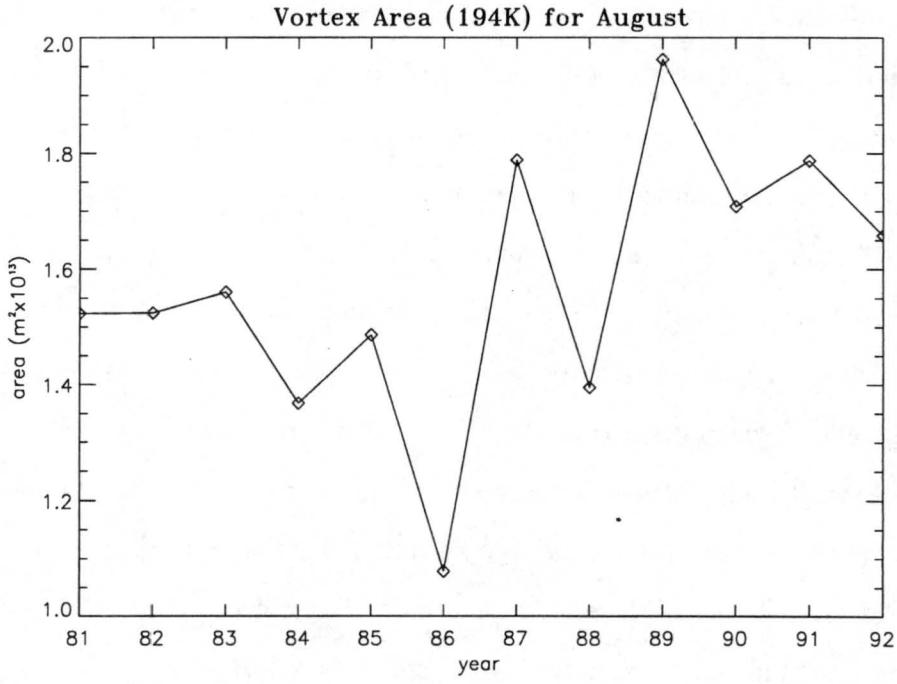


Figure 2.3: Interannual variability of the area bounded by the 194-K isotherm in the Southern Hemisphere. Data courtesy Ian Wittmeyer.

2.5 Circumpolar Vortex Interannual Variability

Thus far the discussion of the vortex has focused on how vertically propagating waves from the troposphere affect the mean state in the stratosphere, specifically their tendency to decelerate the jet-level winds. This is a key idea in understanding why the winter-spring temperatures within the vortex are warmer than those expected from a purely radiative model. If these waves had the same effect on the vortex year after year, causing it to evolve in the same manner, then we could simply focus on a single year's evolution. Nature, however, is not that predictable; there exists a marked interannual variability in the strength of the stratospheric wave forcing, hence the evolution of the circumpolar vortex. A large portion of the variance in polar vortex intensity may possibly be explained by a QBO modulation of extratropical fields. In this section we summarize Tung and Yang's (1994ab, hereafter TYab) discussion of a mechanism which could create such a link. The details of the interannual variability in the seasonal evolution of the austral polar vortex are deferred to Chapter 5.

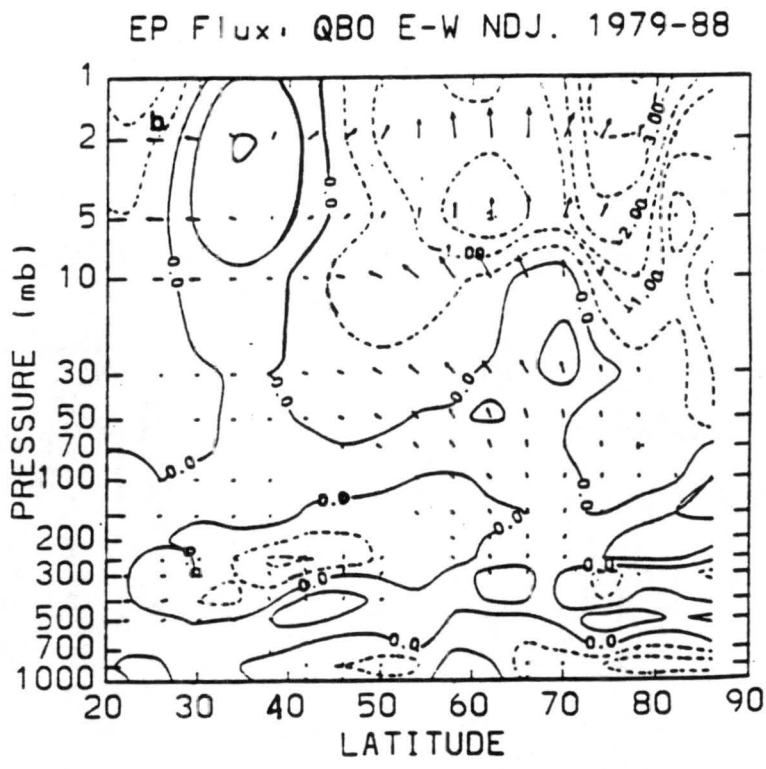


Figure 2.4: Composite difference in NDJ Eliassen-Palm cross-sections in the Northern Hemisphere. Divergence contour interval 0.5 $ms^{-1}day^{-1}$ (From Dunkerton and Baldwin (1991)).

In Fig. 2.3 we see that the vortex area in the Southern Hemisphere varies dramatically from year to year. Holton and Tan (1980) investigated a similar biennial oscillation in the Northern Hemisphere using a 16-year record of analyzed 50-mb geopotential heights. Looking for signals of the QBO in the extratropics in winter, they found zonal-mean geopotential heights to be much higher during the easterly phase of the QBO than during the westerly phase. In addition they also observed a QBO modulation of low-wavenumber planetary wave amplitudes. They concluded that there does exist a significant QBO signal in the extratropical circulation, and hypothesized that it might be related to a modulation of the $\bar{u} = 0$ critical line for stationary waves. While the exact role of the critical line is not completely understood, the modulation of planetary wave activity by the QBO is essential in explaining how a quasi-biennial oscillation could exist at higher latitudes. Dunkerton and Baldwin (1991) investigated this modulation by creating Northern Hemisphere E-P composites for years when the QBO was in its easterly and westerly phases. In Fig. 2.4 we find the difference between the easterly and westerly composites. Throughout the

bulk of the atmosphere the values are close to zero, so there is little difference between the westerly and easterly phases of the QBO. At high latitudes above 20 mb, however, a clear region of convergent E-P flux is visible on the order of -2 $ms^{-1}day^{-1}$, indicating an enhancement of wave forcing in the high-latitude stratosphere during the easterly phase of the QBO.

Other studies, using chemical tracers, have also shown a clear QBO signal in the extratropics. In the discussion of the diabatic circulation in Section 2.4 we found that the wave forcing of the mean flow will effect a downward transport in the polar stratosphere. The stronger the wave forcing, the greater the downward transport. In the case of ozone, which has its source in the mid-stratosphere, this transport is realized as a column ozone anomaly. When the wave forcing is weak(strong), we expect a negative(positive) anomaly. TYa compared the TOMS global and equatorial ozone anomalies with the 30-mb Singapore wind in an effort to establish such a correlation. In Fig. 2.5 we see that the global ozone anomaly and QBO are 180 degrees out of phase. The equatorial ozone anomaly, however, is in phase with the QBO (see Fig. 2.6). Since the global ozone anomaly is made up of both equatorial *and* extratropical anomalies, there must exist a large extratropical anomaly 180 degrees out of phase with the QBO in order to produce the observed global correlation shown in Fig. 2.5. This would then imply that during the easterly(westerly) phase of the QBO, a positive(negative) ozone anomaly exists at high latitudes, consistent with enhanced(reduced) downward transport. Based on these results, we can draw two conclusions pertaining to the dynamics of the circumpolar vortex: First, during the westerly phase of the QBO, the vortex is more isolated and closer to a radiative balance. Second, during the easterly phase the vortex is more perturbed, causing temperatures to depart significantly from T_r (O'Sullivan and Young, 1992; O'Sullivan and Salby, 1990; Holton and Tan, 1980). Although these conclusions are substantiated to some degree by the observations (especially in the Northern Hemisphere), it is important to have a sound dynamical theory to explain how the QBO manages to modulate the chemical species amounts, stratospheric circulation, and circumpolar vortex strength. The reader may wish to refer to Fig. 2.7, which schematically describes the major processes discussed in TYb, and summarized below.

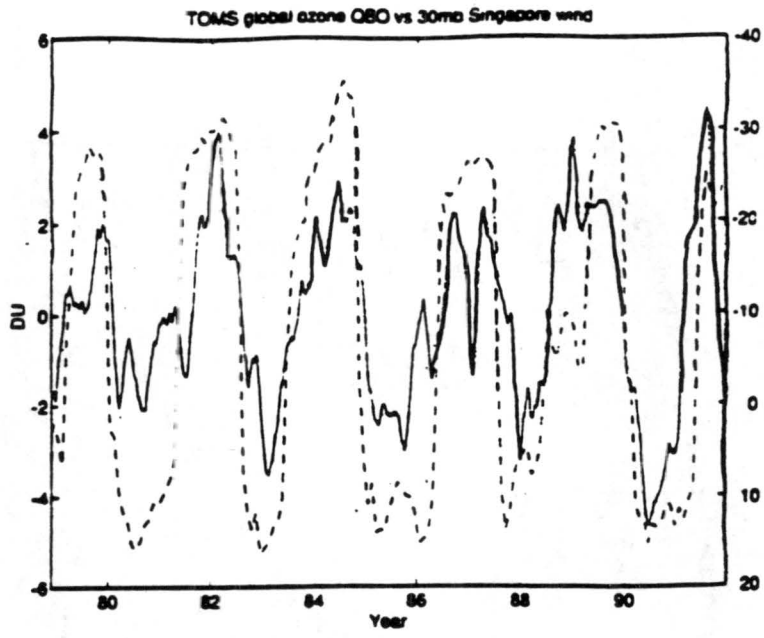


Figure 2.5: TOMS global ozone anomaly (solid line, Dobson units) and 30-mb Singapore wind (dashed line, ms^{-1}) (From Tung and Yang (1994a)).

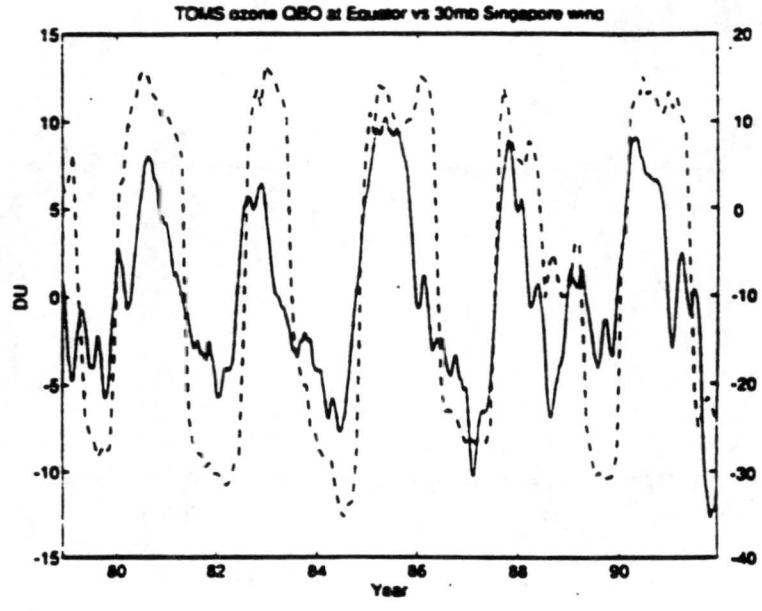


Figure 2.6: TOMS equatorial ozone anomaly (solid line, Dobson units) and 30-mb Singapore wind (dashed line, ms^{-1}) (From Tung and Yang (1994a)).

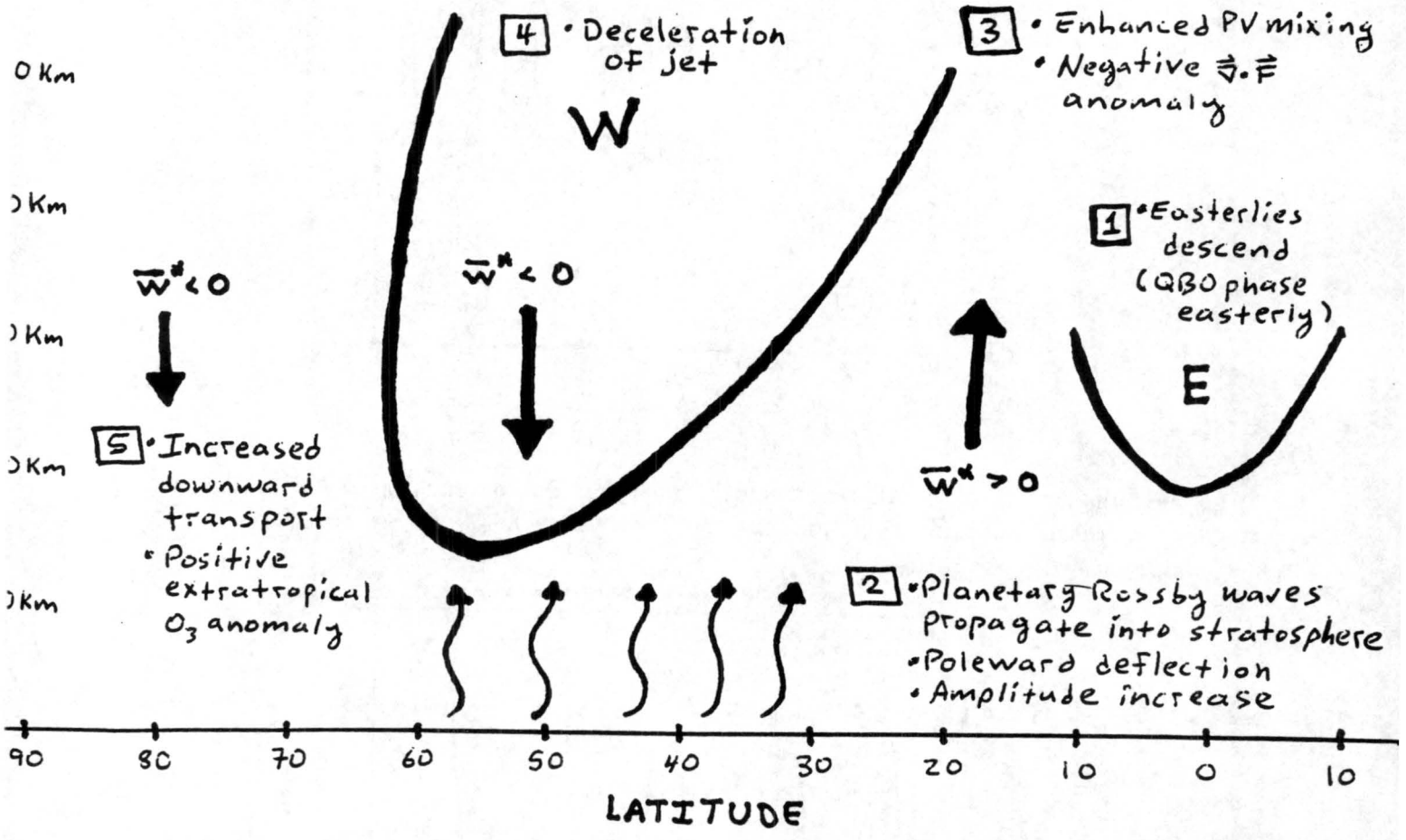


Figure 2.7: Schematic diagram of the processes involved in the QBO modulation of extratropical fields.

2.5.1 The Quasi-biennial Oscillation

The quasi-biennial oscillation (QBO) is a roughly periodic reversal in stratospheric winds, from westerly to easterly, confined to about 12 degrees either side of the equator (Reed et al., 1961). The period is not regular, and averages around 27 months. Two kinds of equatorially-confined waves with vertical energy propagation are known to initiate the downward progression of the easterlies and westerlies. Vertically propagating Kelvin waves transport westerly momentum from the lower levels up to the level where $\bar{u} - c = 0$ (c is the positive phase speed of the Kelvin wave). Here, the Kelvin waves deposit their westerly momentum, accelerating the mean flow, causing the $\bar{u} - c = 0$ level to descend. The easterly phase of the QBO comes about as Yanai (Rossby-gravity) waves propagate vertically. Although these waves also transport westerly momentum upward, they have a strong vertical component of the E-P flux associated with a poleward transport of heat (Holton, 1992). At the level where $\bar{u} - c = 0$ (c is the negative phase speed of the Yanai wave), the vertical derivative of the heat flux is negative. By the TEM momentum equation, (2.14), this implies a deceleration of the mean flow, and subsequent descent of the easterlies. It should be emphasized that, because the Kelvin and Yanai waves are confined to the Tropics, they will not contribute to the wave forcing in the extratropical stratosphere.

2.5.2 Waveguides and Poleward Focusing

Recall from Section 2.3 that planetary Rossby waves *will not propagate* through a region of easterlies. Solving Eq. (2.11) for the vertical wavenumber, we find

$$m^2 = \frac{N^2}{f_0^2}\left[\frac{\bar{q}_y}{\bar{u}_g - c} - k^2 - l^2 - \frac{{f_0}^2}{4H^2N^2}\right]. \tag{2.44}$$

The quantity, m, is often referred to as the refractive index, in analogy to geometrical optics. Where the refractive index is real, wave energy propagates freely. Where it is imaginary, wave energy propagation is inhibited, and the waves may actually be reflected.

In the stratosphere $\bar{q}_y$ is generally positive, so regions of imaginary refractive index are most likely to be found where the mean flow is easterly. Thus, during the easterly

phase of the QBO we expect to find an enhancement of the area spanned by imaginary m. The region of stratospheric westerlies where propagation is favored, referred to as the polar westerly waveguide, is subsequently narrowed.

According to the climatological cross-section of E-P flux (see Fig. 2.4), waves forced in the midlatitudes tend to deflect equatorward with height. The narrowing of the waveguide, however, will cause waves to be deflected slightly more poleward. That the waves are redirected to higher latitudes does not guarantee an increased disturbance of the circumpolar vortex; we also require that the waves be of sufficient amplitude. The equation for wave-activity density, (2.24), provides the conditions under which significant increases in wave amplitude can be expected. In the absence of wave-energy convergence the wave-activity density is materially conserved following the group velocity. Thus, as the waves propagate upward into a lower-density environment, $\overline{q'^2}$, hence wave amplitude, must increase. For wave-activity density to grow substantially following the group velocity, there must also exist a pronounced convergence of wave energy. This is especially important in the Southern Hemisphere where waves must propagate up to much higher altitudes (than in the Northern Hemisphere) to amplify to the point of breaking. By narrowing the waveguide, we restrict the space through which the waves can propagate, enhancing the wave-energy convergence. The amplitudes of the waves approaching the edge of the polar vortex during the easterly phase of the QBO are more prone, then, to becoming large and causing noticeable disruptions of the vortex at lower levels over seasonal timescales. TYb also note that enhanced planetary wave breaking may occur as the criteria for instability defined by Garcia (1991) is satisfied in the mixing region.

2.5.3 Forcing of the Mean Flow

The irreversible mixing of material contours at the vortex edge leads to a down-gradient flux of PV, and by Eq. (2.37) a negative E-P flux divergence anomaly. With more wave activity, and subsequent breaking, during the easterly phase of the QBO, we expect to find this anomaly to be even greater than during the westerly phase. Caution is advised, however, in blindly making this connection in the Southern Hemisphere. Unlike in the Northern Hemisphere, forced planetary Rossby waves are not found in abundance, and

their amplitudes are not as large. Therefore, during the easterly QBO phase, the increase in wave activity may not always be statistically significant. Nonetheless, an increasing number of studies claim to observe a QBO signal in the dynamical fields of the austral polar stratosphere.

The direct effect of enhanced wave breaking on the vortex circulation is the deceleration of the jet winds. For years in which the QBO is in its easterly phase we expect a greater forcing at the vortex edge, and a subsequently less-developed vortex. The ozone depletion chemistry acts over smaller areas, resulting in less spring-time depletion of ozone. In years when QBO is in its westerly phase, waves are more inclined to follow their climatological equatorward path, leaving the vortex relatively undisturbed. Its evolution under these conditions is observed to be closer to that expected of a radiatively-driven vortex. Typically, the jet winds in the lower stratosphere will maintain their late-winter maxima into early spring until the Sun warms the polar regions.

The negative tendency in the vortex winds is partially offset by the Coriolis force, which establishes a counteracting mean meridional circulation. By the downward control principle discussed in Section 2.4, the circulation will be downward at high latitudes. As we showed in the introduction to this section, positive anomalies in column ozone provide evidence of an enhanced circulation during the easterly phase of the QBO.

2.6 Summary

We have demonstrated how planetary Rossby waves forced in the troposphere can propagate into the stratosphere and influence the the mean flow there. If the Eliassen-Palm flux associated with these waves does not converge (i.e., the waves are steady, conservative, and linear), then the poleward transport of heat by the waves will exactly cancel the equatorward transport due to the wave-induced Eulerian-mean circulation. The temperatures will approach a radiative equilibrium in this case. In the actual stratosphere this exact cancellation does not occur due to the dissipation of waves at the circumpolar vortex. The result is a deceleration of the polar night jet and warming in the polar regions. The zonally-averaged residual circulation responsible for the departures from radiative

equilibrium conditions was shown in the steady-state to be controlled solely by the wave forcing. Furthermore, it was demonstrated that the temperature departures and residual circulation at any given level are controlled by the wave forcing above, the so-called downward control principle.

The aforementioned residual circulation is associated with the transport of tracers in the steady-state limit (on the order of a few months in the stratosphere). This circulation is believed to be modulated by the tropical QBO. We showed how the polar westerly waveguide in the stratosphere is narrowed during the easterly phase of the QBO, forcing waves to propagate to higher latitudes (i.e., closer to the circumpolar vortex) and increase their amplitudes through enhanced convergence of wave energy. During the westerly phase of the QBO, the waves tend to propagate more towards the equator, leaving the vortex relatively undisturbed. In Chapter 5 further observational evidence supporting this theory will be presented.

Chapter 3

MSU DATA

3.1 Introduction

An observational study of the polar stratospheric circulation requires a data set with global coverage and, because of the important role that interannual variabilities play (Mechoso et al., 1985), a long temporal record. Analyzed data sets from the NMC and ECMWF satisfy the above criteria, but are a mixture of data from many sources. While this assimilation of observations is not necessarily disadvantageous, one might find the assessment of data quality to be more manageable if only one source is involved. We have chosen to use primarily the brightness temperature data from the Microwave Sounding Unit (MSU), flown aboard the NOAA TIROS–N series of polar orbiting satellites. The MSU data is, in fact, used by the aforementioned data centers in their analyses. Independent studies of the MSU data, then, not only further our understanding of the MSU itself, but also the data sets which incorporate it.

The MSU samples the Earth's atmosphere globally, eliminating the need for analysis over regions which are traditionally sparsely sampled, like oceans and the polar latitudes. Its temporal record is continuous, with eighteen years of data from the time of its launch in 1978. Among satellite-based instruments this continuity makes the MSU a unique resource. Vroman and Stephens (1989; hereafter VS) explored the capacity of the MSU to depict various general circulation features. By comparing against ECMWF data, VS established the consistency of the MSU fields and were also able to detect a sudden stratospheric warming using the lower stratospheric channels. It was concluded that the MSU could be a useful tool in other general circulation applications.

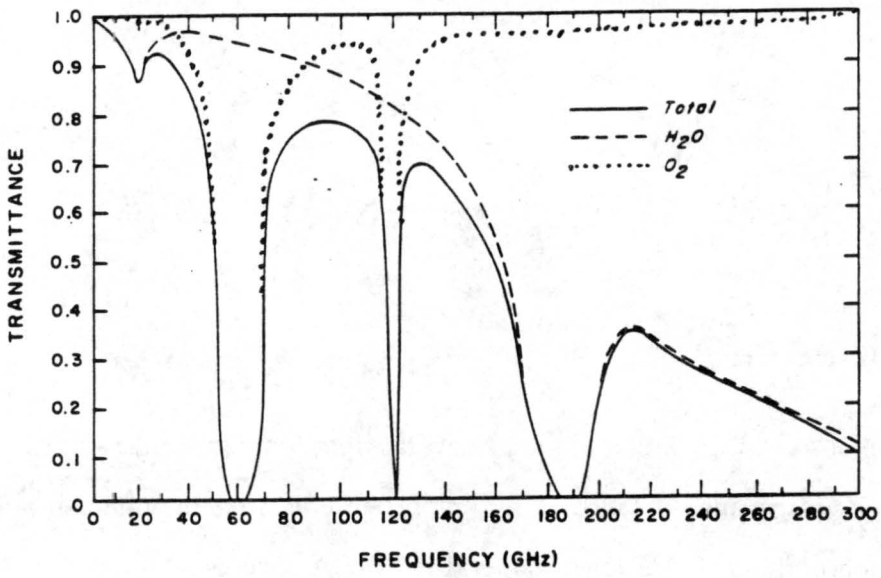


Figure 3.1: Oxygen and water vapor absorption bands (From Liou (1980)).

In this chapter we introduce the MSU data and discuss its source from the perspective of radiative transfer. After a brief discussion of data quality the MSU brightness temperature data is related to ECMWF temperatures and thicknesses for the purpose of determining roughly what layers the MSU is sampling. Finally, a detailed look at the high-latitude data is presented, pointing out certain flaws of the MSU as well as its consistency with data reported by other studies. In light of these findings we discuss the feasibility of a detailed study of the polar stratosphere using MSU data alone.

3.2 Data Source

The MSU is a remote temperature sounder with the ability to penetrate clouds. It was first launched in 1978, and has since flown aboard the NOAA TIROS–N series of polar orbiting satellites (VS). The four channels available sample the entire troposphere and lower stratosphere. In this study we use the brightness temperatures obtained from the upper two tropospheric channels and the lower stratospheric channel.

Figure 3.1 shows the strong oxygen absorption band centered about 60 GHz which the MSU uses to construct its temperature sounding. Emissions from water vapor are

small in this band, so the spectral lines are primarily due to oxygen rotational transitions associated with the relative orientations of its electronic spins and molecular rotation (Liou, 1980). The spacing between the rotational energy states, $h\nu$, is much smaller than the thermal energy, kT, so many rotational quantum states are excited, reminiscent of classical behavior. In this continuous limit the observed emissions depend on the absolute temperature of the medium in an approximately linear fashion, as can be shown from classical physics (e.g., Feynman, 1963). This fact can also be arrived at by taking the quantum mechanically correct Planck function (spectral radiance, $\mathrm{Wm^{-2}Hz^{-1}}$),

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1}, \tag{3.1}$$

in the limit $h\nu << kT$,

$$B_\nu(T) \approx \frac{2k\nu^2}{c^2} T, \tag{3.2}$$

where h is Planck's constant, ν is the emission frequency, k is Boltzmann's constant, c is the speed of light, and T is the absolute temperature.

The observed emissions at the top of the atmosphere come from the atmosphere itself and the surface, as can be shown by solving the radiative transfer equation for a nonscattering atmosphere. As a beam of radiation of intensity, I_ν, passes through a grey atmosphere in thermodynamic equilibrium, various absorption and emission processes, due to the presence of matter in the atmosphere (e.g., CO_2, O_2, O_3, H_2O, etc.), act to change the intensity (Liou, 1980):

$$\frac{dI_\nu}{ds} = -\sigma_{abs}[I_\nu - B_\nu], \tag{3.3}$$

where ds is a path length (just dz for vertical paths) and σ_{abs} is the absorption coefficient ($\mathrm{m^{-1}}$). The first term on the right-hand side represents extinction, while the second term represents an increase in intensity due to emission. What we would like to find is an expression for upwelling radiance (intensity) at the top of the atmosphere, which is what the MSU actually measures.

First, upon defining an optical depth increment,

$$d\tau_\nu(s) = -\sigma_{abs,\nu}(s)ds, \tag{3.4}$$

the radiative transfer equation can be rewritten as

$$\frac{d}{d\tau_\nu}\left[I_\nu e^{-\tau_\nu}\right] = -B_\nu e^{-\tau_\nu}. \tag{3.5}$$

Integrating (3.5) vertically through the atmospheric column, and defining a transmissivity,

$$\mathcal{T}_\nu = e^{-\tau_\nu}, \tag{3.6}$$

the following equation for upwelling radiance at the top of the atmosphere results:

$$I_\nu(z=\infty) = I_\nu(z=0)\mathcal{T}_\nu(0,\infty) + \int_0^\infty B_\nu[T(z)]\frac{\partial \mathcal{T}_\nu(z,\infty)}{\partial z}dz. \tag{3.7}$$

The hydrostatic equation allows (3.7) to be written as

$$I_\nu(0) = I_\nu(p_s)\mathcal{T}_\nu(p_s,0) + \int_{p_s}^0 B_\nu[T(p)]\frac{\partial \mathcal{T}_\nu(p,0)}{\partial p}dp. \tag{3.8}$$

The surface radiance contribution, $I_\nu(p_s)$, is comprised of two parts:

$$I_\nu(p_s) = \varepsilon_\nu B_\nu(T_s)\mathcal{T}_\nu(p_s,0) + (1-\varepsilon_\nu)\int_0^{p_s} B_\nu[T(p)]\frac{\partial \mathcal{T}_\nu(p_s,p)}{\partial p}dp, \tag{3.9}$$

where ε is the surface emissivity. The first term on the right-hand side is the contribution due to surface emissions. It is multiplied by ε, because the surface of the Earth is not necessarily a perfect absorber in the microwave region. The second term represents the downwelling atmospheric radiation reflected by the Earth's surface. It is of the same form as the upwelling atmospheric radiance in (3.8), except for the factor of $1-\varepsilon_\nu$, which is just the surface reflectivity.

Now define a brightness temperature, T_B, such that

$$I_\nu = \frac{2k\nu^2}{c^2}T_B. \tag{3.10}$$

Combining Eqs. (3.8) and (3.9), and using (3.2), gives the expression for brightness temperature (Liou, 1980):

$$\begin{aligned} T_B(\nu) &= \varepsilon_\nu T_s \mathcal{T}_\nu(p_s,0) + (1-\varepsilon_\nu)\mathcal{T}_\nu(p_s,0)\int_0^{p_s} T(p)\frac{\partial \mathcal{T}_\nu(p_s,p)}{\partial p}dp \\ &\quad + \int_{p_s}^0 T(p)\frac{\partial \mathcal{T}_\nu(p,0)}{\partial p}dp \end{aligned} \tag{3.11}$$

Because emissions of a particular frequency come from an atmospheric layer, the brightness temperature should not be viewed as representing a single level. It is more appropriate to think of it as a weighted mean-layer temperature, defined by the convolution of a layer temperature and a weighting function, $-d\mathcal{T}_\nu(p)/d(lnp)$. The weighting function determines how the temperature at each level contributes to $T_B(\nu)$. Ideally we would like the weighting function to take the form of a delta function (i.e., peaked at a single level) so that $T_B(\nu)$ can be equated with the actual temperature. For broader, more realistic weighting functions we may identify $T_B(\nu)$ with the temperature at the level of the peak of $-d\mathcal{T}_\nu(p)/d(lnp)$. It should be remembered, however, that emissions above and below this level will also contribute to $T_B(\nu)$, creating an apparent cold- or warm-bias when compared to the actual temperatures.

The daily fields of brightness temperature we use have a 2.5° grid spacing in both latitude and longitude. The precision of the data is quite good, as reported by Spencer and Christy (1993). Upon comparing Channel 4 aboard the NOAA–6 and NOAA–7 satellites during overlapping periods, they found a monthly gridpoint precision of 0.05°C in the tropics and 0.10°C at high latitudes. The signal to noise ratio was generally over 500 at all latitudes. Aside from the aforementioned difficulty in relating MSU brightness temperatures with actual temperatures, the data is quite useful.

During certain years, however, some data is missing, making analysis difficult. Based on a suggestion by Bailey et al. (1993), we reject any day's data, if more than 25 percent is missing. Figure 3.2 shows the percentage of days for which this occurs for each month and year of our study. The quantity of data available for each month is most variable prior to 1989. The primary reason for the high percentages scattered throughout this period is missing grids of data. For example, June through August of 1987 were each missing nearly 1/3 of their days. The days without data typically occur consecutively, suggesting a malfunctioning channel, although we were not able to confirm this. The general procedure was to throw out those days that did not meet the criteria and use the remaining days to form the monthly averages. Some months, as mentioned, were quite deficient in "good" days, so the criteria had to be relaxed to allow more days in

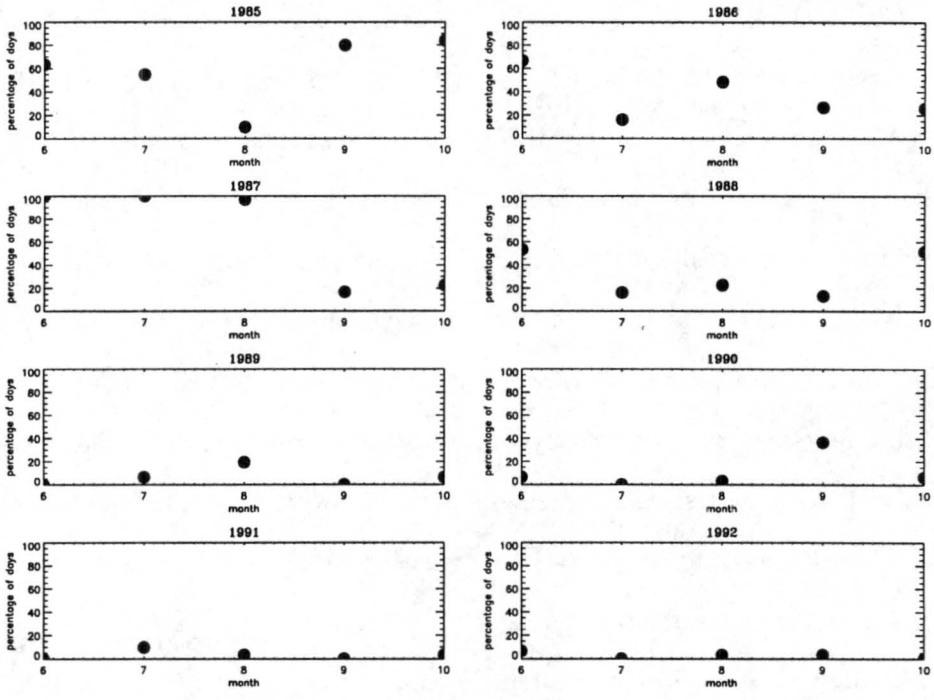


Figure 3.2: Percentage of days in each month with more than 25 percent of their data points missing.

the monthly averages. Therefore, the early years of our study may not have the quality of the latter years, although some attempt to remedy the situation was made using an interpolation scheme.

3.3 Comparison of MSU and ECMWF Data

The brightness temperatures measured by each channel can be shown to be proportional to geopotential thicknesses, although some trial and error is involved in exactly matching the two:

The hydrostatic equation,

$$\frac{dp}{dz} = -\rho g, \tag{3.12}$$

can be rewritten in term of geopotential by noticing that $d\Phi = gdz$. Using the ideal gas law, $p = \rho RT$, (3.12) becomes

$$d\Phi = -RTd\ln p. \tag{3.13}$$

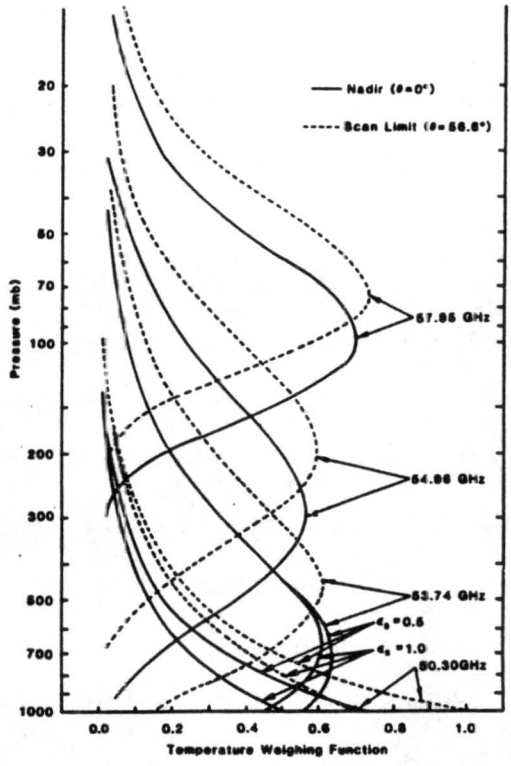


Figure 3.3: Weighting functions for Channels 2, 3, and 4 for nadir and scan-limit view angles (From Grody (1983)).

Integrating this vertically from p_0 to p_1, we obtain a relation between thickness and mean layer temperature, $\overline{T}$:

$$\delta\Phi = R \ln\left(\frac{p_0}{p_1}\right)\overline{T}, \tag{3.14}$$

where

$$\overline{T} = \frac{\int_{p_1}^{p_0} T d\ln p}{\int_{p_1}^{p_0} d\ln p}. \tag{3.15}$$

Although MSU T_B is not a true mean layer temperature in the sense that each level is weighted equally, it can be used, with care, in (3.14) for $\overline{T}$.

The three channels used in this study are Channel 2 (53.74 GHz), Channel 3 (54.96 GHz), and Channel 4 (57.95 GHz). The weighting function for each channel is shown in Fig. 3.3 for nadir (0° from the azimuth) and scan limit (56.6° from the azimuth) positions. Channel 2 peaks at around 700 mb, and has been shown to give a very good approximation to the ECMWF 850-300-mb geopotential thickness (VS). Notice, however, that its weighting function peak is very close to the surface and, where terrain is elevated, the peak may actually intersect the ground. The last two terms in Eq. (3.11) contain the weighting

function. Thus, the Channel 2 brightness temperature may be significantly biased by the surface temperature. Inhomogeneities in the surface emissivity will also affect the Channel 2 brightness temperature through the first two terms in Eq. (3.11). The atmospheric transmissivity of Channel 2 is about 0.1 (Grody, 1982), since it is displaced from the center of the oxygen absorption band. The first two terms of (3.11), which are both multiplied by the atmospheric transmissivity, will then be nonzero. Channels 3 and 4 are closer to the band center (see Fig. (3.1)), so this effect will be greatly diminished. Given these surface effects, care must be taken in interpreting the Channel 2 brightness temperatures, as they may not necessarily be representative of actual *atmospheric* temperatures.

The weighting function for Channel 3 peaks at around 300 mb in the upper troposphere. Although VS did not determine an explicit relationship between Channel 3 temperature and ECMWF thickness, it is plausible that, based on the weighting function, it could represent a 300-200-mb mean-layer temperature. For the purposes of this dynamical study, exact agreement in magnitude between the ECMWF and MSU is not crucial. Since velocity and vorticity are differentials of the temperature field, as shown in Chapter 4, it is only necessary that their spatial derivatives agree. As a consistency test we compared Channel 3 temperature in the Southern Hemisphere with the ECMWF 300-200-mb geopotential thickness for June through October of 1992 (not shown). We found reasonable agreement in form, except for slightly stronger meridional gradients in the MSU temperatures near 60°S. Thus, in the following chapters Channel 3 will be used as a 300-200-mb mean-layer temperature.

Channel 4's weighting function for nadir-viewing peaks at about 100 mb. VS found that its brightness temperatures are related to the ECMWF 200-100-mb geopotential thicknesses, and also closely resemble the ECMWF 100-mb temperatures. At the time of their study, however, analyzed fields of geopotential from the EC were not available above 100 mb. In this work we use the the ECMWF 50-mb geopotential to compute the 200-50-mb geopotential thickness. We were able to establish a better relationship with Channel 4 brightness temperature. The resemblance of the Channel 4 temperatures to the ECMWF 100-mb temperatures is still accurate. Other studies have used Channel 4

to investigate the thermal structure of the lower stratosphere. For example, Yulaeva et al. (1994) used it to demonstrate that on a monthly timescale near compensation exists between lower stratospheric temperature changes in the tropics and high latitudes. They interpret this as being due to the wave-driven circulation discussed in Section 2.4, with temperature rising above T_r in the extratropics and falling below T_r in the tropics.

3.4 High-Latitude Data

Previous studies using the MSU data have taken advantage of its global coverage to discuss aspects of the global-scale circulation. Data problems in particular regions have been remedied without having much effect on the overall results. This study, on the other hand, focuses on the polar regions of the Southern Hemisphere, making data quality there of the utmost importance. In this section we identify potential problems with the high latitude data that could influence our results and demonstrate the consistency of the data with other observations of the polar Southern Hemisphere.

Figure 3.4 shows the mean-layer temperatures in the Southern Hemisphere for all three channels. T_B for Channel 2 was found by VS to be best correlated with ECMWF 500-mb temperature. The temperature contours shown, however, deviate dramatically from those usually observed at 500 mb. They appear to roughly follow the Antarctic topography. Based on our previous discussion of Channel 2, this behavior is to be expected since the peak of the weighting function is close to the surface. Much of the Antarctic terrain, which is mostly snow- and ice-covered, in fact, intersects the 700-mb peak. The high albedo also contributes to the anomalously low brightness temperatures. The resulting large temperature gradients imply nonphysical values of the velocity and vorticity, making Channel 2 unusable in dynamical studies of the polar regions. For these reasons, we found it necessary to replace MSU Channel 2 data with ECMWF analyzed 300-mb geopotential data.

The thermal pattern of Channel 3 shows relatively weak meridional temperature gradients, consistent with temperatures near the tropopause. The high frequency oscillations in the data are most likely due to measurement error. Near the pole, say within 10 degrees, much data is missing, so only a coarse-grained picture of reality can be expected there.

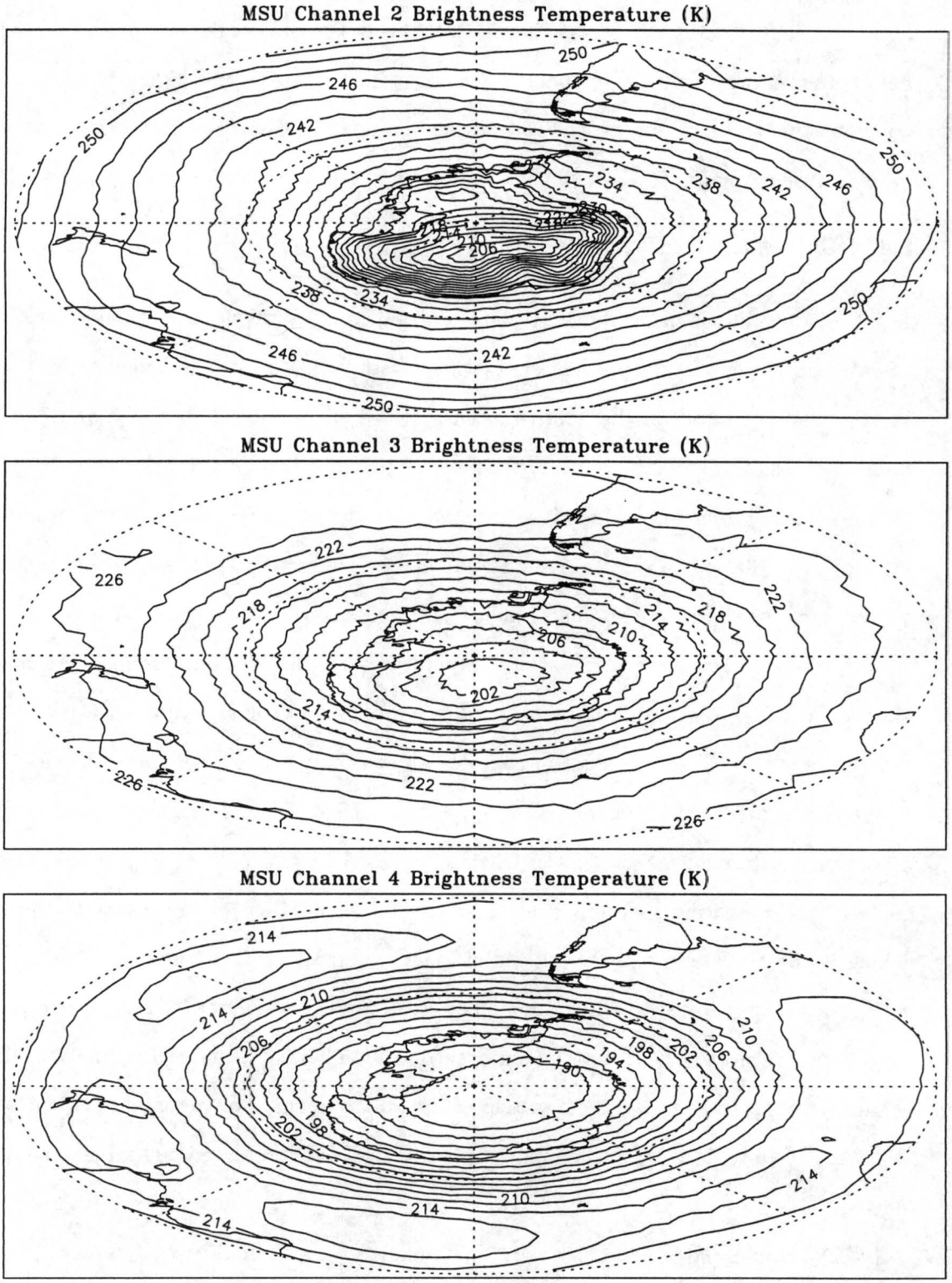


Figure 3.4: MSU brightness temperature for Channels 2 (top), 3, and 4 (bottom) for July 1985. Contour interval 2 K.

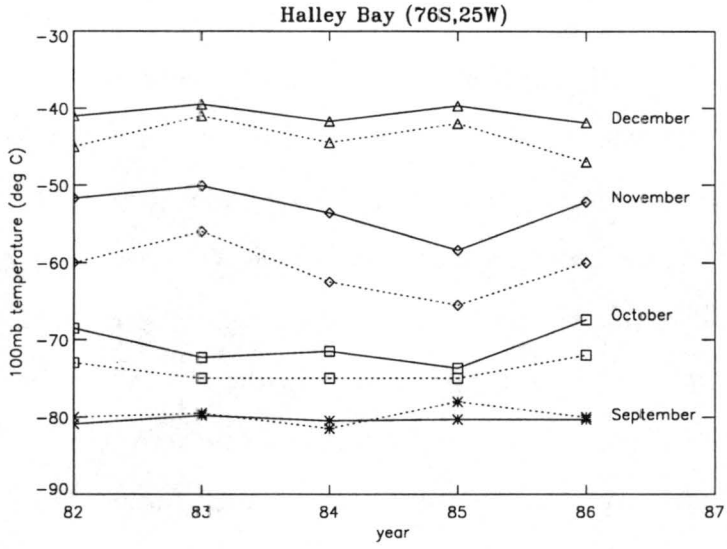


Figure 3.5: A comparison of MSU Channel 4 temperature (solid) and radiosonde data at 100 mb (dotted) over the Halley Bay station in Antarctica. Radiosonde data adapted from McIntyre (1989, Fig. 9).

Channel 4, which samples the lower stratosphere, shows a noticeable signature of the lower polar vortex. The strong temperature gradients around 60°S give some indication of the vortex edge. The center of the vortex is broad with uniform temperatures on the order of 190 K. These features are all consistent with previous observations (e.g., Schoeberl et al., 1992).

As a means of establishing consistency between the MSU data and other observations of the polar regions, we first compare the Channel 4 data with 100-mb radiosonde data above the Halley Bay station in Antarctica. Figure 3.5 shows the two data sets of monthly averaged temperatures. While both instruments agree in their representations of the temperature trends for each month, the magnitudes are not the same. The MSU temperatures are consistently warmer than the actual temperatures. Given that Channel 4 actually measures different levels of the atmosphere centered about the 100-mb weighting function peak, this discrepancy is not at all surprising. In September the polar temperatures directly above 100 mb are slightly colder. In October through December, however, the temperatures above 100 mb are warmer (a 10-K difference between 100 mb and 50 mb in November). Since the MSU Channel 4 weighting function extends up to 30 mb, it is likely that these warmer temperatures are producing the aforementioned

warm-bias. According to a time-height cross-section of temperature at McMurdo Sound (77.9°S, 166.7°E), vertical temperature gradients in September, October, and December are relatively weak above 100 mb (see Fig. 7 of Trenberth and Olson, 1989). Thus, one might expect the weighted-mean temperatures produced by the MSU to be close to the actual 100-mb temperatures. This is apparent in Fig. 3.5, where the maximum discrepancy during these months is observed to be on the order of 3-4 K. In November the vertical gradient in temperature is much greater, so the warm-bias will also be greater. The largest departures from the radiosonde temperature occur in November and are on the order of 6-9 K. Our study focuses on the late Winter and early Spring period (i.e., June through October), so, based on these results, we may assume the MSU brightness temperature to be a close approximation to the actual temperature.

As a further test of the consistency of the high latitude MSU data, we next compare Channel 4's depiction of interannual variability in the size of the austral polar vortex with observed variations in column ozone in the Southern Hemisphere polar region. The formation of Polar Stratospheric Clouds (PSC's) is a key step in the ozone depletion process. The number of PSC's is greatest when the vortex is cold and isolated (Schoeberl et al., 1992). The larger the region of cold temperatures, the greater the spring-time depletion of ozone. The interannual variability of the area bounded by the 194K contour, which is a measure of the interior vortex size, is depicted for the month of August in Fig. 3.6. The bottom Fig. 3.7 shows the TOMS October mean column ozone poleward of 70°S. From 1985 to 1986 the vortex area drops to the lowest value of our study. Correspondingly, we observe an increase in column ozone. The most dramatic rise in vortex area occurs between 1986 and 1987, increasing to one of its largest values. Consistent with the large cold area in August is a record low column ozone value in October. A similar increase in area occurs between 1988 and 1989, and we observe a large drop in column ozone here as well. Thus, we see that the year-to-year variations in vortex area as measured by the MSU are consistent with the observed anomalies in polar column ozone.

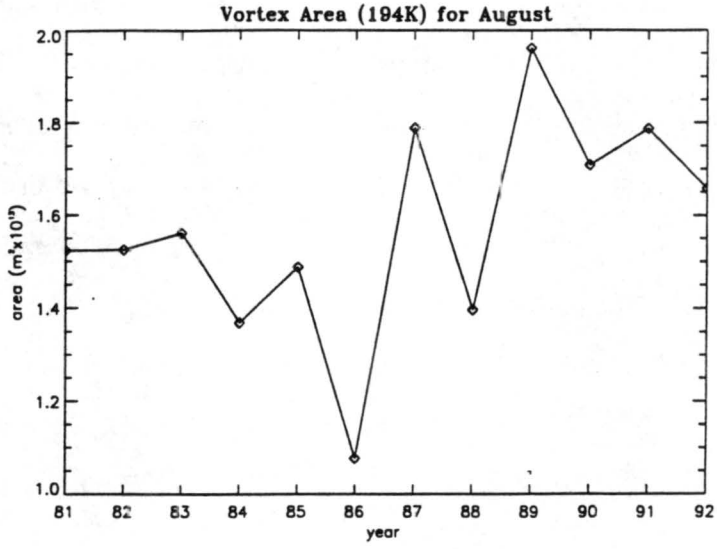


Figure 3.6: Interannual variability of the area bounded by the 194-K isotherm in the Southern Hemisphere. Data courtesy Ian Wittmeyer.

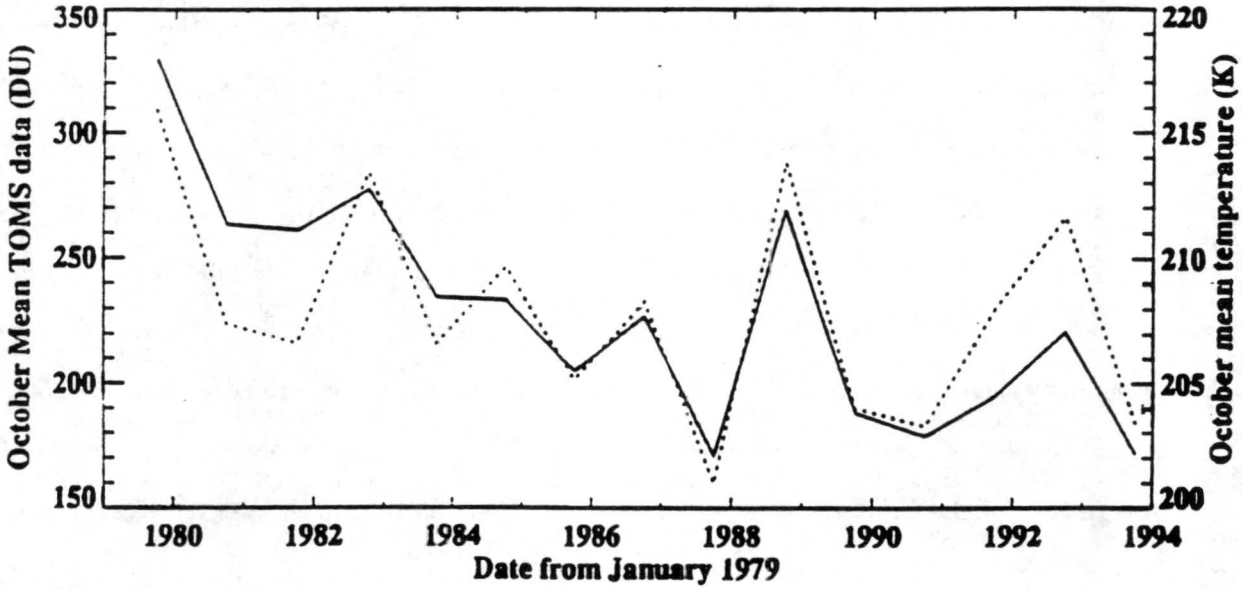


Figure 3.7: Solid curve: TOMS October mean column ozone (Dobson units) poleward of 70°S (From Butchart and Austin (1996)).

3.5 Summary

The MSU brightness temperature was shown to represent a deep-layer atmospheric temperature, defined by the convolution of a layer temperature and weighting function. The weighting function of MSU Channel 2 was found to intersect the surface in the Antarctic, introducing a surface bias into the resulting temperatures. Because the Channel 2 temperatures do not represent actual atmospheric temperatures in the polar regions, we concluded that they could not be used in our dynamical investigation of the circumpolar vortex. MSU Channel 3 brightness temperatures were found to closely resemble the ECMWF 300-200-mb geopotential thicknesses. MSU Channel 4 brightness temperatures were related to ECMWF 200-50-mb thicknesses. This is a slight augmentation of the relation used by VS. The relation between Channel 4 temperatures and ECMWF 100-mb temperatures reported by VS is still accurate.

The brightness temperature fields were shown to depict the circumpolar vortex in the lower stratosphere. A comparison of 100-mb radiosonde temperatures above the Halley Bay station with the MSU Channel 4 temperatures showed a consistency in their representations of temperature trends for the years, 1982 through 1986. Except in September, the Channel 4 temperatures were greater than the radiosonde temperatures, with a maximum discrepancy on the order of 6-9 K in November. The discrepancy was accounted for by noticing that the temperatures above 100 mb are warmer than at 100 mb from October to December. Channel 4's weighting function extends well above 100 mb, and will give weight to these warmer levels, producing the observed warm-bias. In conclusion we also demonstrated a consistency between the size of the austral circumpolar vortex and the magnitude of the October ozone anomaly. Years in which the vortex was large (and presumably isolated) the column ozone reached low levels, while the opposite occurred when the vortex was small.

Chapter 4

TECHNIQUE

4.1 Introduction

In Chapter 3 it was found that the upper tropospheric and lower stratospheric channels of the MSU are capable of depicting the circumpolar vortex. A consistency between polar column ozone anomalies and fluctuations in vortex area was also demonstrated using Channel 4 brightness temperature. While the thermal field certainly does provide some insight into the fundamental characteristics of the vortex, kinematic and dynamical quantities like velocity and vorticity are necessary to examine the vortex dynamics summarized in Chapter 2.

In the polar stratosphere the winds are in approximate geostrophic balance, except in regions of strong flow curvature. In this section it is shown that the MSU temperature, because of its deep layer characteristic, can be used to estimate the geostrophic winds up to 50 mb. This allows for the determination of the jet location in the lower stratosphere. Related quantities like vorticity and circulation are also derived, and provide further insight into the evolution and year-to-year variation of the polar vortex.

Additional details regarding our method are left to Appendix B. There, a discussion of the spectral method used to compute the velocity and vorticity is found. The spectral method allows us to not only filter the data, but to also decompose various fields in terms of different wavenumber components. Such a spectral decomposition will be used in Chapter 5 to investigate the waves forcing the austral circumpolar vortex.

4.2 Computation of Dynamical Quantities

The polar stratospheric winds are in approximate geostrophic balance,

$$f\mathbf{v} = \mathbf{k} \times \nabla_p \Phi, \tag{4.1}$$

where $\mathbf{k}$ is the vertical unit vector and Φ is the geopotential. Differentiating (4.1) with respect to pressure yields

$$\frac{\partial \mathbf{v}}{\partial \ln p} = \frac{1}{f}\mathbf{k} \times \nabla_p \frac{\partial \Phi}{\partial \ln p} = -\frac{R_d}{f}\mathbf{k} \times \nabla_p T, \tag{4.2}$$

where the second identity invokes hydrostaticity and assumes the ideal gas law for dry air. Equation (4.2) is simply the thermal wind equation in pressure coordinates. Vertically integrating (4.2) gives an expression for the geostrophic winds at level p_1,

$$\mathbf{v}(p_1) = \mathbf{v}(p_0) - \frac{R_d}{f} \ln\left(\frac{p_1}{p_0}\right)\mathbf{k} \times \nabla_p \overline{T}, \tag{4.3}$$

where $\overline{T}$ is the mean layer temperature (see Eq. (3.15)). Since the MSU measures well-defined deep layers of the atmosphere, the brightness temperature, T_B, can be used in (4.3) to compute the winds at p_1, given the winds at lower level p_0. Chapter 3 showed that the three MSU channels span the troposphere up to 50 mb. Thus, given the surface winds, successive applications of (4.3) allow us to compute the geostrophic winds in the lower stratosphere.

The first step uses Channel 2 brightness temperature, which is an 850-300-mb mean-layer temperature, to compute the geostrophic winds at 300 mb. Zero winds are assumed at the surface. The resulting zonal winds at 300 mb (not shown) are large and nonphysical in the vicinity of the Antarctic landmass. The anomalously large temperature gradient associated with the surface bias of Channel 2 is responsible. We conclude that Channel 2 is unacceptable for use in a dynamical study of the polar regions. We are then forced to use the ECMWF 300-mb geopotential to compute the geostrophic winds at 300 mb.

Channel 3 brightness temperature, which resembles the 300-200-mb geopotential thickness, is next used to compute the winds at 200 mb. Figure 4.1 shows the 200-mb geostrophic winds for July of 1992 derived from the ECMWF 200-mb geopotential. Also

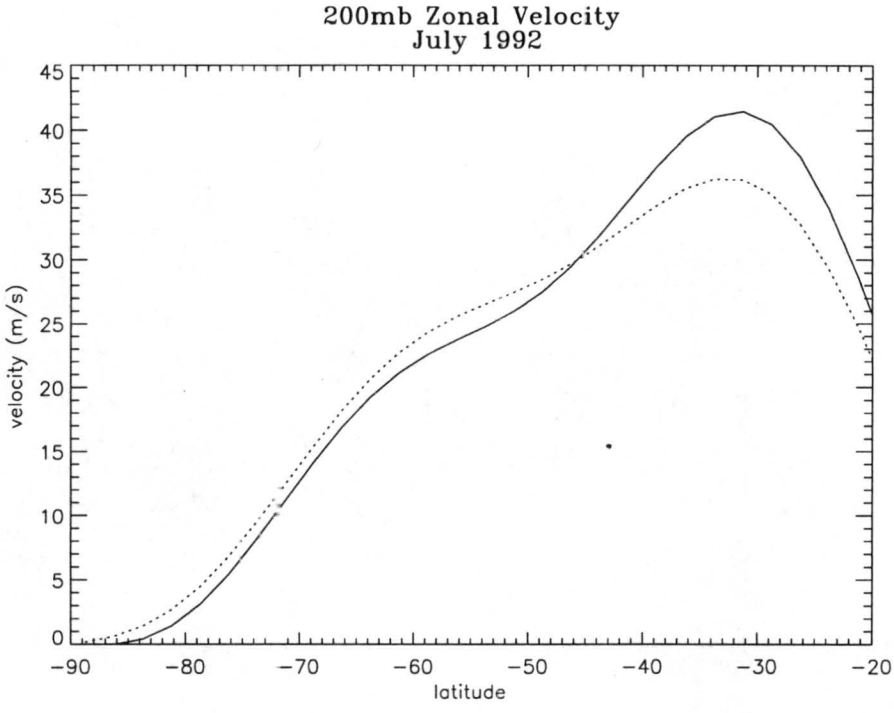


Figure 4.1: Profiles of 200-mb ECMWF geostrophic wind (solid) and MSU geostrophic wind (dotted).

shown are the winds derived using the above method. Between the south pole and 45°S, the two agree remarkably well. The difference is on the order of 1-2 ms^{-1}. Near the jet maximum (about 32°S) the geostrophic wind estimates deviate more substantially, with a maximum discrepancy of 5 ms^{-1}. The winds out to 40°S, however, are most pertinent to the study of the circumpolar vortex presented here. Thus, for the latitudes of concern, we conclude that the meridional gradient of MSU Channel 3 may be used with confidence to compute the 200-mb geostrophic winds.

Finally, we use Channel 4 brightness temperature, which was found in Chapter 3 to represent the 200-50-mb geopotential thickness, to compute the winds at 50 mb. The results for July are shown in Fig. 4.2, along with the ECMWF actual and geostrophic winds. While all three estimates of the wind agree quite nicely in form, each depicting the jet maximum around 59°S, their representations of its magnitude are dissimilar.

One contributing factor to the disagreement is related to our assumption of geostrophy. A more appropriate approximation to the actual winds, including the effects of both the Coriolis and centrifugal force, is gradient balance. In polar coordinates the radial

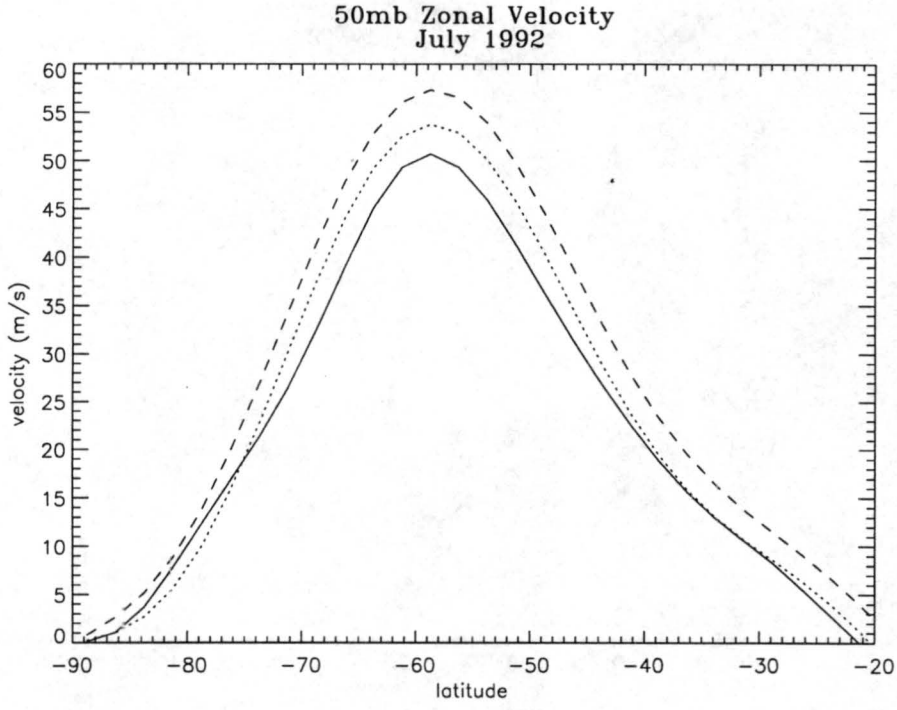


Figure 4.2: Profiles of ECMWF actual wind (solid), ECMWF geostrophic wind (dotted), and MSU geostrophic wind (dashed).

momentum equation for steady zonal flow is

$$\frac{u^2}{R} + fu = fu_g, \tag{4.4}$$

where u is the zonal wind and R is the radius from the origin. Dividing (4.4) by fu_g, gives the following ratio of geostrophic and gradient winds:

$$\frac{u_g}{u} = 1 + \frac{u}{fR}. \tag{4.5}$$

At the radius of maximum winds (RMW) of the polar vortex $f \approx 1.25 \times 10^{-4}$ s^{-1}, $u \approx 50$ ms^{-1}, and $R \approx 3300$ km. Thus u/fR (the Rossby number) is approximately 0.12. Accordingly, the geostrophic winds should be no more than 12 percent greater or less than the gradient winds. Upon comparing the ECMWF actual (gradient) and geostrophic winds in Fig. 4.2, we see that this does hold true. The source of the discrepancy between the ECMWF and MSU geostrophic winds is not as straightforward.

Closer examination of the ECMWF and MSU geostrophic winds shows a 3-4 ms^{-1} difference at nearly all latitudes. Assuming the ECMWF winds to be correct, this suggests

that our method of "stepping-up" the geostrophic winds using T_B may not be entirely accurate. For better results, a closer match between Channel 4 brightness temperature and ECMWF thickness would need to be made. This is not a trivial task, since the MSU temperatures are weighted mean-layer temperatures. The 5-10 percent error introduced using the present relations, however, will not significantly affect the qualitative results of the climatology presented in Chapter 5.

In Chapter 2 much of the discussion of the polar vortex dynamics revolved around the implied fluxes and mixing of PV. It would be beneficial, then, if we could compute the PV using the MSU data. Unfortunately, since the expression for QG PV (see Eq. (2.6)) contains vertical derivatives of temperature, we need another channel above Channel 4 to obtain it at 50 mb. Clough et al. (1985) used the MSU data to compute such PV maps, but also had to use Stratospheric Sounding Unit (SSU) data, high resolution infrared sounder (HIRS-2) data, and NMC analyses. Since our study aims to test the usefulness of the MSU in a dynamical investigation, the use of multiple data sets is inappropriate. We can, however, compute those components of the QG PV involving horizontal derivatives, namely the absolute vorticity.

Taking $\mathbf{k} \cdot \nabla_p \times$ (4.3) gives

$$\zeta(p_1) = \zeta(p_0) - \frac{R_d}{f} \ln\left(\frac{p_1}{p_0}\right) \nabla_p^2 \overline{T}, \tag{4.6}$$

where ζ is the vertical component of the (geostrophic) relative vorticity. The relative vorticity at 50 mb can be evaluated following the same methodology for determining the geostrophic winds.

Although we cannot compute the full potential vorticity, one might wonder to what extent the QG PV can be approximated by the absolute vorticity. A simple scale analysis of the QG PV suggests whether this is plausible. Denoting the ratio of the "stretching" term to the relative vorticity by 'I', we find, after scaling all variables,

$$[I] = \frac{[\Phi]}{L_R^2} / \frac{[\Phi]}{L^2} = \frac{L^2}{L_R^2}, \tag{4.7}$$

where $L_R = NH/f$ is the Rossby radius of deformation and L is the characteristic horizontal scale of disturbances. Thus, if the horizontal scale is much less than L_R, we can

neglect the term involving the vertical derivatives of temperature, and equate PV and absolute vorticity. In the stratosphere, around 60°S, $N \approx 0.022$ s^{-1} and $H \approx 7$ km, so $L_R \approx 1200$ km. The characteristic horizontal scales must be much less than 1200 km to make the approximation. For planetary waves, which are typically on the order of 10000 km, this is not reasonable. Only for the smallest disturbances will this be appropriate. The movement of larger absolute vorticity anomalies, however, may still roughly follow their PV counterparts on time scales short enough to be considered adiabatic (less than a week in the polar stratosphere).

If the wind computed by (4.3) is integrated around a closed loop with tangent path element **dl**, i.e.,

$$\Gamma(p_1) = \oint \mathbf{v}(p_1) \cdot \mathbf{dl}, \tag{4.8}$$

then the circulation is obtained. The circulation can be used as a dynamical index of polar vortex strength, analogous to the temperature-area index shown in Chapter 3. Assuming the polar vortex winds to be purely zonal, **dl** and **v** are parallel. Equation (4.8) then simplifies to

$$\Gamma(p_1) = 2\pi a \bar{u}(p_1) \cos\theta, \tag{4.9}$$

where $\bar{u}$ is the zonally averaged wind, θ the latitude, and 'a' the radius of the earth. Chapter 5 uses this scalar index to investigate the interannual variability of the circumpolar vortex.

4.3 Summary

In this chapter the variables necessary for an investigation of the dynamics of the polar vortex were derived. It was shown that the MSU brightness temperature can be used in the thermal wind equation to compute the geostrophic winds at different levels in the troposphere and lower stratosphere. A comparison of the MSU-derived winds with the ECMWF analyzed winds demonstrated that errors will be introduced into our results through the assumption of geostrophy and, more importantly, by how we choose to relate the MSU brightness temperatures to actual temperatures.

An expression for vorticity was derived in terms of the MSU brightness temperature in the hope of computing the quasi-geostrophic analogue of Ertel's potential vorticity. The vertical resolution of the MSU proves to be too low to calculate the vertical derivatives of the "stretching" term in the stratosphere. When the Advanced Microwave Sounding Unit (AMSU) is launched sometime in the 1990s, this should no longer be an obstacle; it is heralded to have 11 channels in the 50-60 GHz frequency range (VS). Nonetheless, the vorticity can still be used as a diagnostic variable and, under special conditions, a dynamical tracer.

Finally, we showed how the MSU-derived winds can be used to compute the circumpolar vortex circulation. This scalar measure of vortex strength will be compared interannually in the next chapter in an attempt to explain why the vortex is larger and stronger in some years than others. Understanding this variability is crucial to the problem of the "ozone-hole" phenomenon, and may help us to better separate the dynamical and anthropogenic contributions to ozone depletion. Ultimately we would like to be able to predict the intensity of the circumpolar vortex and, hence, the "ozone hole." First, however, we must understand the basic dynamical mechanisms responsible for the observations, both seasonal and interannual.

Chapter 5

CLIMATOLOGY OF THE AUSTRAL POLAR VORTEX

5.1 Introduction

Thus far we have motivated the formation of the circumpolar vortex through radiative considerations and the forcing processes which shape it through wave-mean-flow interaction. Using the MSU data and its derived fields, we now explore these ideas from an observational standpoint. Since a complete study of the vortex really requires more than a single vertical level at 50 mb, we focus on the lower stratosphere, briefly mentioning more thorough studies to place our results in a broader context.

First we consider the seasonal evolution of the polar vortex. Using MSU Channel 4 brightness temperatures, we discuss the thermodynamical aspects of the vortex development in the lower stratosphere. The thermal field, as found in Chapter 2, can also be used to infer wave forcing by comparing the observed and radiative equilibrium temperatures (see Section 2.3.1). In the lower stratosphere, however, wave forcing is not very large, so the measured departures may result from forcing at higher levels through downward control. Nonetheless, evidence of wave activity at lower levels can be deduced by careful analysis of PV maps (McIntyre and Palmer, 1983). Although we are generally unable to compute PV from MSU data alone, we can perform an exploratory analysis of daily maps of MSU-derived absolute vorticity. Undulations of the vortex edge signify the presence of waves. The question of which waves are most prominent in the lower stratosphere is then answered through a spectral decomposition of Channel 4 temperatures into their zonal wavenumber components.

As mentioned above, the significant wave forcing occurs in the middle and upper stratosphere. This forcing causes dramatic local changes in the zonal mean flow, with

a diminishing effect at lower levels. In addition to low-level departures from radiative equilibrium, evidence of wave forcing at higher levels can also be obtained through the wind field in the lower stratosphere. As spring is approached, a dipole pattern of zonal-mean-flow acceleration and deceleration forms. The characteristic signature of zonal-mean-flow acceleration at high latitudes and deceleration at lower latitudes is evident in the 50-mb winds derived from the MSU data

The data used to discuss the seasonal evolution is averaged over the eight years of our study (with the exception of the daily plots), providing a reliable climatological picture of the vortex. Inspection of the fields from each individual year, however, reveals an often dramatic variability in the appearance and development of the polar vortex. As others have reported, a biennial oscillation in the strength of the vortex is observed in the MSU data. In Chapter 2 we discussed a plausible mechanism linking oscillations in vortex strength to the equatorial QBO. We now investigate this link using the MSU data and attempt to establish a qualitative correlation. Wave forcing, which essentially transmits the QBO signal to higher latitudes, is weak in the lower stratosphere of the Southern Hemisphere, so irrefutable proof of a connection is not obtainable with the MSU data alone.

5.2 Seasonal Evolution

The winter-spring evolution of the austral circumpolar vortex is fairly regular in the sense that certain dynamical processes take place each year, contributing to the observed thermal and wind structures. The timing and intensity of these processes, however, show a marked interannual variability, the reasons for which are discussed in Section 5.3. Here, we create a dynamical portrait of the austral vortex from early winter (June) to late spring (October), emphasizing the role of waves in governing the climatological vortex evolution.

In Fig. 5.1 we find the MSU Channel 4 brightness temperatures for each month, averaged over the eight years of our study. Channel 4, as mentioned in Chapter 3, samples the lower stratosphere, and for our purposes, the base of the polar vortex. The meridional gradients of temperature, which define the geostrophic zonal wind, increase from June to

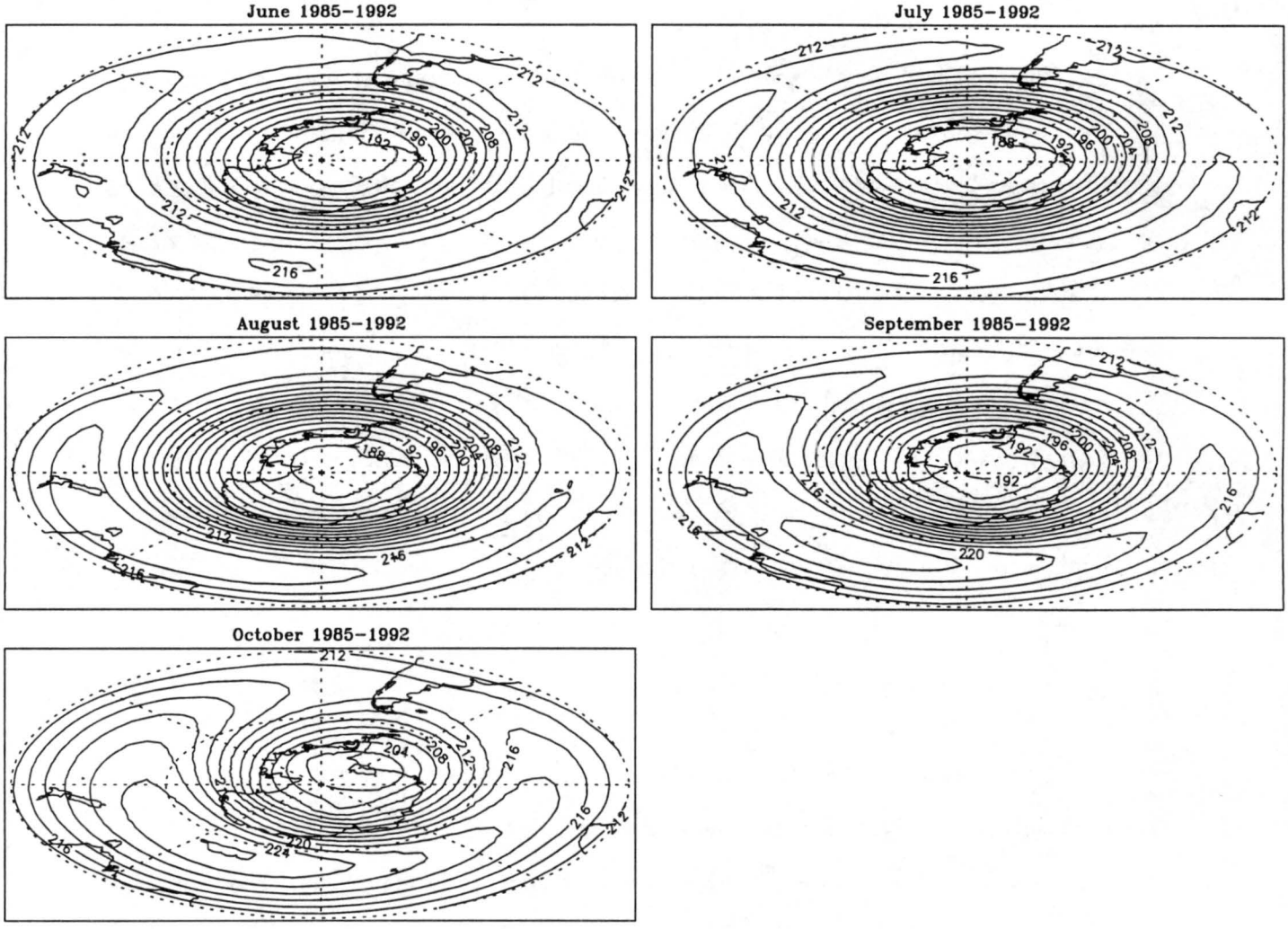


Figure 5.1: MSU Channel 4 brightness temperatures depicting the thermal evolution of the lower austral circumpolar vortex. Contour interval 2 K.

July, implying an intensification of the polar jet. As the jet in the upper stratosphere begins to weaken in August, the jet in the lower stratosphere shows continued signs of strengthening. In September, however, the warm pool centered near 50°S and 100°E starts to encroach on the pole, displacing the vortex and slightly weakening it. By October the vortex is much reduced in size, with the warm pool dominating the thermal field.

Within the vortex we observe a region of uniform temperatures on the order of 190 K (except in late spring). The corresponding winter temperatures in the Northern Hemisphere are on the order of 200 K. This is an important distinction between the two hemispheres, since it is crucial to explaining why ozone depletion is more of a problem in the Southern Hemisphere. Heterogeneous reactions occurring on the surfaces of polar stratospheric clouds (PSC's) are known to remove nitrogen species from the lower stratosphere, thus enhancing the spring-time destruction of ozone by chlorine radicals. PSC's form only at very low temperatures, however, generally less than 195 K. In the Northern Hemisphere temperatures occasionally fall below this value, but the areas involved are inconsequential for widespread depletion to occur (Newman et al., 1990). The affected areas in the Southern Hemisphere (determined solely by temperature, see Fig. 3.6) are large, on the order of 10 to 20 million square kilometers.

Consistent with the aforementioned increase in jet winds is a drop in the central temperatures of the vortex from June to July. In the absence of wave forcing the temperatures would continue to drop as the vortex relaxed to its radiative state (about 183 K in the central region of the vortex in July). In August, however, the vortex temperatures in the lower stratosphere have ceased to decrease. In the middle and upper stratosphere a warming on the order of 0.3-0.4 K day^{-1} now extends from the midlatitudes into the polar regions (Mechoso et al., 1985). This is associated with an erosion of the circumpolar vortex (i.e., deceleration of the westerly jet). The central temperatures in the lower stratosphere increase rapidly after the destruction process has reached 100 mb. By October they jump some 12 to 14 K. Ozone-depleted air, produced within the vortex as the Sun reappears in September, now mixes with the unprocessed air outside the vortex. The result is a reduction in column amounts at lower latitudes, causing some concern for the inhabitants of the Southern Hemisphere.

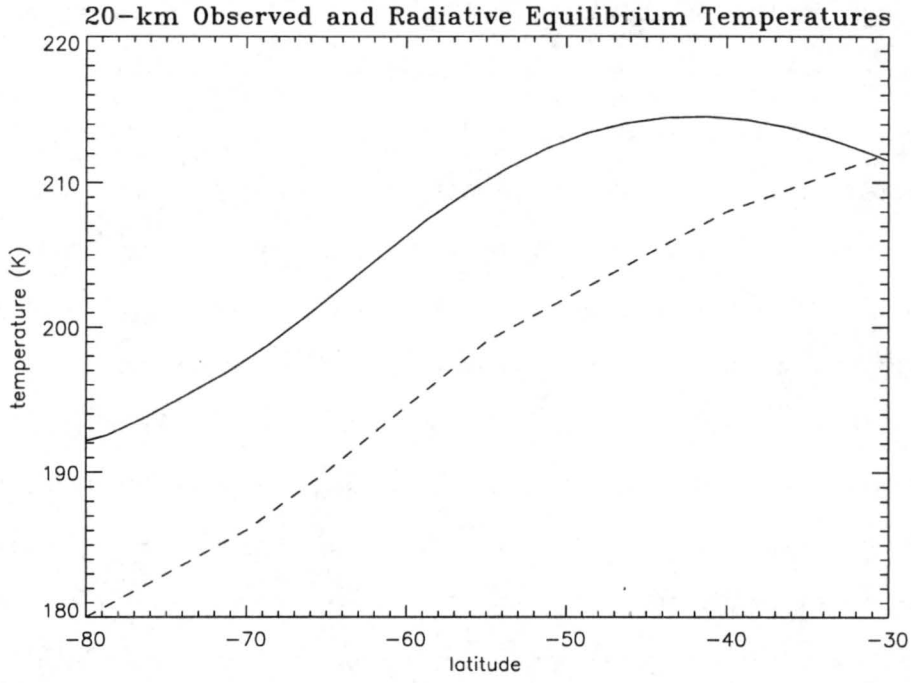


Figure 5.2: Comparison of MSU Channel 4 brightness temperature (solid) and the 20-km radiative equilibrium temperature (dashed) for June. MSU data is averaged over 1985 to 1992. Data for the equilibrium temperature adapted from Wehrbein and Leovy (1982).

To explain this sequence of events we must investigate the the factors which significantly influence the vortex, i.e., the waves. The MSU 100-mb temperatures can be used for this purpose. According to Eq. (2.35) a meridional gradient of the E-P flux divergence translates directly into a departure from radiative equilibrium conditions. If the observed temperatures deviate from T_r, then the amount of wave forcing (resulting from the propagation of waves from remote regions of the atmosphere as well as in-situ instabilities) may be inferred. The forcing responsible for the temperature departure, contrary to Eq. (2.35), may not be local; by downward control, deviations from T_r also depend on the wave activity at higher levels (see Eq. (2.36)). Wave forcing in the lower stratosphere is, in fact, very small (see Fig. 2.1). Therefore, the observed departures there must result from wave activity in the middle and upper stratosphere. At the edge of the polar vortex we expect a strong convergence of the E-P flux associated with downgradient transport of PV. Poleward of the forcing region, one might naively assume a diminishing E-P flux convergence. What we actually observe is a region of divergent E-P flux, creating what is

often referred to as a "dipole structure" (Mechoso et al., 1985). The origins of this positive anomaly are not well understood, although some have tied it to in-situ instabilities poleward of the polar jet. The important thing, however, is that the meridional gradient at the vortex edge is positive, implying that temperatures will exceed T_r. Figure 5.2 shows the MSU Channel 4 temperatures and T_r for June. Indeed, the actual temperatures do exceed the radiative equilibrium temperatures, but the latitudinal structure of the departure is somewhat counterintuitive. We might expect a maximum departure from radiative equilibrium conditions near the maximum wave forcing around 40°S to 50°S. What we observe is a maximum departure closer to the pole, around 70°S to 80°S. This may be due to the fact that the radiative damping rates in the lower stratosphere are small at the pole and increase out to about 40°S, as discussed in Section 2.3.1. Thus, temperature departures below the upper-level forcing will be damped out more quickly than at the pole, producing the observed diminishing discrepancy at lower latitudes.

Direct evidence of waves in the lower stratosphere can be found by analyzing maps of Ertel PV on isentropic surfaces (e.g., McIntyre and Palmer, 1983; Clough et al., 1985). For adiabatic timescales (approximately a few days in the stratosphere) PV is materially conserved. Thus, one can follow isolines of PV in time as if they were material tracers. Although we showed in Chapter 4 that the absolute vorticity is generally not equivalent to PV, we should be able to infer the presence of waves through distortions of the vortex edge as defined by the absolute vorticity. In the absence of any waves the isolines of absolute vorticity will be circumpolar. Figures 5.3 and 5.4 show the 50-mb absolute vorticity every four days for August and September of 1992. The vortex edge near 60°S undulates in time, responding to the superposition of various wavenumber forcings. The vorticity anomalies associated with these waves are clearly visible inside and outside the band of large meridional gradients. There appears to be some continuity in the evolution of these anomalies, suggesting they may be physical, and not just the result of data analysis errors or other data-related problems. For example, the cyclonic anomaly located near 180° longitude and 70° latitude on August 9th shows signs of eastward propagation through the 17th. The daily evolution plots (not shown) confirm this.

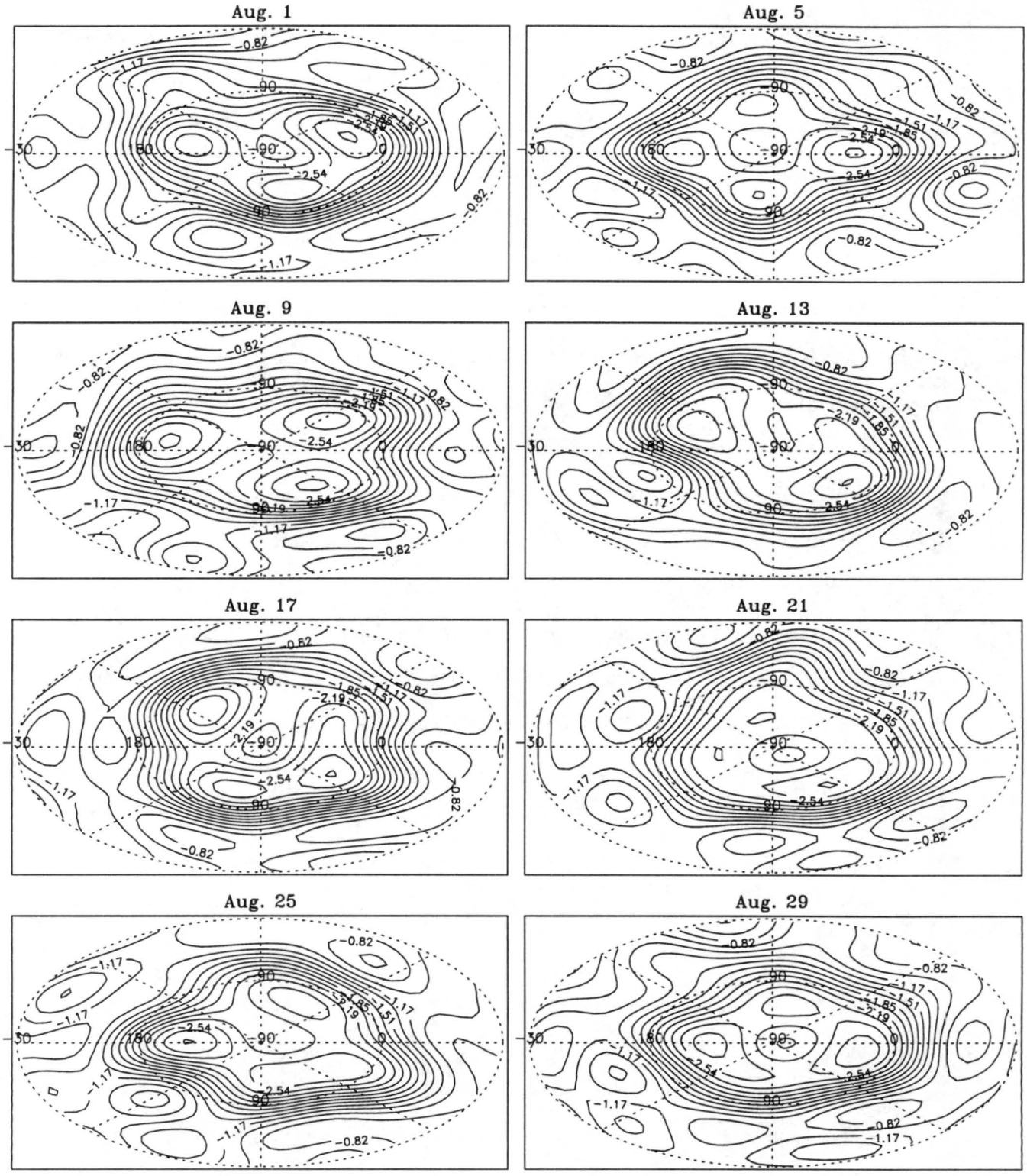


Figure 5.3: 50-mb absolute vorticity derived from August 1992 MSU data. The contour interval is $1.25 \times 10^{-5} s^{-1}$ scaled by Ω.

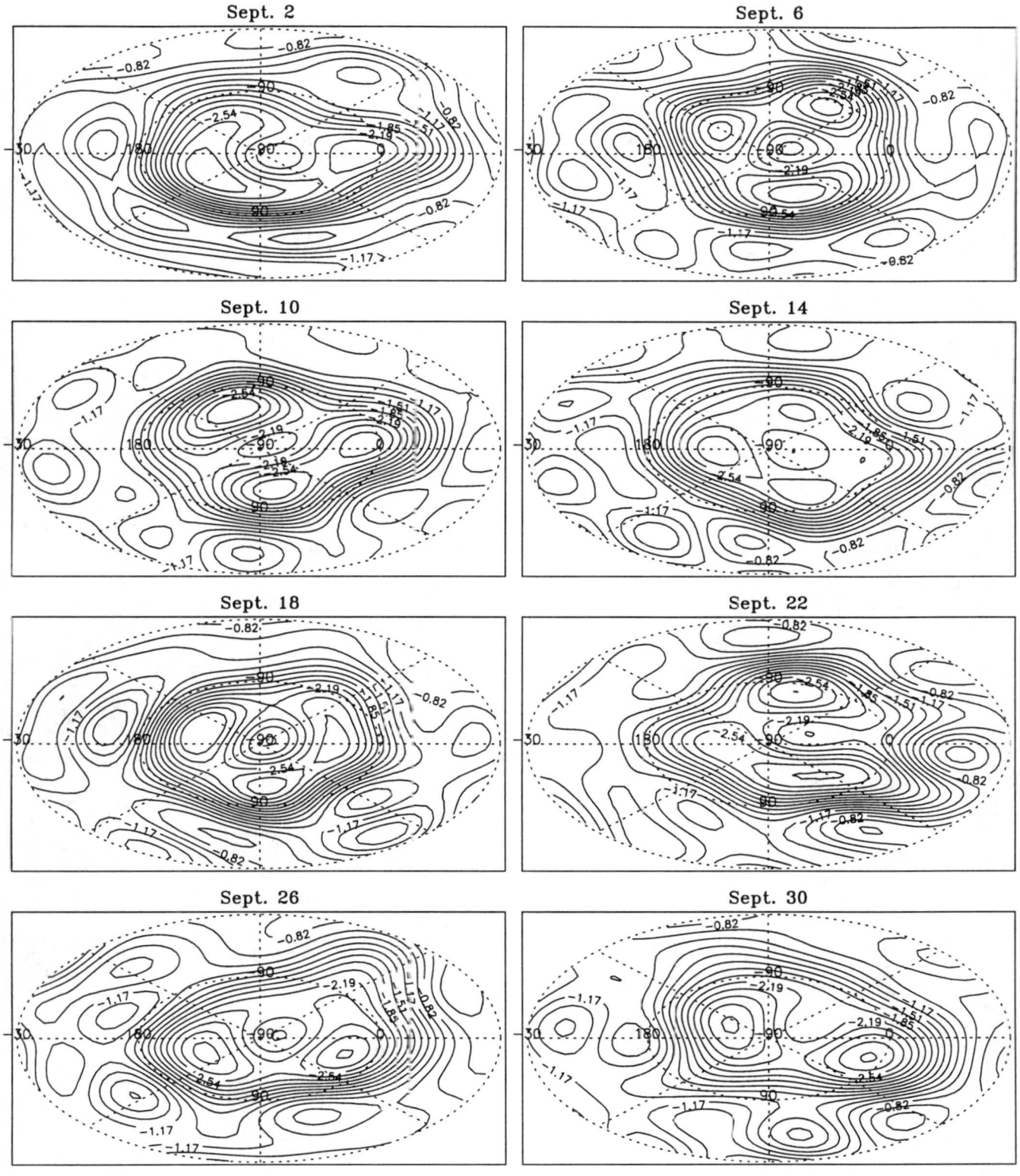


Figure 5.4: 50-mb absolute vorticity derived from September 1992 MSU data. The contour interval is $1.25 \times 10^{-5} s^{-1}$ scaled by Ω.

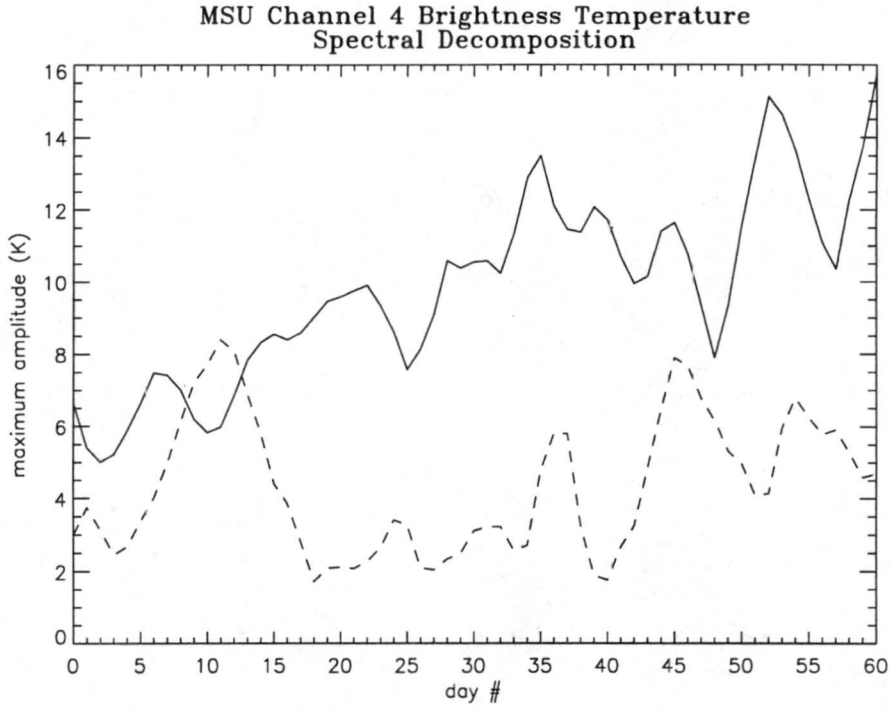


Figure 5.5: Spectral decomposition of MSU Channel 4 brightness temperature into wavenumber one (solid) and wavenumber two (dashed) components for August and September of 1992.

The waves responsible for the forcing at the edge of the vortex are planetary in scale and of low zonal wavenumber. The major contributors to the observed E-P flux convergence anomalies are quasi-stationary wavenumber one and eastward traveling wavenumber two (Hartmann, 1976). The eastward traveling wave has a period of about two to three weeks in the stratosphere. Shiotani and Hirota (1985, hereafter S&H) in their one-year investigation of the polar stratosphere found that, in the upper stratosphere, wavenumber two dominates during late winter (July-August) and wavenumber one dominates from early spring on (September on). Although the relative strength of the various wavenumber components may vary with height due to the different mean-flow structures, and hence potential barriers to propagation, wavenumbers one and two should also dominate in the lower stratosphere. To see whether we can obtain findings similar to S&H and others using the MSU, we spectrally decompose the MSU Channel 4 temperature field into its zonal wavenumber one and two components. The results are shown in Fig. 5.5 for the early spring period, August through September. Indeed, wavenumber one is prominent

throughout the period, with the exception of a brief burst of wavenumber two activity during early August. From the beginning of August to the end of September wavenumber one increases, attaining a maximum amplitude of approximately 16 K. This is consistent with Quintanar and Mechoso's (1995) observation that quasi-stationary wavenumber one generally peaks around September or October in the upper troposphere. The increase in amplitude of the wavenumber two component between the 5th and 15th of August is followed by a two-week lull in activity, with much smaller amplitudes on the order of 2 K. Following the lull, wavenumber two peaks three more times in September. The maximum peak occurs around mid-September and is followed by a smaller jump in amplitude ten days later. Such wavenumber two activity in September was also observed by Manney et al. (1991) in other years. Their time-height cross-sections of wavenumber two geopotential height amplitude show the largest growth in the upper stratosphere. The larger wavenumber two events had a vertical structure extending down into the upper troposphere, suggesting baroclinic instability in the troposphere as a source. The weaker events did not appear to be connected with the troposphere, and may have resulted from local instabilities of, say, the polar night jet. Through late September and October the magnitude of these temporary increases in amplitude diminish. While this trend becomes apparent in the MSU wavenumber two temperatures during the latter half of September, the trend in October must be known to confirm their observations. We conclude that the qualitative characteristics of the primary wavenumbers in the lower stratosphere, as observed by the MSU, are consistent with previous observations.

As discussed in Chapter 2, we recall that the forced tropospheric waves generally increase in amplitude as they propagate into the stratosphere due to the decrease in density and narrowing of the waveguide. By the time they reach the middle to upper stratosphere, the amplitudes often become large enough to "break", leading to a deceleration of the mean flow. When the wave forcing occurs at the edge of the circumpolar vortex, a weakening of the jet ensues. We have already observed the consequences of this upper-level phenomenon in the lower stratosphere through departures from radiative equilibrium. As the waves continue to disrupt the vortex, the jet eventually moves poleward and downward. The displacement of the jet into the lower stratosphere completes the life-cycle of

the circumpolar vortex. We now investigate this phenomenon, again using the MSU data to establish consistency with the observations of previous studies.

For the winter-spring period of 1981, S&H investigated wave-mean-flow interactions in the Southern Hemisphere stratosphere. S&H outlined the seasonal evolution of wave activity and zonal wind, which we now summarize here. During early winter they observed weak wave activity. The polar jet in the upper stratosphere was located near 40°S, while in the lower stratosphere it was observed in the vicinity of 60°S. As winter progressed, wavenumber two activity started to increase at upper levels, occurring in periodic bursts. Following this activity, the polar jet suddenly shifted poleward and downward (about 20° in a week). Soon after, wavenumber one activity increased in the stratosphere, leading them to conclude that the poleward shifting of the jet created a more favorable mean-flow structure for wave propagation. In a later study Shiotani et al. (1990) found that the 1983 jet shifted after a burst of wavenumber one activity. Although the jet shift may result from any number of different wave forcings, observations show that the end result is usually an enhancement of wave activity in the stratosphere.

In a four-year climatology of wave-mean-flow interaction in the Southern Hemisphere, Mechoso et al. (1985, hereafter M) found that the initial shifting of the jet typically occurs in one or more stages, followed by a monotonic descent through September. Their plots of zonal-mean wind tendency (see their Fig. 2) lend some insight into what we should expect as far as changes in the mean-flow at 50 mb. In June they show acceleration throughout most of the stratosphere. In July a "dipole structure" in zonal-wind tendency appears, related to the structure in E-P flux divergence previously discussed, with maximum deceleration centered about 40°S in the upper stratosphere and maximum acceleration centered about 55°S in the middle stratosphere. The region of acceleration, however, is not confined to the upper levels; it extends down to 50 mb, where it has a magnitude (averaged over the four years of their study) of about 0.2 $ms^{-1}day^{-1}$. In August the region of deceleration extends down to 50 mb. There the line of demarcation between acceleration and deceleration falls near 45°S.

In Fig. 5.6 we have plotted our 50-mb winds at 40°S, 50°S, 60°S, and 70°S for the period from August to September. Although our results are only for a single year, and

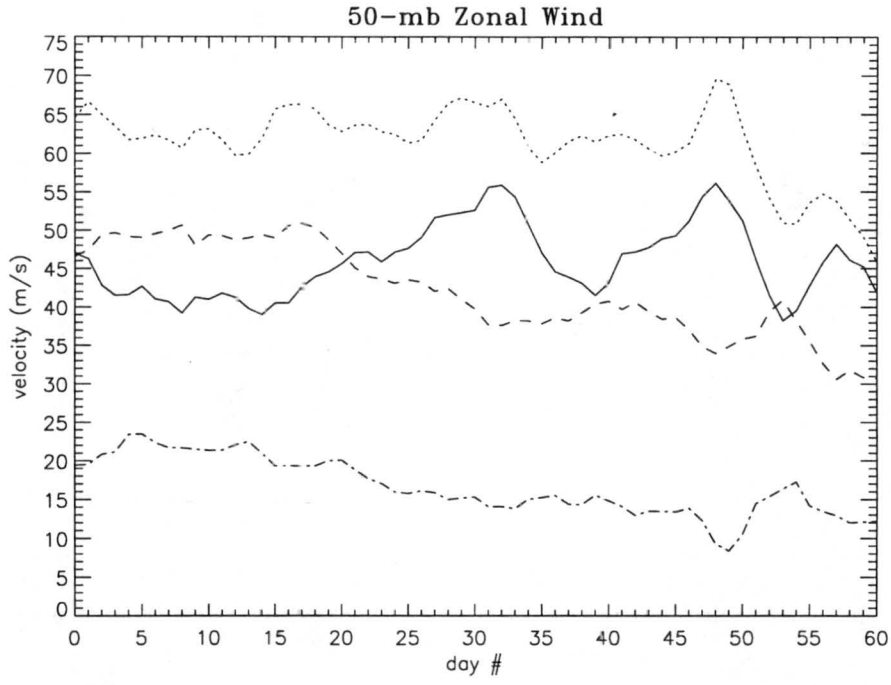


Figure 5.6: 50-mb zonal wind (ms^{-1}) for August-September of 1992 at radii inside and outside the radius of maximum wind, 58°S. Winds at 40°S (dash-dot), 50°S (dashed), 60°S (dotted), and 70°S (solid) are shown.

the timing of the accelerations and decelerations at 50 mb may not agree with the climatological observations of M, the latitudinal structure should be consistent. At 40°S we certainly do observe a deceleration of the winds throughout August on the order of 0.2 $ms^{-1}day^{-1}$. Poleward of the line of zero wind acceleration at 50°S, we observe nearly constant winds during the first 15 to 20 days of August, followed by a dramatic decrease through the rest of the month. Further poleward we should observe an increase in the zonally-averaged winds. At 60°S we do not observe any noticeable trend, but at 70°S there is a clear acceleration through the latter half of August.

By September M found that the entire stratosphere is dominated by a negative tendency in zonally-averaged winds. At 50 mb the maximum amplitude occurs around 50°S and is of the order -0.5 $ms^{-1}day^{-1}$. Our results at 50°S show a tendency of -0.4 $ms^{-1}day^{-1}$ at 50 mb, in agreement with their climatological value. Elsewhere we also observe a negative acceleration for September, although the large variability at the highest latitudes makes this difficult to quantify.

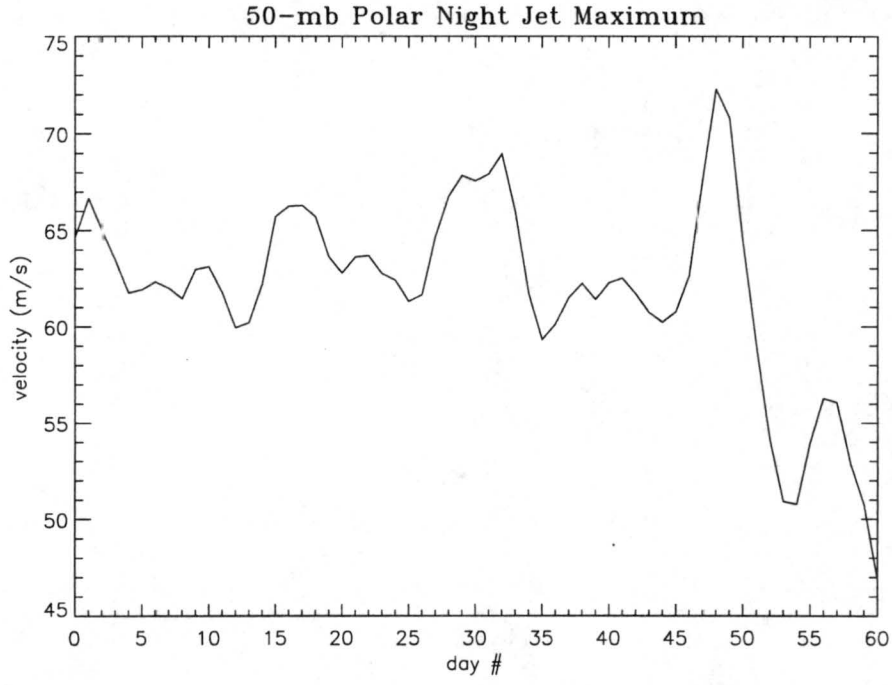


Figure 5.7: Southern Hemisphere PNJ winds (ms^{-1}) at 50 mb for August-September of 1992.

According to S&H, the shifting of the jet in the Northern Hemisphere is followed by an erosion of the westerlies, which are eventually replaced by easterlies. The westerlies do not return until the next winter season. In the Southern Hemisphere, when the jet finally reaches its resting position in the middle to lower stratosphere, it begins to oscillate in strength. In Fig. 5.7 we see that at 50 mb the jet winds cycle between periods of relative inactivity and sudden accelerations. Following the increases in wind speed, the magnitude tends to return to a value between 62 and 63 ms^{-1}, except in late September when the vortex is undergoing rapid destruction. Although it appears that the MSU may have captured the lower-stratospheric extent of this jet oscillation observed by S&H, a further investigation involving more years may be required for conclusive evidence.

5.3 Interannual Variability

The reasons for the observed interannual variability in the seasonal evolution of the austral polar vortex are still under question. Beginning with Holton and Tan's (1980) observation of a possible modulation of the global circulation at 50 mb by the QBO,

much attention has focused on trying to describe the dynamical mechanisms involved. As TYb note, however, current measurement and data analysis techniques are inadequate to conclusively show such a signal in the circulation, although they were able to infer an extratropical QBO signal in column ozone. Furthermore, they demonstrated that the amplitude of this signal must result from an actual extratropical circulation anomaly, and not just a transport from the Tropics by the climatological circulation.

We begin this section with a brief description of the vortex evolution for each of the eight years of our study, comparing our findings with those of the more thorough study of Atkinson (1993, hereafter ATK) [1]. The causal relationship between the intensity of wave activity, vortex strength, and QBO phase becomes apparent. To qualitatively correlate the variability of the vortex with the phase of the QBO, we then use the circulation index derived in Chapter 4.

5.3.1 1985

ATK characterizes the wave activity of 1985 as "relatively high", which, since the QBO is in its westerly phase, appears to contradict our conclusion, based on theoretical grounds, that the vortex should be less disturbed when planetary waves are free to propagate equatorward without obstruction. The temperature fields, however, show that this is an average year (determined from our eight-year climatology) from June to September, as far as vortex evolution is concerned. In late spring ATK observe an unusual lack of strong wavenumber one forcing, which is responsible for the displacement of the vortex from the pole and the subsequent final warming. It follows that the temperatures in October are around 6 K colder than average. The vorticity gradients remain relatively strong, implying a delayed or very slow vortex destruction.

5.3.2 1986

Below 40 mb the stratospheric winds are westerly throughout both seasons. Above this level the descending easterlies of the easterly phase of the QBO are evident. It is

---

[1] Refer to Appendix A for monthly averages of Channel 4 brightness temperatures and MSU-derived 50-mb absolute vorticity in the Southern Hemisphere. The climatological reference is found in Fig. 5.1.

unclear how far down the easterlies must extend, and how early in the season they must appear, to have a significant effect on the direction of wave propagation. The presence of easterlies above 40 mb may be enough to explain the high level of continuous wave activity observed by ATK. The continuous forcing apparently takes its toll on the vortex, as it erodes quite rapidly. The Channel 4 temperatures are slightly warmer than average during August and September, and around 4 K warmer in October. The vorticity gradients in October are very weak, evidence of the erosion discussed by ATK. Usually accompanying the final warming of the vortex is a noticeable displacement off the pole. In this year, however, the vortex remains circumpolar through October.

5.3.3 1987

1987 is the well-documented year of record ozone depletion in the austral stratosphere. Numerous studies find the polar vortex to be isolated with a broad region of temperatures below 194 K (necessary for the formation of PSC's). Prior to July, the QBO below 30 mb is in its easterly phase. Throughout the rest of the period it is in its westerly phase at most significant levels. From June to August the temperatures are near average. In September, when the vortex usually warms significantly, the area bounded by the 194-K isotherm observed in our temperature plot is much greater than on average. By October the vortex is still visibly circumpolar and the temperatures are on the order of 4 K colder than average. The maps of absolute vorticity also show the signature of a relatively undisturbed vortex, with vorticity isolines arranged almost zonally in September and October. Most other years show dramatic undulations in the absolute vorticity at the vortex edge by this time, perhaps the signature of wave activity. The fact that the vortex appears to be isolated lends credence to the theoretical argument that planetary waves are more prone to follow their climatological equatorward path when tropical stratospheric westerlies are present.

5.3.4 1988

During the winter-spring period, the equatorial stratospheric winds are easterly above 30 mb and westerly below. This is somewhat similar to that observed in 1986. As in 1986, we also observe a highly disturbed vortex. ATK characterizes the wave activity throughout

the spring as "outstandingly high." Temperatures within the vortex are close to average during early winter. In August, however, ATK finds stationary wavenumber one activity to be very intense. This is clear in our plot of lower stratospheric temperatures. The displacement of the vortex from the pole appears to have begun in August, earlier than any year of our study. By October the vortex is clearly eroded with weak absolute vorticity gradients, as in 1986. Another striking deviation from the climatological mean is the temperature of the warm pool, which is nearly 10 K warmer than average. This is consistent with ATK's observation of a very warm polar stratosphere. All these observations point to enhanced wave activity in the polar regions, expected during the easterly phase of the QBO (or at least middle stratospheric easterlies in the Tropics).

5.3.5 1989

The evolution of the vortex in 1989 is average, with no large wave events observed in the early part of the season, except for October. In June the temperatures are close to the climatological mean, while from July to October they are about 2 K colder. The vorticity maps look similar to those of 1985, with weak disturbances of the isolines at the vortex edge, unlike the more dramatic undulations in 1986 and 1988. Correlation with the QBO is not as straightforward for this year. Tropical stratospheric winds are easterly, except for descending westerlies at 10 mb in August and 20 mb later in the season. The dominance of easterlies throughout the season in the 30-50-mb layer would suggest a more disturbed vortex. One possible reason for the seeming contradiction may be the (climatologically-speaking) weak wave activity observed in the Southern Hemisphere compared to the Northern Hemisphere. Less of a distinction between westerly and easterly phases of the QBO may then be expected in the Southern Hemisphere. This, along with other possible reasons for discrepancies between observation and theory, will be explored in Section 5.3.9.

5.3.6 1990

As far as the authors are aware, little information on wave activity in the Southern Hemisphere is available for the years following 1989. We can, however, still give a

qualitative discussion of how each year's vortex evolution compares with climatology, and whether the observed strength is consistent with the phase of the QBO.

The QBO in 1990 is in its westerly phase, especially in the 30-50-mb layer. As expected of a more isolated vortex, the temperatures from June to October are about 2 K colder than the climatological values. The absolute vorticity also appears to be relatively undisturbed. In September the vortex does move off the pole, but unlike other years, it returns in October.

5.3.7 1991

1991 provides another apparent contradiction to the theory presented in Chapter 2. Above 50 mb the equatorial winds are easterly throughout the period. The vortex in June is slightly warmer than average, but becomes progressively colder than the climatological vortex. In October the vortex is well off the pole, but still colder than average. It should be noted, however, that the warm pool is 4 to 6 K warmer than average in September and October.

5.3.8 1992

The equatorial winds in the stratosphere shift from easterly to westerly at 30 mb during the period of observation. Consistent with enhanced wave activity, the vortex is warmer than average throughout the winter-spring evolution. A perhaps more convincing sign of large wave forcing in the stratosphere is the presence of a significant wavenumber one pattern in October, quite similar to that of 1988. The absolute vorticity gradients at the vortex edge appear to be very weak, as expected of a vortex undergoing an early erosion.

5.3.9 Polar Vortex Modulation

In the Northern Hemisphere studies show a nearly consistent correlation between the dynamics of the polar stratosphere and the phase of the tropical QBO (Holton and Tan, 1980; Labitzke, 1982; Holton and Austin, 1991). As shown in our year-by-year analysis of wave activity in the austral stratosphere, such a correlation is not as evident. While in

some years the vortex is indeed more disturbed when the tropical stratospheric winds are easterly (e.g., 1986), in other years like 1991, we cannot draw such a clear-cut conclusion. Using the circulation index derived in Chapter 4, the correlation is further investigated. Possible reasons for discrepancies are then discussed.

Figure 5.8 shows the vortex circulation averaged over the period from June to September. Overlaid are the 30-mb Singapore winds for July. The circulation index, while it does capture changes in vortex intensity, is biased when the vortex is far from circumpolar. In the Southern Hemisphere the assumption of polar symmetry is reasonable, but tends to break down as the spring increase in wavenumber one activity shifts the vortex off the pole. Therefore, we use this index primarily for qualitative comparisons of vortex strength.

From 1985 to 1992 the average circulation oscillates in strength with a period of roughly two years. Consistent with the theory linking the strength of the vortex to the equatorial winds, we then expect a similar oscillation in the phase of the QBO. For the first few years we do observe such a correlation. In 1989, however, the circulation increases in spite of the tropical easterlies. Similar discrepancies occur in the following years.

Further insight into the circulation is obtained when we look at each month individually. Figure 5.9 shows the circulation for June though September (denoted by the numerical labels, 6 through 9) for each year. Of interest is how the circulation changes from August to September. Years in which the percentage change is small signify a resilient vortex, i.e., one that is relatively isolated from wave activity, and is thus long-lived. If, on the other hand, large drops in circulation take place between August and September, then we may infer an enhanced erosion of the vortex by waves.

The only years for which the vortex circulation does not change dramatically are 1985, 1987, and 1989. Both 1987 and 1989 were characterized by weak wave activity, and 1985 was found to be lacking in strong wavenumber one forcing. The weak forcing and delayed erosion in 1989, however, are not expected given the narrowed polar westerly waveguide found during the easterly phase of the QBO. All other years show a significant drop in circulation. The early erosions in 1986 and 1988 are consistent with the high wave activity and easterly QBO phase reported by ATK. Without detailed accounts of wave

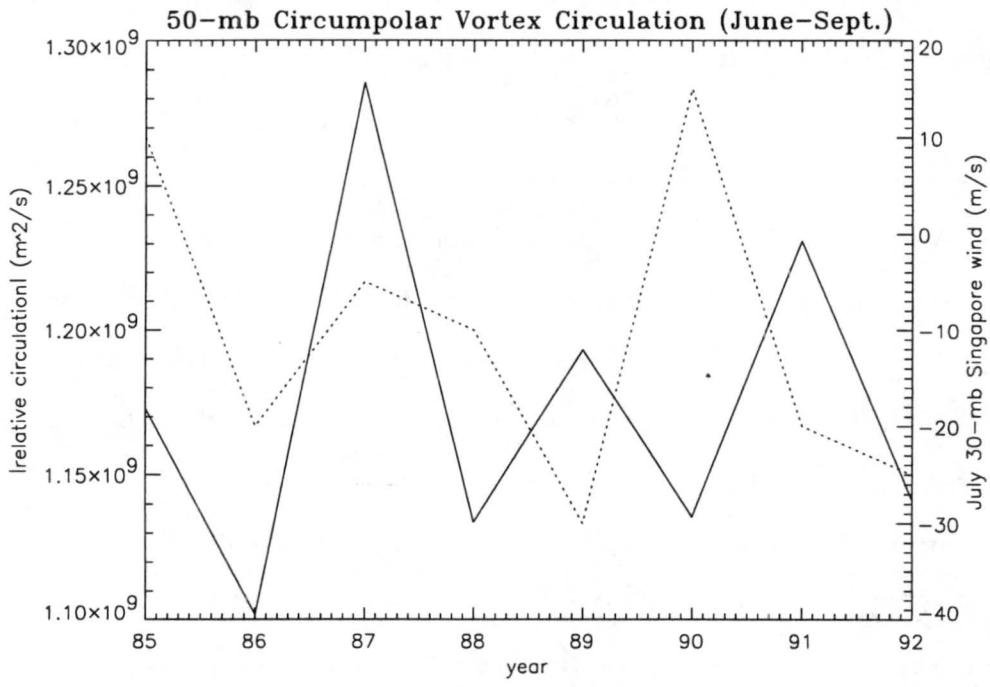


Figure 5.8: 50-mb circulation bounded by the RMW of the austral polar vortex, averaged over the period June-September (solid). Also shown are the 30-mb equatorial winds for July adapted from Holton (1992) (dotted).

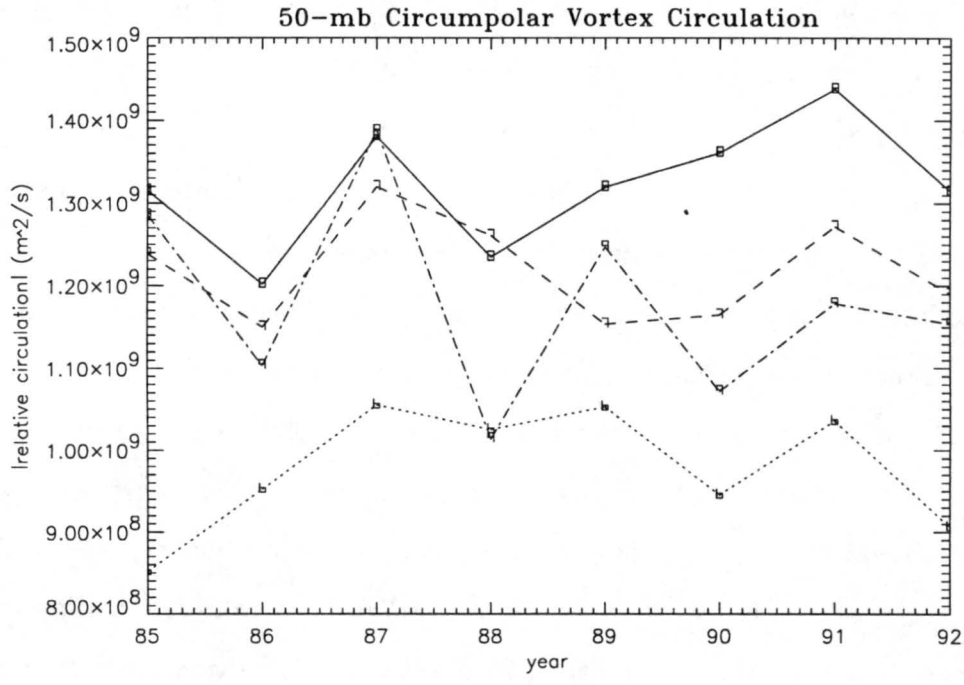


Figure 5.9: 50-mb circulation bounded by the RMW of the austral polar vortex. Each month's circulation is shown separately for June-September (6-9).

activity for the years following 1989, it is difficult to speculate why 1990, 1991, and 1992 all observe large decreases in circulation. The decrease in circulation between August and September in 1990 may be due to an early displacement of the vortex off the pole. In October the vortex actually moves back to the pole. The relatively strong meridional gradient in absolute vorticity observed in October is similar to that observed in other years when the QBO is in its westerly phase.

Although a firm link between the phase of the QBO and the strength of the circumpolar vortex is not clear in the present analysis, this does not invalidate the dynamical mechanism for interannual variability discussed in Chapter 2. A strong signal of the QBO in the wave field may exist, but may fail to be realized in the forcing of the mean flow. This could occur, for example, if the tropospherically-forced wave amplitudes are too small to permit large growth and mechanical dissipation (i.e., breaking) in the stratosphere. A minimum amplitude may be required in order to distinguish between westerly and easterly QBO phases. The greater consistency observed between the phase of the QBO and low-frequency variability in dynamical fields of the Northern Hemisphere, where wave amplitudes are larger, suggests such a constraint. Also, the stronger westerly winds associated with the austral circumpolar vortex may inhibit vertical propagation, causing waves to dissipate at lower levels before their amplitudes have a chance to grow substantially.

Other sources of interannual variability may also contribute to our circulation index. Tropospheric sources of variability, like the Southern Oscillation, could be important. Internal sources of stratospheric variability should also be taken into consideration (Pawson and Kubitz, 1996).

Even if the wave activity is significant and the QBO does play a principle role in the modulation of the vortex, the question of which level and month of the QBO to use in the correlation remains to be answered. In the Northern Hemisphere Holton and Tan (1980) found the best correlation between the PNJ strength and the equatorial winds at 50 mb. In the Southern Hemisphere Lait et al. (1989) used the August-September 30-mb, 50-mb, and 30-50-mb winds in their computations of correlation coefficients. Garcia and Solomon (1987) used the October 50-mb winds. At this time there does not appear to be a consensus

on which level and month of the QBO to use. As Gray and Ruth (1993) suggest, there may not be a unique level and month appropriate for every year; the correlation is highly sensitive to the descent rate of the QBO phase and the time at which the descent takes place. It is agreed, however, that when correlating the QBO with a dynamical field, like circulation, in the polar regions, the strongest QBO anomaly will be found in the middle to upper stratosphere where the wave forcing is greatest. Discrepancies highlighted in our qualitative correlation with the circulation in the lower stratosphere may have resulted from the small signal there.

For quantitative correlations of extratropical fields (dynamical or chemical) with the QBO, one must also consider the coupling of the annual cycle with the tropical QBO. Subtracting the frequency of the tropical QBO ($\approx$ 1/30 mo) from the frequency of the annual cycle (1/12 mo), one obtains a signal with a period of about 20 months. TYa note that many studies in the past only retained the tropical QBO signal in the extratropics and filtered away the 20 month signal. As TYa demonstrate, however, this signal actually increases toward the pole, becoming as large as the tropical QBO signal. The combined signals still do not fully account for the observed ozone anomaly amplitudes in the extratropics ($\approx$ 20-30 Dobson units), but TYa suggest that this is probably due to the fact that the actual QBO is not a pure harmonic. Passing only the 20 and 30 month signals will certainly lead to an underestimation of the exact contribution of the QBO to interannual variability in the extratropical stratosphere.

5.4 Summary

We have successfully utilized the MSU brightness temperatures in a dynamical study of the polar stratosphere. The thermal field represented by the Channel 4 temperatures accurately depicted the climatological seasonal evolution of the circumpolar vortex. We found a strengthening of the lower stratospheric vortex through August, followed by a 12-14 K warming by October. The temperatures did not fall to their radiative equilibrium values, suggesting that wave forcing of the vortex maintained the temperatures above those expected of a radiatively driven vortex. It is interesting to note that the latitudinal profile

of the discrepancy did not show a maximum below the region of large wave forcing (around 40°S to 50°S). We believe that this may be due to the fact that the lower-stratospheric damping rates in the midlatitudes are greater than in the polar regions, resulting in a more prolonged anomaly at high latitudes.

Our investigation of the waves responsible for the forcing in 1992 showed quasi-stationary wavenumber one to be dominant through late August and September. The second largest contribution, wavenumber two, was found to increase dramatically in early August, followed by a second burst in mid-September. Wavenumber two showed signs of a decline in amplitude through late September, but we were not able to confirm this. Observations of mean-flow changes induced by these waves were also presented and found to be consistent with previous studies.

Following a year-by-year climatology of the circumpolar vortex for 1985–1992, we presented our MSU-derived circulation index to investigate the theoretical link between the QBO phase and the strength of the circumpolar vortex. The early years of our study appeared to show a correlation between weak(strong) circulation and easterly(westerly) QBO phase. In 1989 and the years that followed, the correlation became less clear, and even contradictory. We concluded that the primary reason for the inconsistency between theory and observation may have been the weak wave forcing in the lower stratosphere. Although by downward control reasoning the signal of stronger wave activity in the upper stratosphere should make its way down into the lower stratosphere, it will be greatly reduced. The resulting QBO signal in the lower stratosphere may be statistically insignificant. Future work on this problem would certainly have to focus on the middle and upper stratosphere.

Chapter 6

CONCLUSION

The MSU data has been shown to be a useful tool in a dynamical study of the polar stratosphere, although not a sufficient one. The data in the polar regions, with the exception of Channel 2, was found to be both reliable and consistent with other observations. Based on these findings as well as prior studies using the MSU (Vroman and Stephens, 1989; Clough, et al., 1985), we have developed a method, taking advantage of the deep-layer measurements of the MSU, to construct dynamical fields (e.g., vorticity, velocity, and circulation) in the lower stratosphere. The computed zonally-averaged velocity was similar to that from the ECMWF, with some discrepancies resulting from our assumption of geostrophy and mean-layer representation of the brightness temperatures. We then used these fields to explore the austral circumpolar vortex, discussing previous observations and addressing current theoretical ideas.

We found that the vortex evolution is shaped by wave forcing in the extratropics, causing temperatures to rise above their radiative equilibrium values. Wavenumbers one and two were found to dominate in the austral lower stratosphere in 1992, as others had observed in more general studies of waves in the stratosphere. Wavenumber one was prevalent in late August and September and was responsible for shifting the vortex off the pole in October. We could not estimate the amount of wave forcing in the lower stratosphere, but, based on other studies and Fig. 2.1, presumed it to be small. Changes in the thermal and wind structure in the lower stratosphere were then related to more significant wave forcing in the upper stratosphere through the downward control principle.

We concluded our discussion of the circumpolar vortex with observations of interannual variability in its circulation. According to current thought (summarized by Tung

and Yang (1994b)), the QBO in the tropics modulates the polar westerly waveguide in the stratosphere, leading to enhanced wave activity in the polar stratosphere during the easterly phase. During the westerly phase the barrier to equatorward propagation is reduced, allowing waves propagating from the troposphere into the stratosphere to follow their natural equatorward path. While we did observe some consistency between the phase of the QBO and the circulation of the circumpolar vortex, there were notable exceptions (e.g., 1989 and 1991). We concluded that the primary reason for the discrepancies may be due to the small wave forcing in the lower stratosphere. It is expected that the results would improve in the upper stratosphere where wave forcing is a maximum (although this signal may still be small compared to that observed in the Northern Hemisphere).

6.1 Suggested Future Work

According to Vroman and Stephens (1989), an advanced version of the MSU is due to be launched sometime in the 1990s. This vertical sounder would have 11 channels in the 50–60 GHz frequency range with sharper weighting functions and much better vertical resolution. With this increased vertical resolution we could then compute potential vorticity and perform much more detailed analyses of wave activity in the stratosphere.

Even without this advanced MSU, there are still some things we can do to improve our results. First, we could try to find better relations between the MSU brightness temperatures and the geopotential thicknesses. This would increase the accuracy of the dynamical fields and allow for a more quantitative study. Second, we could include a vortex-centered average for computing zonal-mean quantities. Current zonal averages are taken about the pole. We do have a spectral space rotation algorithm, based on work by Rochas et al. (1991), that works quite well for computing zonal averages of temperature about the vortex center. We did not, however, extend it to velocity and vorticity. If we could extend it to the dynamical fields, errors introduced into our results by taking (pole-centered) zonal averages of non-circumpolar vortices would be greatly reduced.

Much work must still be done to quantify the correlation between the QBO and the interannual variability of extratropical fields. Evidenced by the number of studies using

different levels and months of the QBO, it is not yet agreed exactly how this correlation should be performed. Additional difficulties arise in studies which attempt to demonstrate the presence of a QBO in the extratropical circulation through indirect means, like ozone transport. Some claim that extratropical QBO ozone anomalies may result from transport directly from the Tropics (e.g., Gray and Ruth, 1993), rather than in situ transport, as the dynamical mechanism summarized in Chapter 2 would suggest. To resolve this debate there must be, as Tung and Yang (1994b) point out, a direct measurement of the modulated mean meridional circulation. This would confirm the presence of a wave-induced extratropical circulation modulated by the QBO, and lend credence to the current indirect observations of such an anomaly. Current techniques, however, are incapable of such measurements. Once these issues are resolved, and a complete picture of interannual variability is constructed, we may one day be able to actually predict the oscillations in the strength of the transport circulation, the size of the circumpolar vortex, and the intensity of the "ozone hole."

Appendix A

FIGURES

The following are figures referred to in the climatological section of Chapter 5. The first figure is a time-height plot of equatorial winds used to determine the phase of the QBO. Although the 30-mb winds are used in the correlation of the QBO phase and circumpolar vortex circulation, we felt it imperative that the reader understand that different results will be obtained based on the level chosen. The wind anomaly amplitude appears to be the greatest near 30 mb, hence our reason for choosing it.

The second set of figures, A.2–A.9, show the MSU Channel 4 brightness temperatures for each month and year of our study. The thermal field gives some indication of the size of the vortex. The warmer the central temperatures compared to the climatological values, the weaker and more disturbed the vortex is. We can also compare the degree of shifting off the pole in September and October to infer the strength of wavenumber one forcing.

The final set of figures, A.10–A.17, show the absolute vorticity at 50 mb for each month and year of our study. These are primarily supplements to the temperature plots, but do contain some additional information. Because wave breaking in the lower stratosphere is very weak, we should only expect one of two features at the edge of the circumpolar vortex: If wave *propagation* is weak, the lines of absolute vorticity should be nearly zonal. If waves are present and are causing reversible undulations of the vortex edge, this should be present in the absolute vorticity field. Additionally, we may infer the strength of the vortex by the gradients between 50°S and 60°S. A more developed vortex will have tighter gradients of absolute vorticity in this region associated with strong flow curvature.

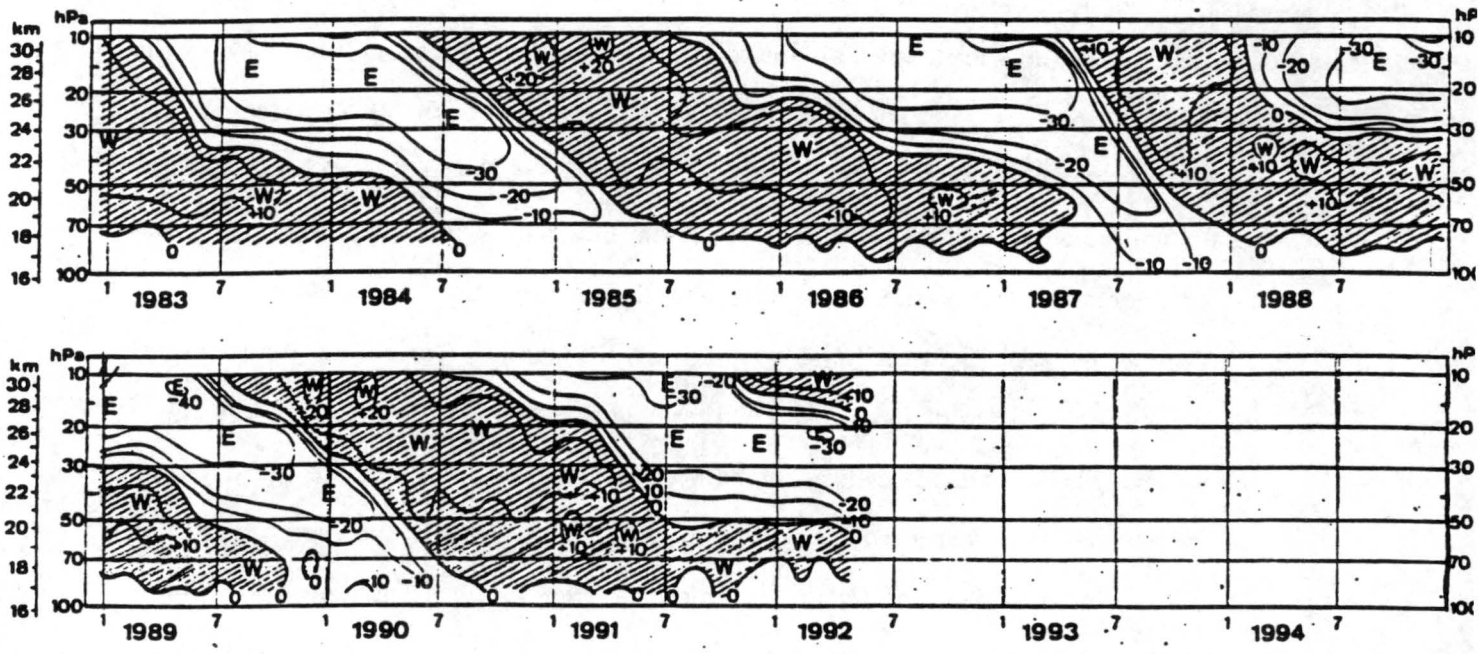


Figure A.1: Time-height plot of equatorial winds. Contour interval 10 ms^{-1}. (From James (1994)).

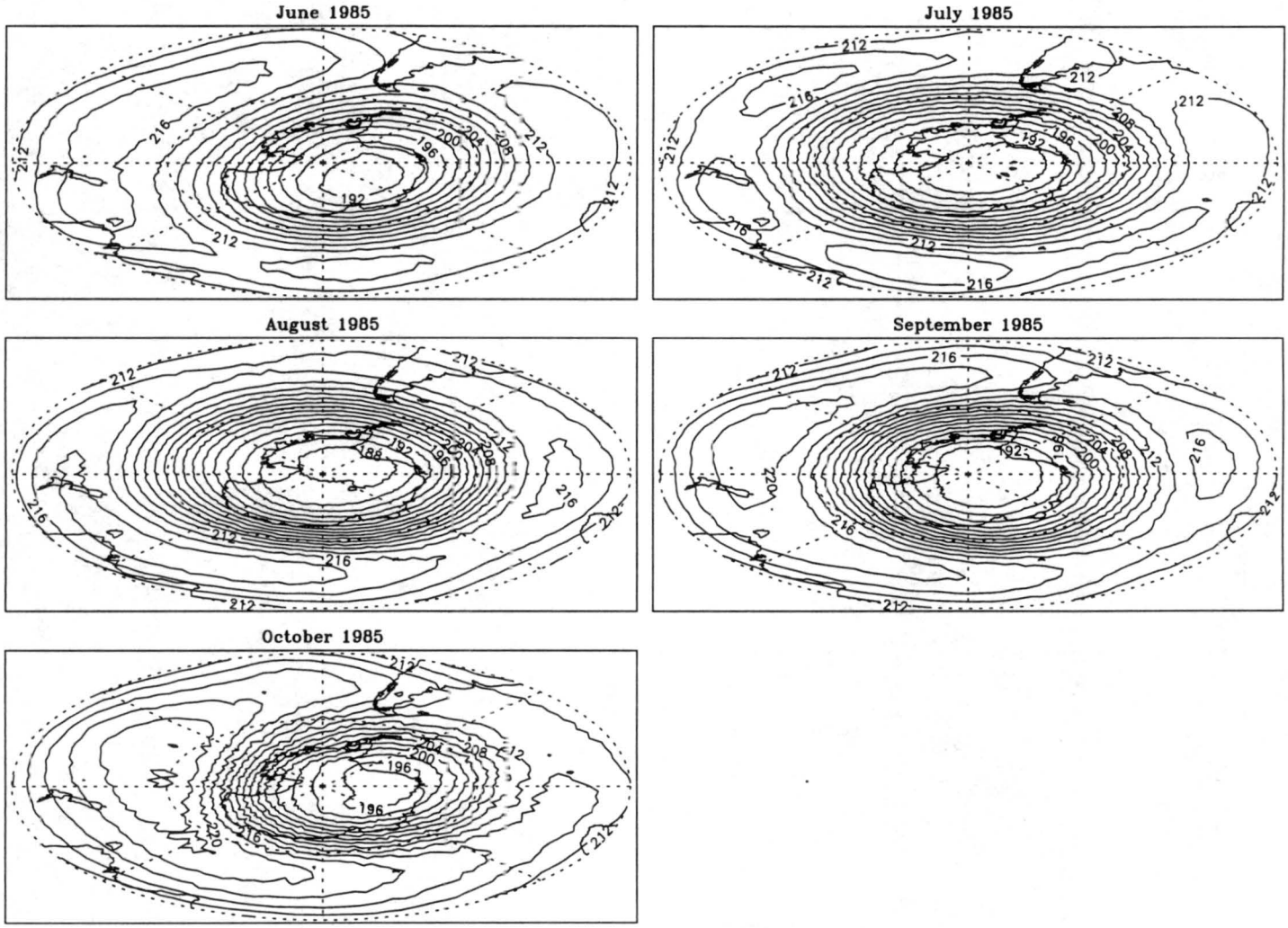


Figure A.2: MSU Channel 4 brightness temperature for 1985. Contour interval 2 K.

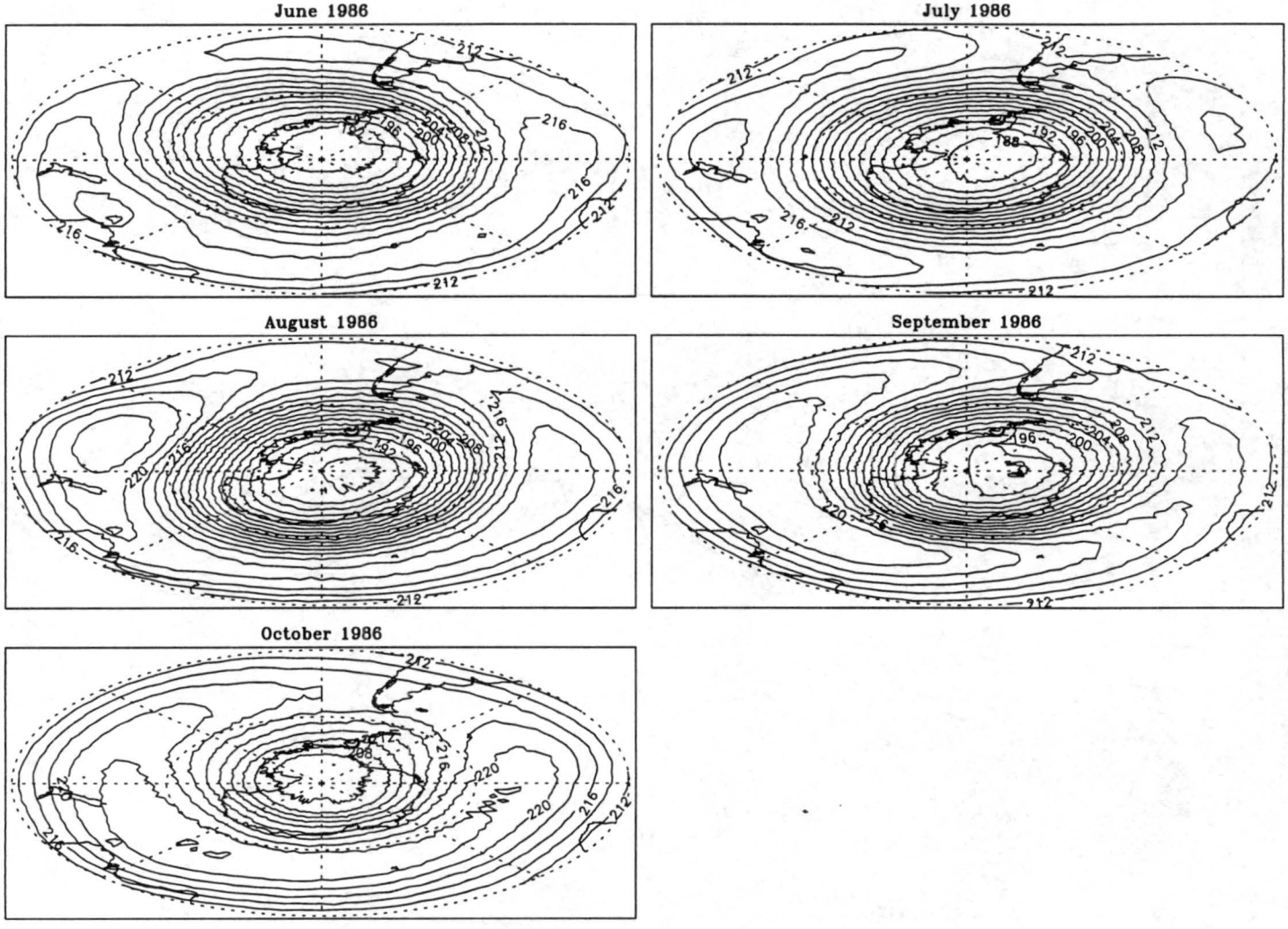


Figure A.3: MSU Channel 4 brightness temperature for 1986. Contour interval 2 K.

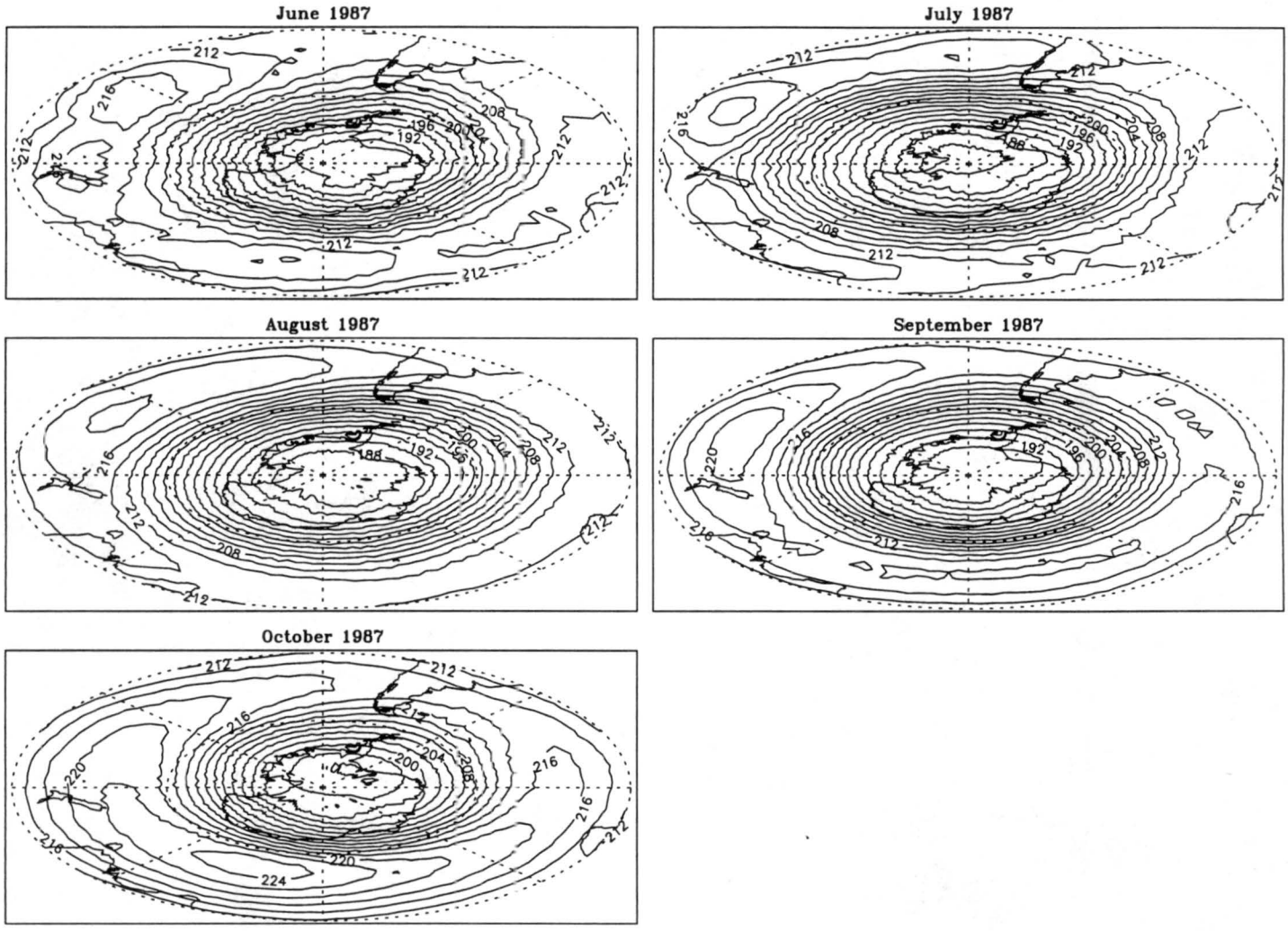


Figure A.4: MSU Channel 4 brightness temperature for 1987. Contour interval 2 K.

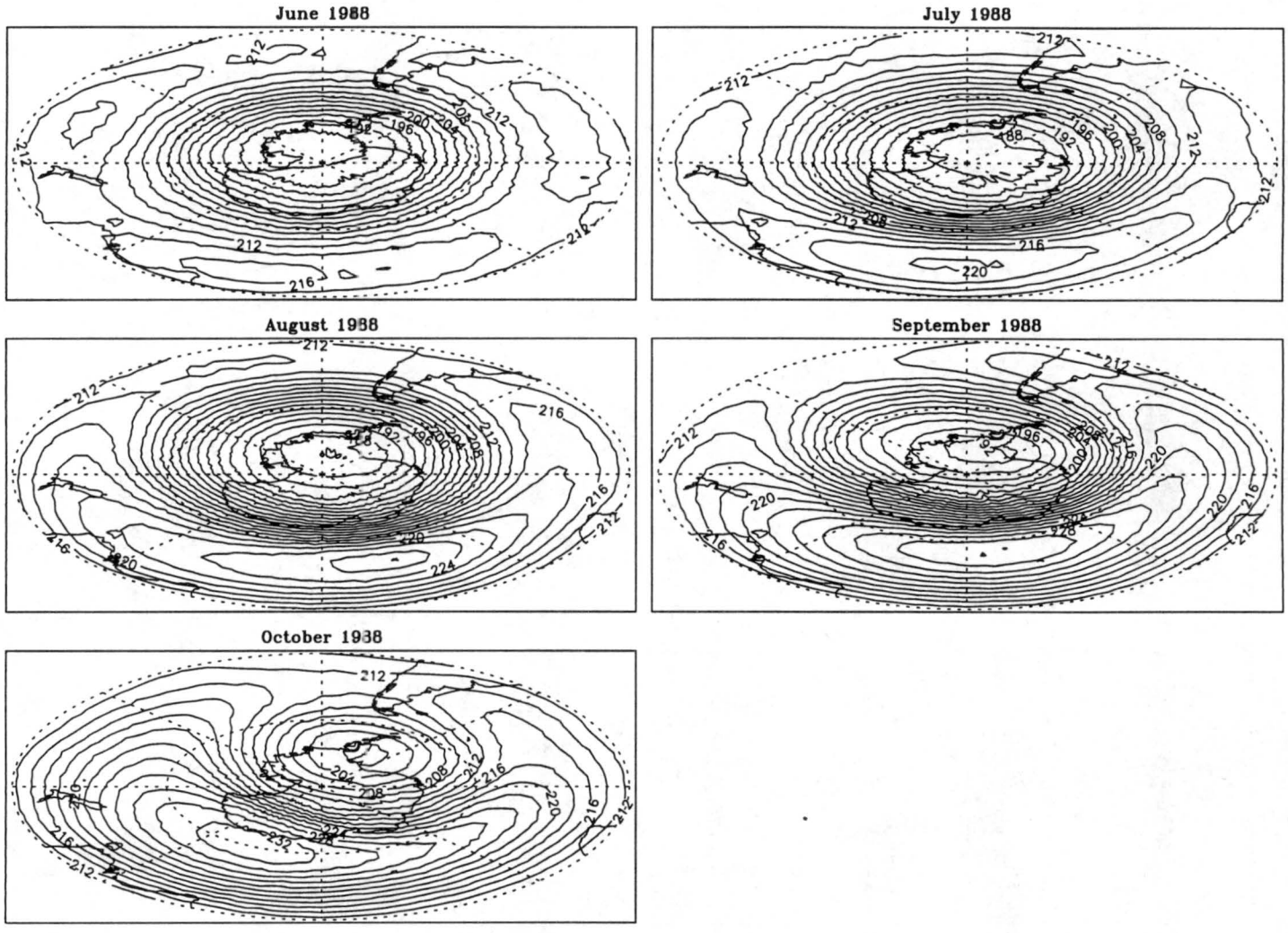


Figure A.5: MSU Channel 4 brightness temperature for 1988. Contour interval 2 K.

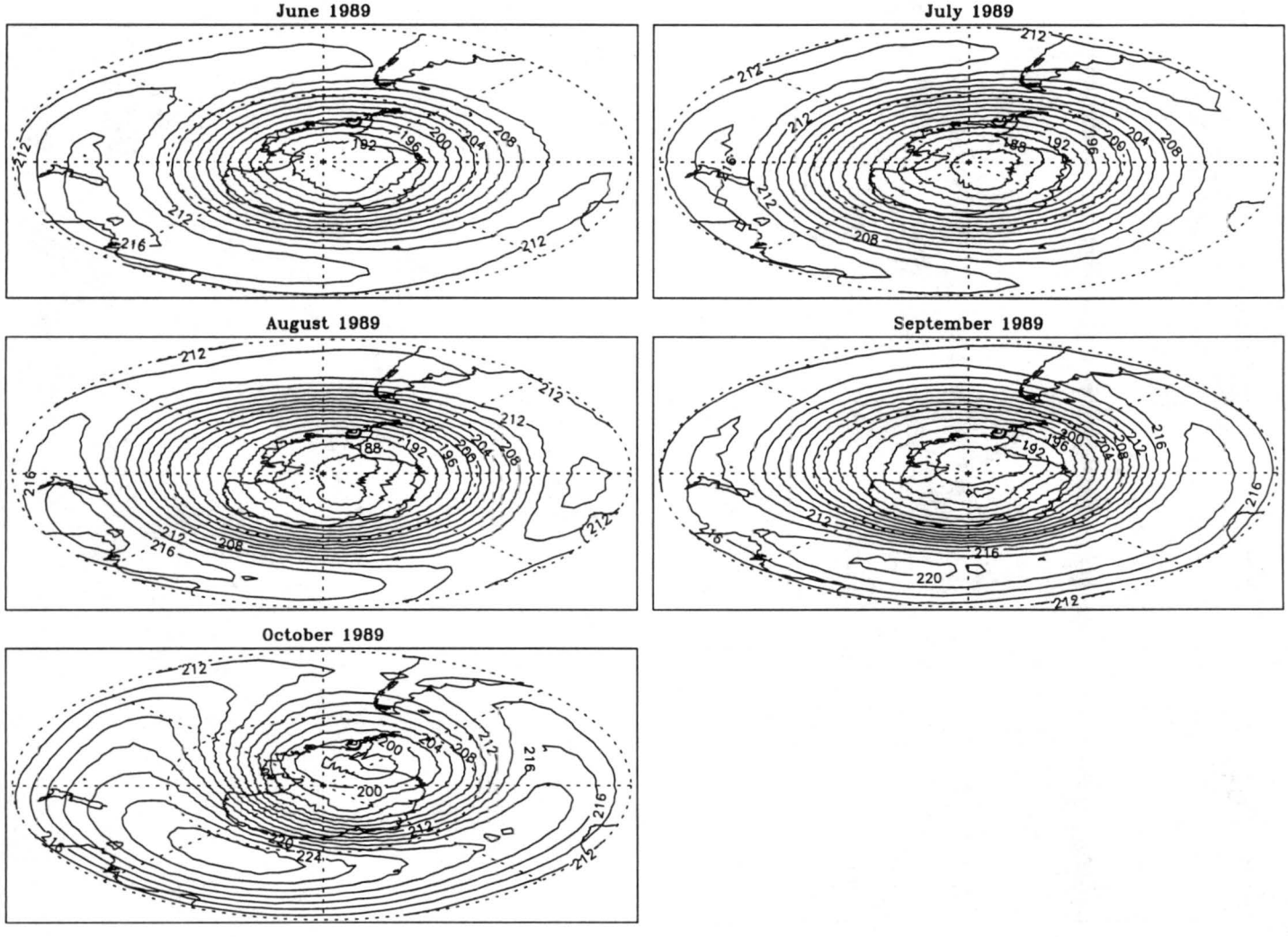


Figure A.6: MSU Channel 4 brightness temperature for 1989. Contour interval 2 K.

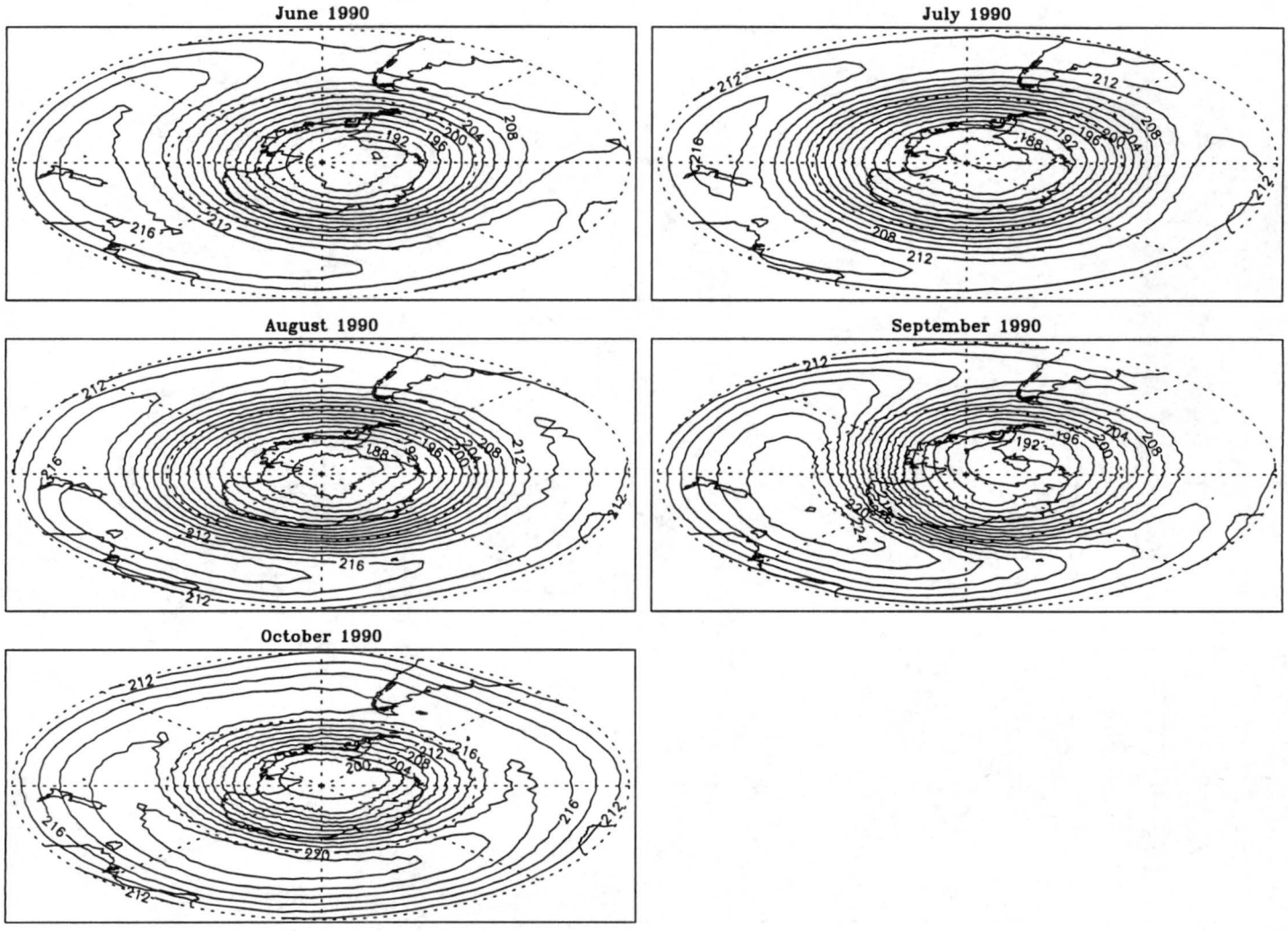


Figure A.7: MSU Channel 4 brightness temperature for 1990. Contour interval 2 K.

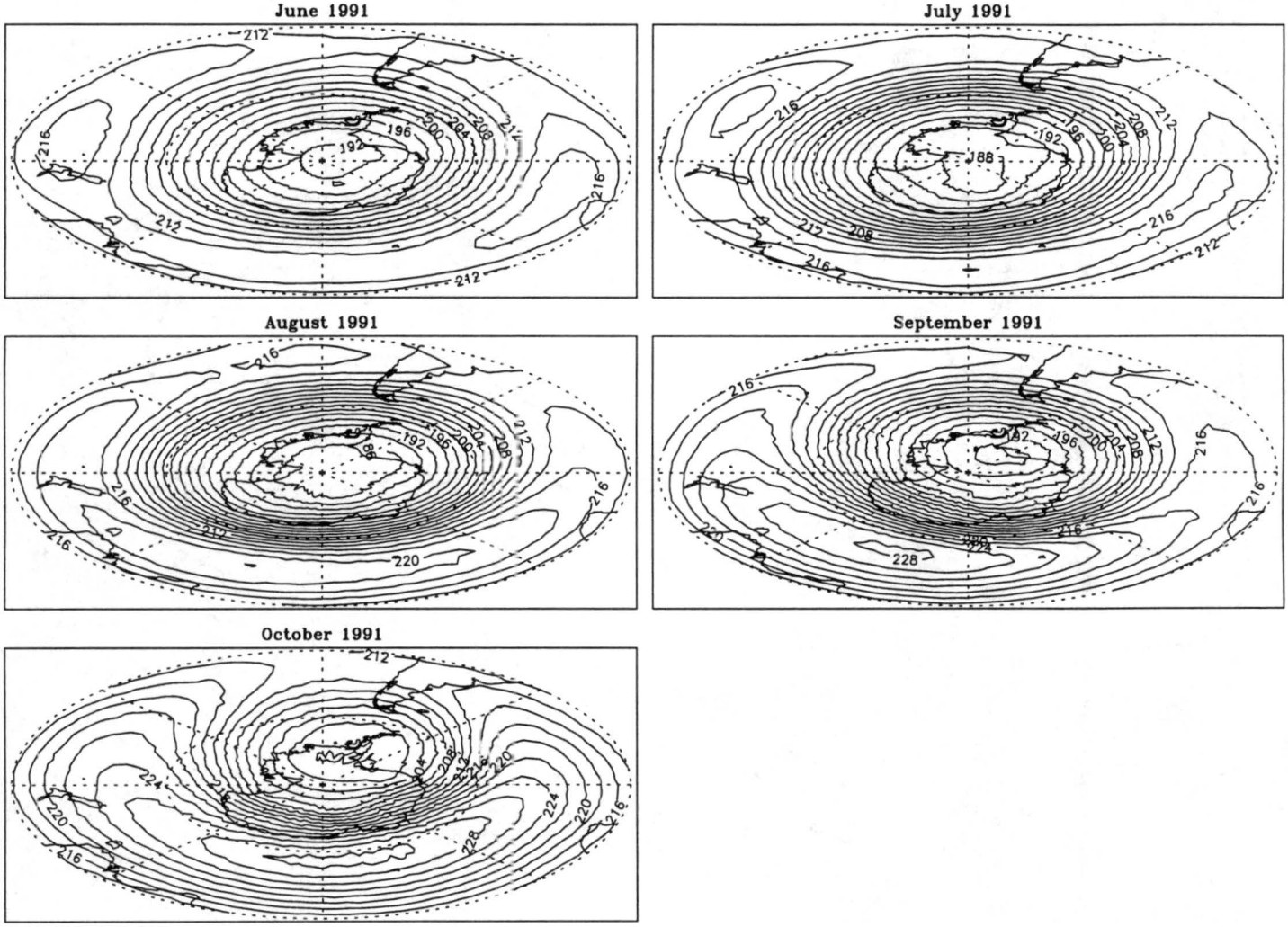


Figure A.8: MSU Channel 4 brightness temperature for 1991. Contour interval 2 K.

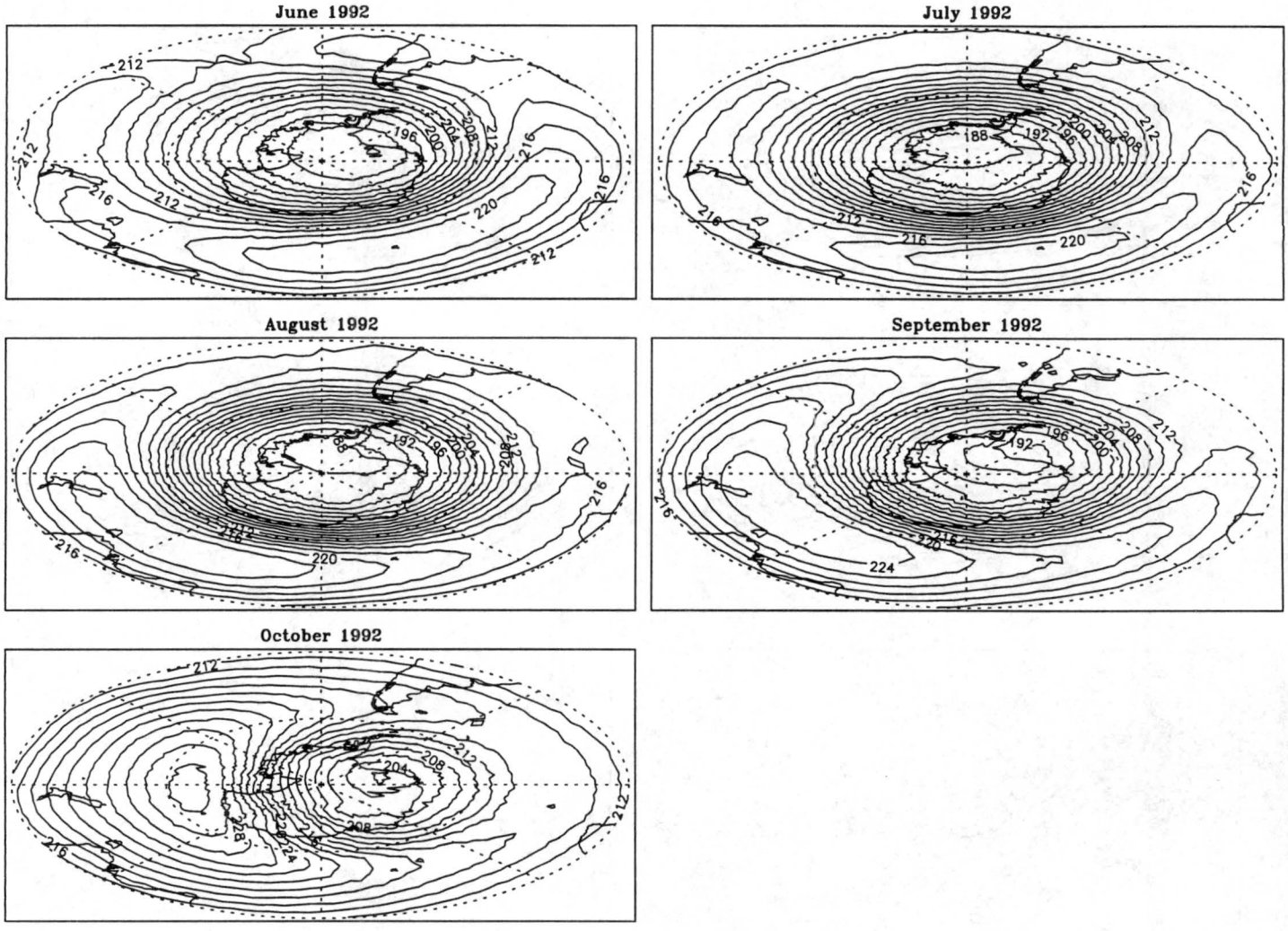


Figure A.9: MSU Channel 4 brightness temperature for 1992. Contour interval 2 K.

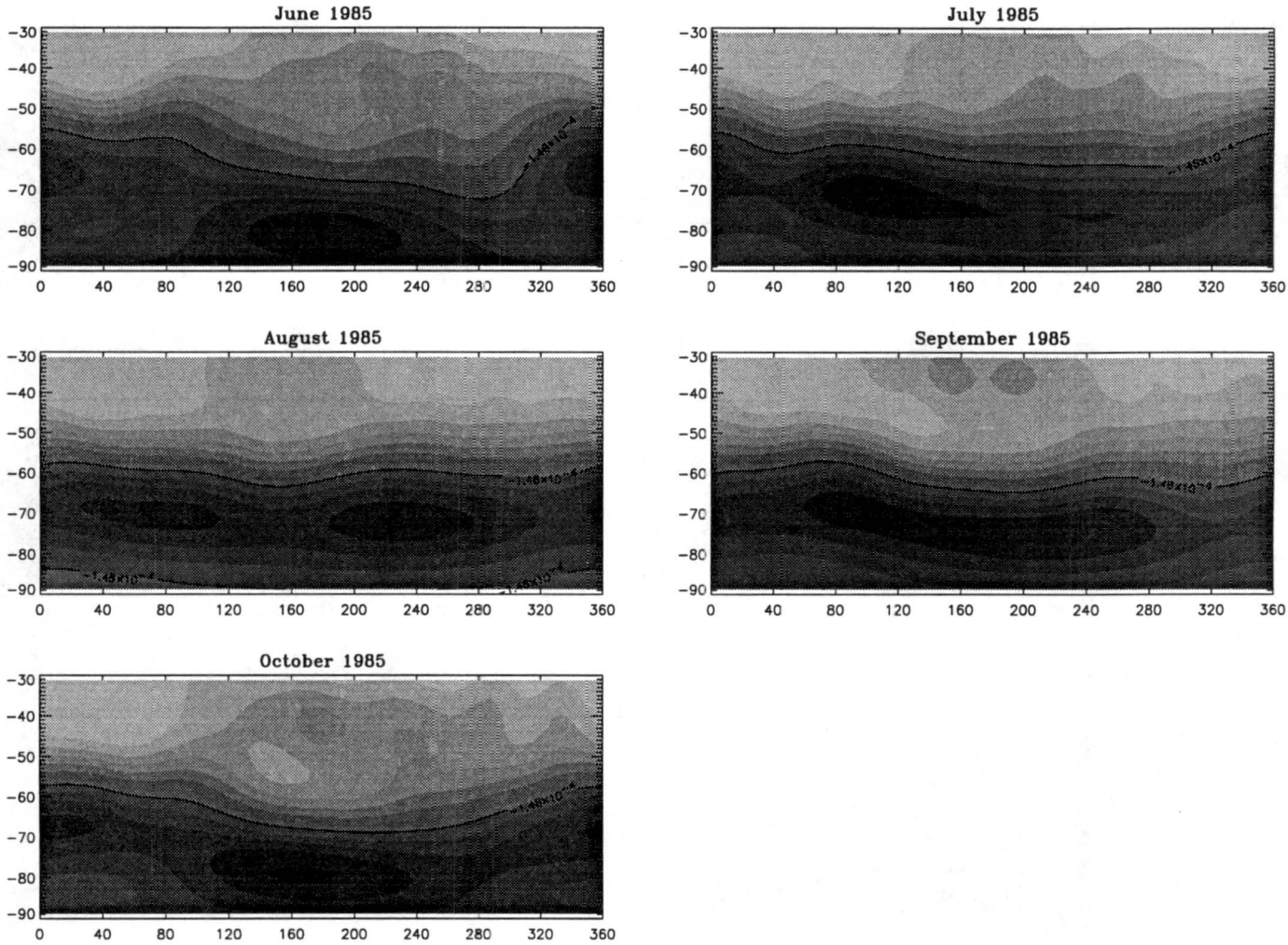


Figure A.10: 50-mb absolute vorticity for 1985. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

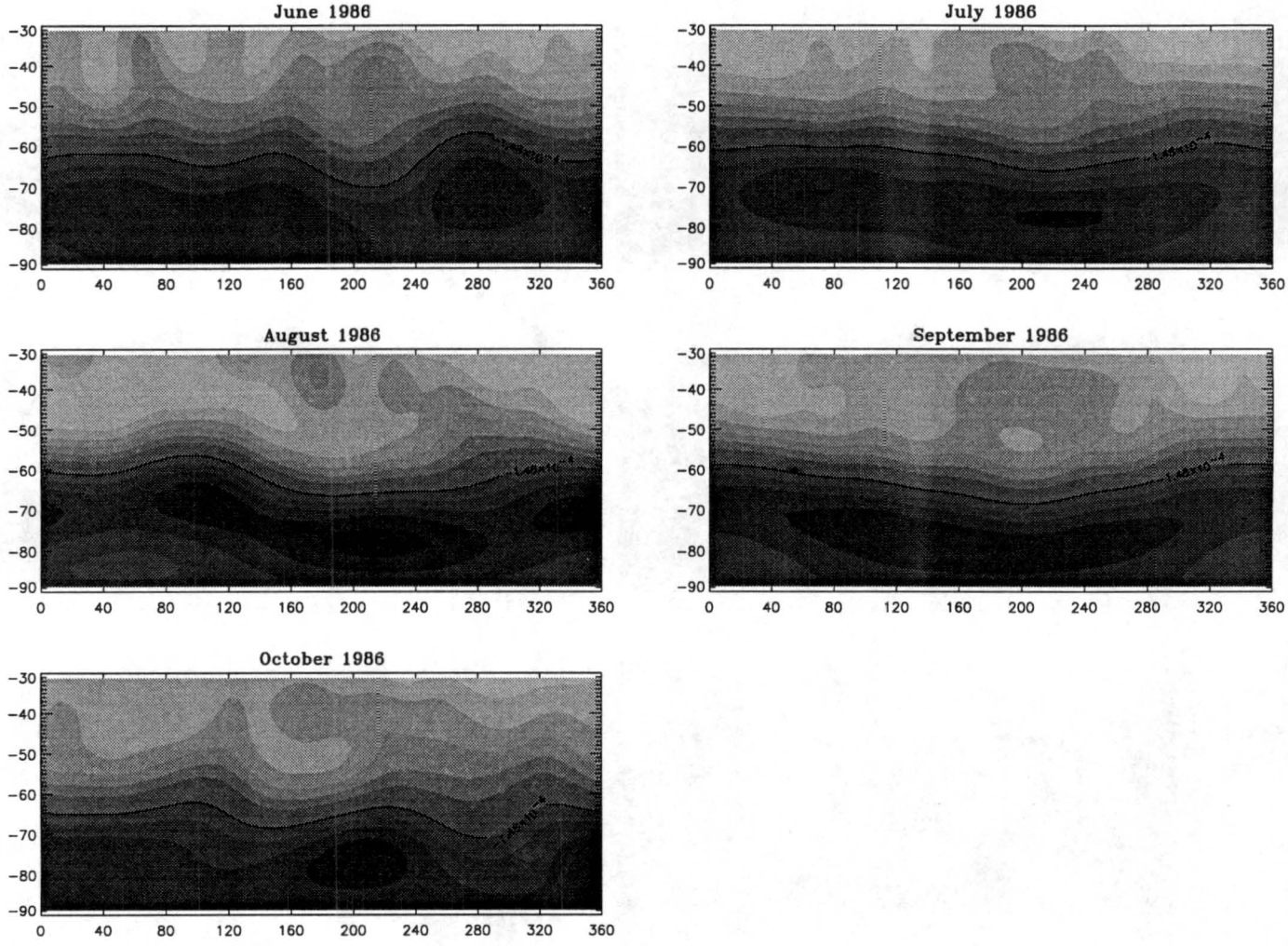


Figure A.11: 50-mb absolute vorticity for 1986. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

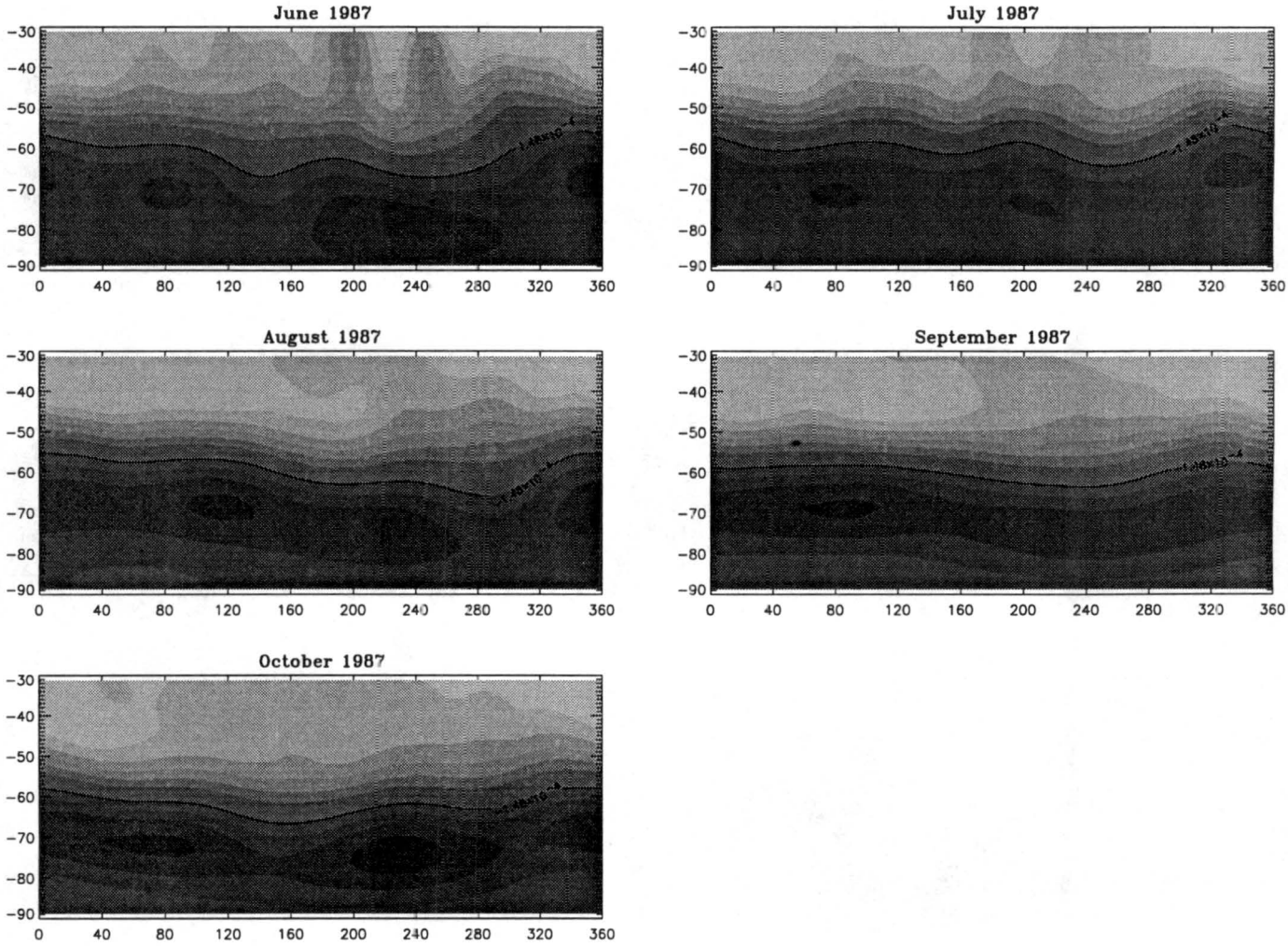


Figure A.12: 50-mb absolute vorticity for 1987. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

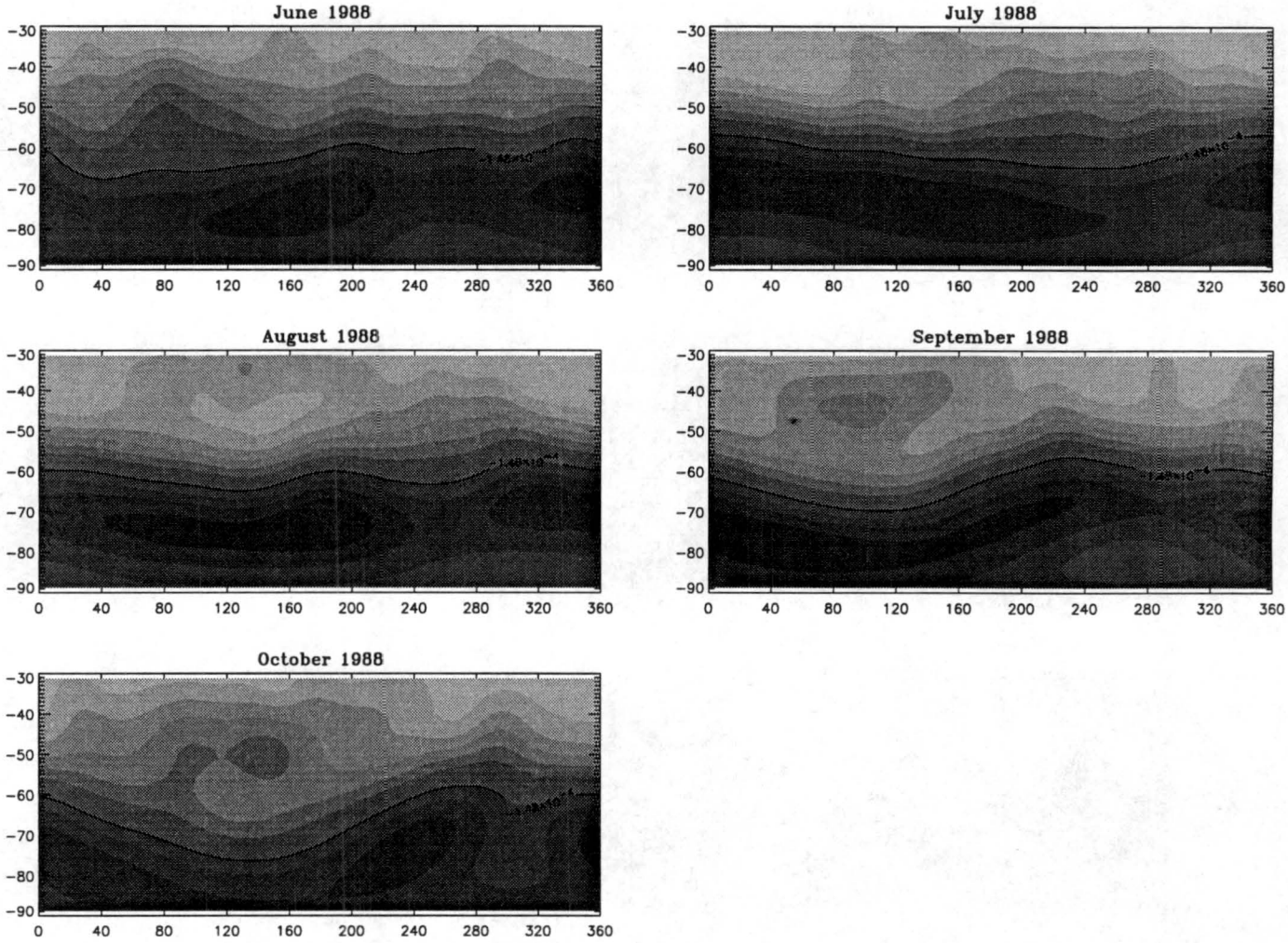


Figure A.13: 50-mb absolute vorticity for 1988. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

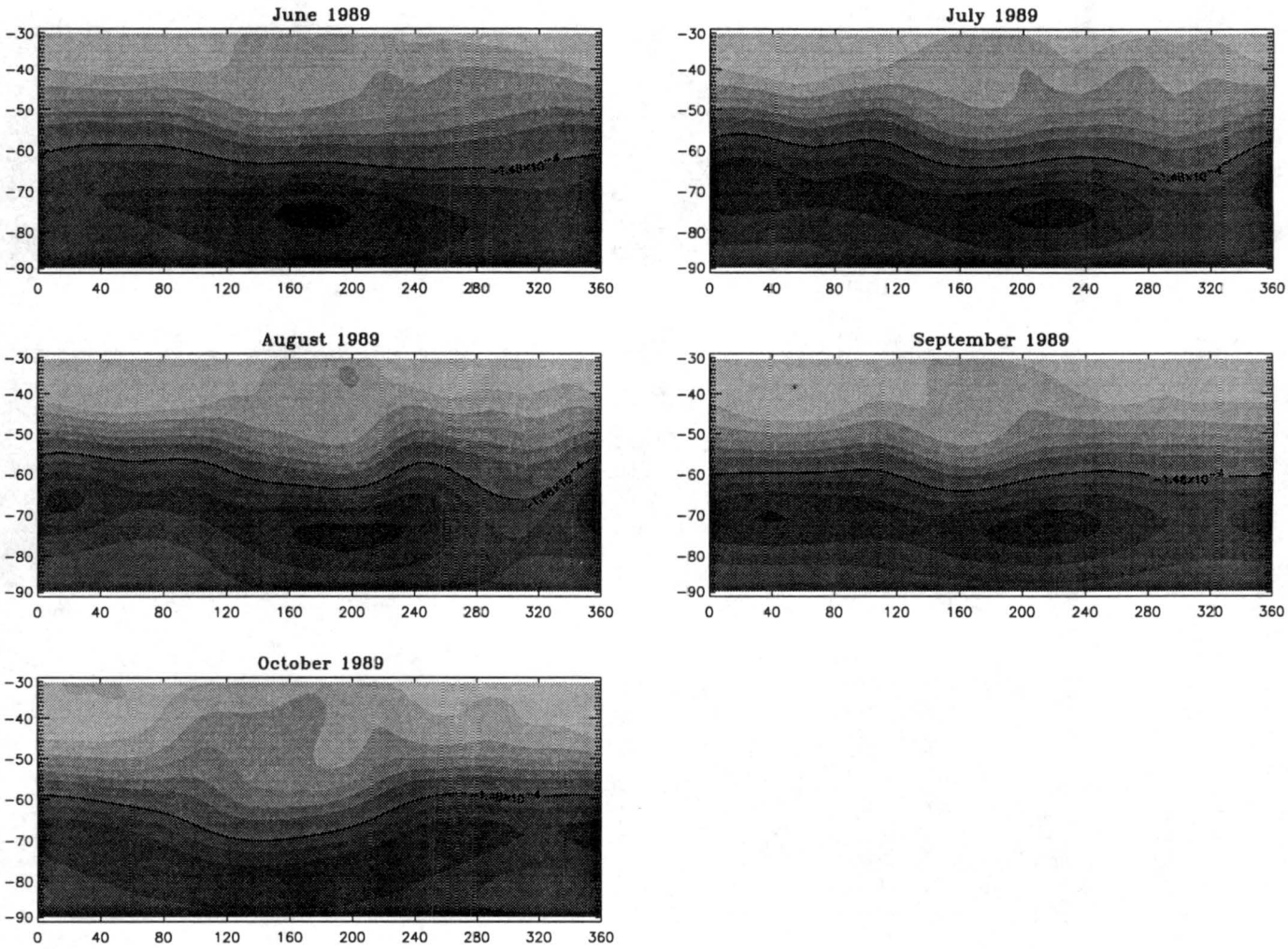


Figure A.14: 50-mb absolute vorticity for 1989. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

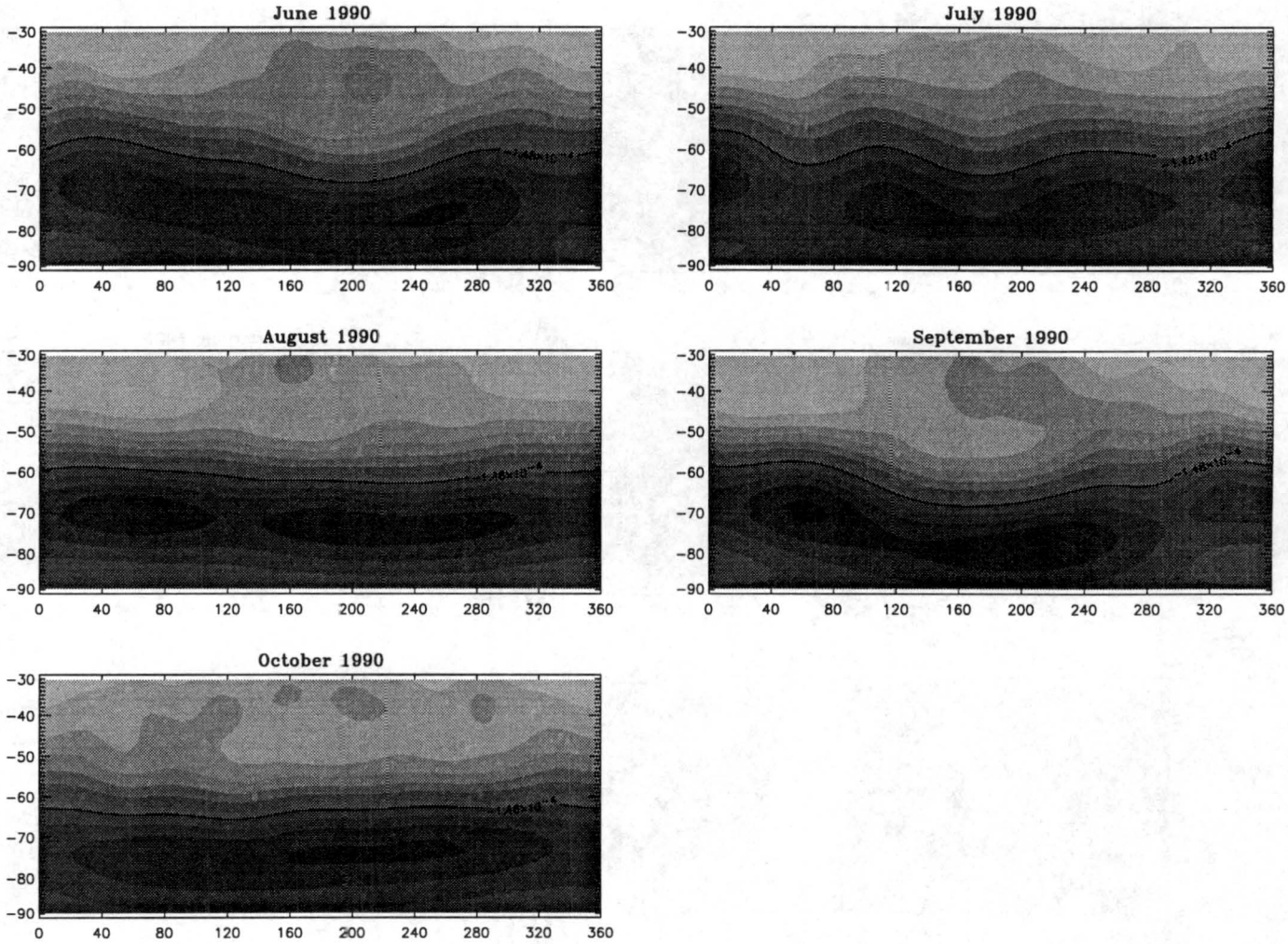


Figure A.15: 50-mb absolute vorticity for 1990. Contour interval $0.125 \times 10^{-4} s^{-1}$.

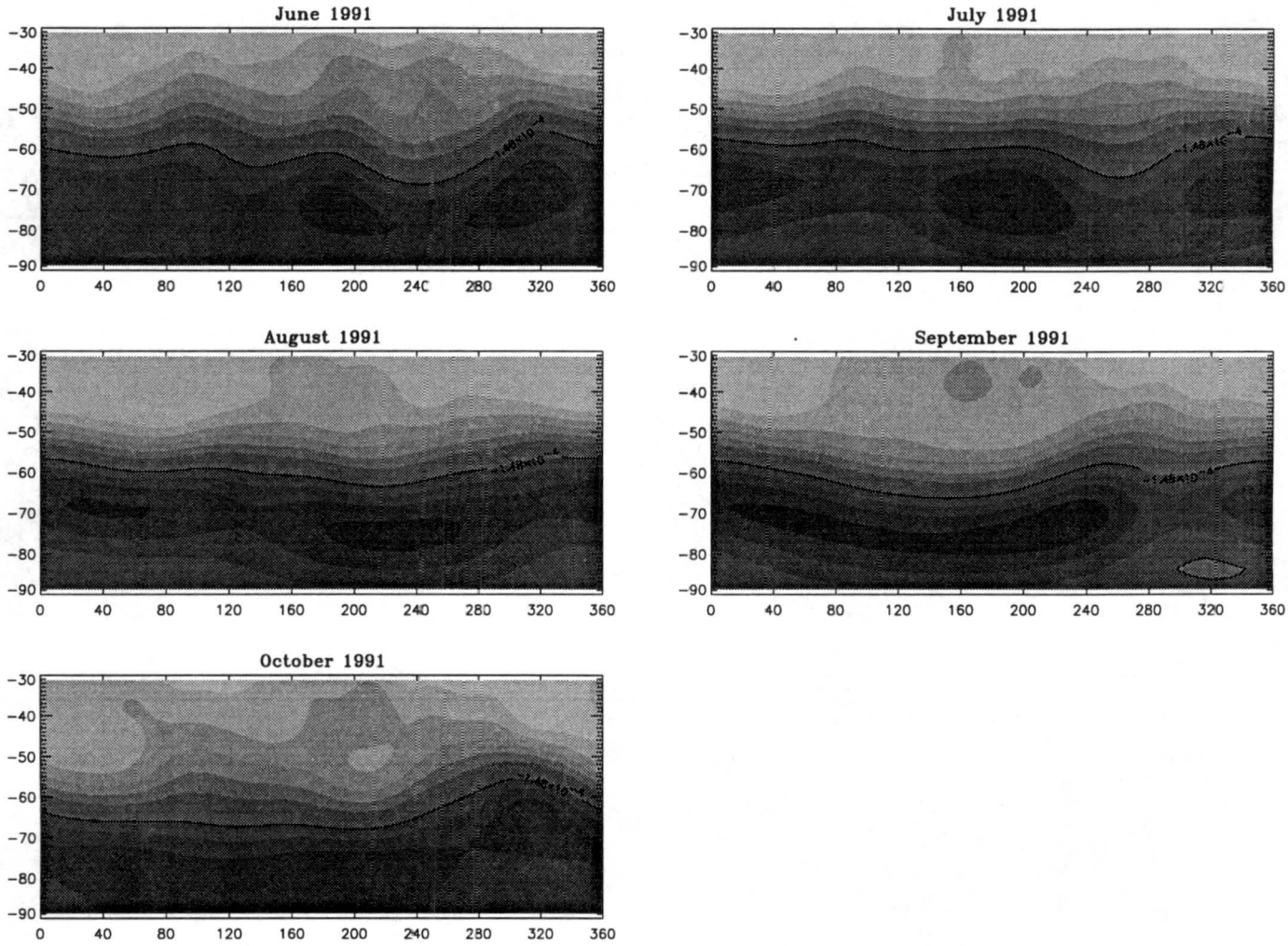


Figure A.16: 50-mb absolute vorticity for 1991. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

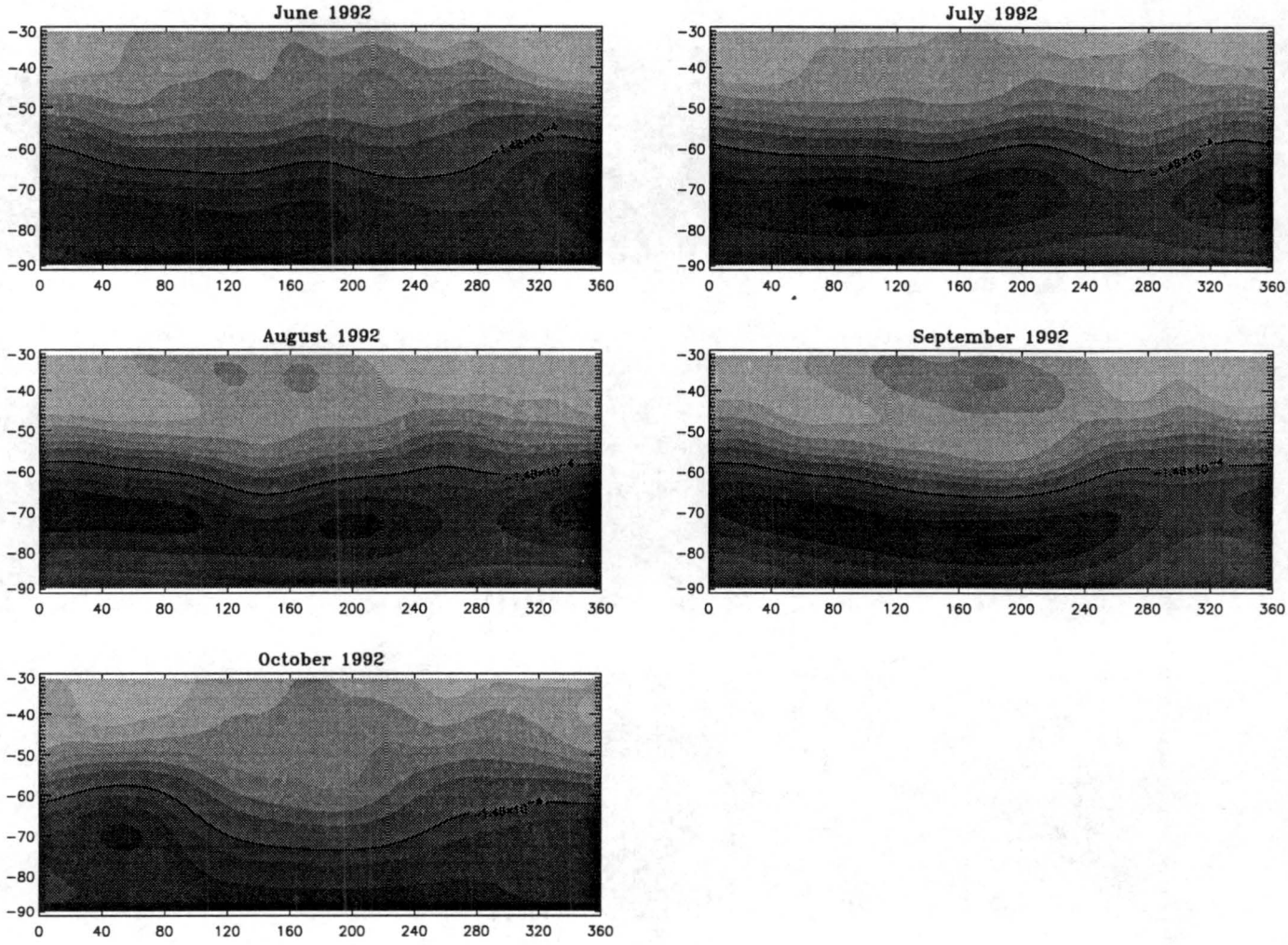


Figure A.17: 50-mb absolute vorticity for 1992. Contour interval $0.125 \times 10^{-4}\mathrm{s}^{-1}$.

Appendix B

THE SPECTRAL METHOD

Suppose there exists some function, $f(\mathbf{r},t)$ defined on a sphere of radius, a, that satisfies the wave equation,

$$\nabla^2 f = \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}. \tag{B.1}$$

Let $f(\mathbf{r},t)$ be separable into spatial and temporal parts, ψ and T, respectively. The differential equation governing the spatial part of the solution, called the Helmholtz equation, takes the form:

$$\nabla^2 \psi = \lambda \psi, \tag{B.2}$$

where λ is a separation constant.

Now let ψ be separable into its polar and azimuthal parts, Θ and Φ, respectively. Assuming periodic solutions in the azimuth, we obtain the following ordinary differential equation for the polar solution:

$$\frac{d}{d\mu}\left[\left(1-\mu^2\right)\frac{d\Theta}{d\mu}\right] + \left[-\lambda a^2 - \frac{m^2}{1-\mu^2}\right]\Theta = 0, \tag{B.3}$$

where m is the azimuthal wavenumber and μ is the sine of the latitude. This is just Legendre's Associated Differential equation. Requiring Θ to be finite at the singular points (i.e., the poles), it can be shown that the solution is of the form

$$\Theta(\mu) = \left(1-\mu^2\right)^{m/2} u(\mu). \tag{B.4}$$

Substituting this into (B.3) we obtain

$$\left(1-\mu^2\right)\frac{d^2u}{d\mu^2} - 2(m+1)\mu\frac{du}{d\mu} - \left(\lambda a^2 + m(m+1)\right)u = 0. \tag{B.5}$$

Solutions to (B.5) may be sought viz. a regular power series

$$u(\mu) = \sum_{r=0}^{r=\infty} \beta_r \mu^r, \tag{B.6}$$

where the β_r are found by a recursion relationship. Away from the poles, the ratio test tells us that the series converges absolutely. At the poles this test is inconclusive, so we must turn to the integral test (e.g., Protter and Protter, 1988). The integral test shows that the series will diverge at the poles. To force convergence, we need a way of truncating the series. This can be accomplished if we let $\lambda_n = -n(n+1)/a^2$, where n is a nonnegative integer. Another way of computing $u(\mu)$ is by taking the mth derivative of the nth order Legendre Polynomial. Hence, the solutions to (B.3) are often called the Associated Legendre Functions, $P_n^m(\mu)$. The index, n, is referred to as the total wavenumber or nodal number.

The Helmholtz equation (B.2) is clearly an eigenvalue-eigenfunction problem. The eigenvalues, λ_n, are as defined above. The eigenfunctions for a particular m and n are obtained by multiplying P_n^m and the periodic azimuthal solution:

$$Y_n^m(\phi, \mu) = P_n^m(\mu) e^{im\phi}. \tag{B.7}$$

These functions are called the Spherical Harmonics, and have two very important properties. The first is that they are orthogonal. The second is that they form a complete set. Thus, any function, $g(\phi, \mu)$, defined on the sphere (say, the Earth) can be decomposed into a series of Spherical Harmonics, since they form a basis for the spherical geometry. We can then write

$$g(\phi, \mu) = \sum_{n=0}^{n=\infty} \sum_{m=-n}^{m=n} \hat{g}_n^m P_n^m(\mu) e^{im\phi}, \tag{B.8}$$

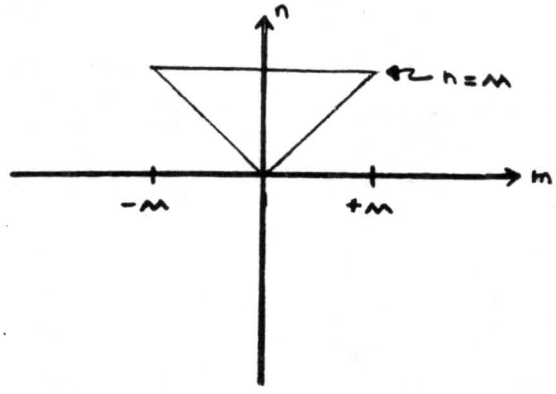
where

$$\hat{g}_n^m = \frac{1}{2} \int_{-1}^{1} \hat{g}^m(\mu) P_n^m(\mu) d\mu. \tag{B.9}$$

The function, $\hat{g}^m(\mu)$, is the Fourier transform of g, and is given by

$$\hat{g}^m(\mu) = \frac{1}{2\pi} \int_0^{2\pi} g(\phi, \mu) e^{im\phi} d\phi. \tag{B.10}$$

The question that now arises is how to perform the integrals in (B.9) and (B.10) numerically. Before we address this question, we must first decide how to truncate the series in (B.8). We choose to use a triangular truncation:



Thus, only azimuthal modes, $m = (-M, M)$, are kept. For our study we use $M = 60$.

If we now let the number of points around a latitude circle equal $2M + 1$, it can be shown that

$$\hat{g}^m(\mu) = \frac{1}{2M+1} \sum_{j=1}^{j=2M+1} g(\phi_j, \mu) e^{-im\phi_j}, \tag{B.11}$$

where $\phi_j = 2\pi j/(2M+1)$. We use a FFT to perform the above summation.

Once $\hat{g}^m(\mu)$ has been calculated (at the appropriate latitudes), we can then compute $\hat{g}_n^m$ using the quadrature formula:

$$\hat{g}_n^m = \frac{1}{2}\int_{-1}^{1} \hat{g}^m(\mu) P_n^m(\mu) d\mu = \sum_{k=1}^{k=M+2} G_k \hat{g}^m(\mu_k) P_n^m(\mu_k). \tag{B.12}$$

The μ_k are called the Gaussian latitudes, and are defined as the zeros of the (M+2)th Legendre Polynomial. The G_k are the Gaussian weights, and are given by

$$G_k = \frac{(1-\mu_k^2)(\ell - 1/2)}{[\ell P_{l-1}(\mu_k)]^2}, \qquad \ell = M+2. \tag{B.13}$$

The utility of the spectral method is easily demonstrated through a practical application. Given the temperature field, the relative vorticity can be computed via Equation (4.6). The Laplacian of the temperature can certainly be evaluated using a finite difference scheme for the derivatives. If, however, the temperature is expanded in a series of Spherical Harmonics, the Laplacian can be evaluated using (B.2). The equation for vorticity then reduces to a simple algebraic relation for the spectral coefficients:

$$\zeta_n^m(p_1) = \zeta_n^m(p_0) + \frac{n(n+1)R_d}{a^2 f} \ln\left(\frac{p_1}{p_0}\right) \bar{T}_n^m. \tag{B.14}$$

Once the vorticity spectral coefficients have been computed, then (B.8) is used to compute the vorticity in physical space. When performing the summation, however, not all the coefficients need be used. For example, if one is only concerned with the zonally symmetric wavenumber component of a field (i.e., $m = 0$), then all coefficients for which m is nonzero can be set to zero. The sum in (B.8) then produces the zonal average of g. This, of course, may also be extended to other wavenumbers, allowing for a convenient means of spectrally analyzing a field.

REFERENCES

Andrews, D. G., and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031–2048.

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

Angell, J. K., and J. Korshover, 1963: Harmonic analysis of the biennial zonal-wind and temperature regimes. *Mon. Wea. Rev.*, **91**, 537–548.

Atkinson, R. J., 1993: An observational study of the austral spring stratosphere: Dynamics, ozone transport and the 'ozone dilution effect.' *Massachusetts Institute of Technology*, Ph.D. thesis.

Bailey, M. J., A. O'Neill, and V. D. Pope, 1993: Stratospheric analyses produced by the United Kingdom Meteorological Office. *J. Appl. Meteor.*, **32**, 1472–1483.

Butchart, N., and J. Austin, 1996: On the relationship between the quasi-biennial oscillation, total chlorine and the severity of the antarctic ozone hole. *Quart. J. Roy. Meteor. Soc.*, **122**, 183–217.

Charney, J. G., and P. G. Drazin, 1961: Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83–109.

Clough, S. A., N. S. Grahame, and A. O'Neill, 1985: Potential vorticity in the stratosphere derived using data from satellites. *Quart. J. Roy. Meteor. Soc.*, **111**, 335–358.

Dunkerton, T. J., and M. Baldwin, 1991: Quasi-biennial modulation of planetary-wave fluxes in the Northern Hemisphere winter. *J. Atmos. Sci.*, **48**, 1043–1061.

Egger, J., 1996: Comments on "On the 'downward control' of extratropical diabatic circulations by eddy-induced mean zonal forces." *J. Atmos. Sci.*, **53**, 2103–2104.

Feynman, R. P., 1989: The Feynman Lectures on Physics (Vol. I, Sect. 41-2). Addison-Wesley.

Garcia, R. R., and S. Solomon, 1987: A possible relationship between interannual variability in Antarctic ozone and the quasi-biennial oscillation. *Geophys. Res. Lett.*, **14**, 848–851.

Garcia, R. R., 1991: Parameterization of planetary wave breaking in the middle atmosphere. *J. Atmos. Sci.*, **48**, 1405–1419.

Gray, L. J., and S. Ruth, 1993: The modeled latitudinal distribution of the ozone quasi-biennial oscillation using observed winds. *J. Atmos. Sci.*, **50**, 1033–1046.

Grody, N. C., 1983: Severe storm observations using the microwave sounding unit. *J. Climate Appl. Meteor.*, **22**, 609–625.

Hartmann, D. L., 1976: The structure of the stratosphere in the Southern Hemisphere during late winter 1973 as observed by satellite. *J. Atmos. Sci.*, **33**, 1141–1154.

Haynes, P. H., 1989: The effect of barotropic instability on the nonlinear evolution of a Rossby wave critical layer. *J. Fluid Mech.*, **207**, 231–266.

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.

Haynes, P. H., M. E. McIntyre, and T. G. Shepherd, 1996: Reply to "Comments on 'On the 'downward control' of extratropical diabatic circulations by eddy-induced mean zonal forces'." *J. Atmos. Sci.*, **53**, 2105–2107.

Holton, J. R., and H.-C. Tan, 1980: The influence of the quasi-biennial oscillation on the global circulation at 50 mb. *J. Atmos. Sci.*, **37**, 2200–2208.

Holton, J. R., and J. Austin, 1991: The influence of the equatorial QBO on sudden stratospheric warmings. *J. Atmos. Sci.*, **48**, 607–618.

Holton, J. R., 1992: Dynamic Meteorology. Academic Press. 511 pp.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.*, **33**, 403–439.

James, I. N., 1988: On the forcing of planetary-scale Rossby waves by Antarctica. *Quart. J. Roy. Meteor. Soc.*, **114**, 619–637.

James, I. N., 1994: Introduction to Circulating Atmospheres. Cambridge University Press. 422 pp.

Kiehl, J. T., and S. Solomon, 1986: On the radiative balance of the stratosphere. *J. Atmos. Sci.*, **43**, 1525–1534.

Killworth, P. D., and M. E. McIntyre, 1985: Do Rossby-wave critical layers absorb, reflect, or over-reflect. *J. Fluid Mech.*, **161**, 449-492.

Labitzke, K., 1982: On the interannual variability of the middle stratosphere during northern winters. *J. Meteor. Soc. Japan*, **60**, 124–139.

Lait, L. R., M. R. Schoeberl, and P. A. Newman, 1989: Quasi-biennial modulation of the Antarctic ozone depletion. *J. Geophys. Res.*, **94**, 11559–11571.

Lindzen, R., and R. Goody, 1965: Radiative and photochemical processes in mesospheric dynamics: Part I, models for radiative and photochemical processes. *J. Atmos. Sci.*, **22**, 341–348.

Liou, K.-N., 1980: An Introduction to Atmospheric Radiation. Academic Press. 392 pp.

Manabe, S., and F. Möller, 1961: On the radiative equilibrium and heat balance of the atmosphere. *Mon. Wea. Rev.*, **89**, 503–532.

Manney, G. L., J. D. Farrara, and C. R. Mechoso, 1991: The behavior of wave 2 in the Southern Hemisphere stratosphere during late winter and early spring. *J. Atmos. Sci.*, **48**, 976–998.

McIntyre, M. E., 1982: How well do we understand the dynamics of stratospheric warmings? *J. Meteorol. Soc. Japan*, **60**, 37–65.

McIntyre, M. E., and T. N. Palmer, 1983: The "surf zone" in the stratosphere. *J. Atm. Terr. Phys.*, **46**, 825–849.

McIntyre, M. E., and T. N. Palmer, 1985: A note on the general concept of wave breaking for Rossby and gravity waves. *Pageoph*, **123**, 964–975.

McIntyre, M. E., 1989: On the Antarctic ozone hole. *J. Atmos. Terr. Phys.*, **51**, 29–43.

McIntyre, M. E., 1992: The Use of EOS for Studies of Atmospheric Physics. North-Holland, 313–386.

Mechoso, C. R., and D. L. Hartmann, 1982: An observational study of travelling planetary waves in the Southern Hemisphere. *J. Atmos. Sci.*, **39**, 1921–1935.

Mechoso, C. R., D. L. Hartmann, and J. D. Farrara, 1985: Climatology and interannual variability of wave, mean-flow interaction in the Southern Hemisphere. *J. Atmos. Sci.*, **42**, 2189–2206.

Newman, P. A., M. R. Schoeberl, and L. R. Lait, 1990: Dynamics, Transport and Photochemistry in the Middle Atmosphere of the Southern Hemisphere. Kluwer Academic Publishers, 71-89.

O'Sullivan, D., and M. L. Salby, 1990: Coupling of the quasi-biennial oscillation and the extratropical circulation in the stratosphere through planetary wave transport. *J. Atmos. Sci.*, **47**, 650–673.

O'Sullivan, D., and R. E. Young, 1992: Modeling the quasi-biennial oscillation's effect on the winter stratospheric circulation. *J. Atmos. Sci.*, **49**, 2437–2448.

Pawson, S., and T. Kubitz, 1996: Climatology of planetary waves in the northern stratosphere. *J. Geophys. Res.*, **101**, 16987–16996.

Protter and Protter, 1988: Calculus with Analytic Geometry. Fourth Edition. Jones and Bartlett Publishers, Boston.

Quintanar, A. I., and C. R. Mechoso, 1995: Quasi-stationary waves in the Southern Hemisphere. Part I: Observational data. *J. Climate*, **8**, 2659–2672.

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

Rochas, M. J., Ph. Courtier, K. Yessad, and Y. Bouteloup, 1991: Rotations in spectral space. *Beitr. Phys. Atmosph.*, **64**, 201–206.

Rodgers, C. D., and C. D. Walshaw, 1966: The computation of infrared cooling rate in planetary atmcspheres. *Quart. J. Roy. Meteor. Soc.*, **92**, 67–92.

Rosenlof, K. H., and J. R. Holton, 1995: Estimates of the stratospheric residual circulation using the downward control principle. *J. Geophys. Res.*, **98**, 10465–10479.

Rosenlof, K. H., 1995: Seasonal cycle of the residual mean meridional circulation in the stratosphere. *J. Geophys. Res.*, **100**, 5173–5191.

Schoeberl, M. R., L. R. Lait, P. A. Newman, and J. E. Rosenfield, 1992: The structure of the polar vortex. *J. Geophys. Res.*, **97**, 7859–7882.

Shiotani, M., and I. Hirota, 1985: Planetary wave-mean flow interaction in the stratosphere: a comparison between northern and southern hemispheres. *Quart. J. Roy. Meteor. Soc.*, **111**, 309–334.

Shiotani, M., K. Kuroi, and I. Hirota, 1990: Eastward travelling waves in the southern hemisphere stratosphere during spring 1983. *Quart. J. Roy. Meteor. Soc.*, **116**, 913–927.

Spencer, R. W., and J. R. Christy, 1993: Precision lower stratospheric temperature monitoring with the MSU: Technique, validation, and results 1979-1991. *J. Climate*, **6**, 1194–1204.

Stewartson, K., 1978: The evolution of the critical layer of a Rossby wave. *Geophys. Astrophys. Fluid Dyn* , **9**, 185–200.

Trenberth, K. E., and J. G. Olson, 1989: Temperature trends at the South Pole and McMurdo Sound. *J. Climate*, **2**, 1196–1206.

Tung, K. K., and H. Yang, 1994a: Global QBO in circulation and ozone. Part I: Reexamination of observational evidence. *J. Atmos. Sci.*, **51**, 2699–2707.

Tung, K. K., and H. Yang, 1994b: Global QBO in circulation and ozone. Part II: A simple mechanistic model. *J. Atmos. Sci.*, **51**, 2708–2721.

Vroman, T. T., and G. L. Stephens, 1989: Microwave brightness temperature and its relation to atmospheric general circulation features. *Colorado State University*, Atmospheric Science Paper No. 454.

Warn, T., and H. Warn, 1978: The evolution of a nonlinear critical level. *Stud. Appl. Math.*, **59**, 37–71.

Wehrbein, W. M., and C. Leovy, 1982: An accurate radiative heating and cooling algorithm for use in a dynamical model of the middle atmosphere. *J. Atmos. Sci.*, **39**, 1532–1544.

Yulaeva, E., J. R. Holton, and J. M. Wallace, 1994: On the cause of the annual cycle in tropical lower-stratospheric temperatures. *J. Atmos. Sci.*, **51**, 169–174.