An Isentropic Model and the Effect of Stratospheric Planetary Waves

Jonathan Stewart Kinnersley

PhD thesis

1991

Department of Meteorology,
University of Edinburgh

DECLARATION

I hereby declare that the work described in this thesis is my own, unless otherwise stated, and that the thesis has been composed by myself.

ACKNOWLEDGEMENTS

I acknowledge with thanks the help of Dr. R.S. Harwood, my supervisor, whose solid advice and encouragement added greatly to the quality of my work, and my enjoyment of it.

I am grateful to my fellow students for their camaraderie through shared misfortune and triumph, and to everyone in the department of Meteorology here, for their friendship, and just for being themselves (and also for giving me a desk in front of the window).

I also owe a lot to my family and friends outside the department who helped me to relax and to enjoy my thesis by stopping it from becoming too important.

This work was payed for by the Natural Environment Research Council.

Abstract

This thesis describes the formulation and operation of a two-dimensional model of the atmosphere from 0 to about 100 km, using an isentropic vertical coordinate above the tropopause. The model is used in this thesis to study the dynamical effects of stratospheric planetary waves, arising from their transport of Ertel's potential vorticity (PV). The use of isentropic coordinates leads to conceptual simplifications, and to practical advantages in parametrising planetary-wave eddy fluxes in the stratosphere. The model has an interactive troposphere and reproduces the observed annual cycle of the equatorial tropopause temperature when an estimate of the planetary wave PV flux is included in the stratosphere. A feature of the radiation scheme used in this model, and apparently not shared by other 2D models, is the inclusion of the effect of the variation of daylight hours with height. This has a significant influence on the dynamical fields, especially during the equinoxes.

The model simulates the observed stratosphere well during autumn and in the southern hemisphere (SH) winter. However, to simulate spring, summer and the NH winter, an accurate estimate of the real PV flux is needed. Since no such estimate was available, three methods were used in this thesis to derive values of the flux for the year from July 1980 to June 1981. One method calculated the flux directly from approximate winds, another estimated the flux due to thermal dissipation of zonal asymmetries in PV, and the third was a variation of an existing method, which finds the PV flux needed to produce the evolution of the observed zonal wind in conjunction with the diabatic circulation. Bearing in mind the significant uncertainties present in all three estimates, the following description of the actual PV flux was put forward. The flux, when large, takes the form of a single negative peak, moving from middle to high latitudes over its life-time. In the NH, the flux is large (greater than 1 $ms^{-1}day^{-1}$ in magnitude) from October till March, reaching a peak of about $-8ms^{-1}day^{-1}$ in January/February. In the SH, the flux is large in April, June and from August till November, the peak value of about $-4ms^{-1}day^{-1}$ occuring in September/October. Thermal dissipation of PV anomalies appears to be the main cause of a PV flux at middle and high latitudes.

In connection with 2D models and the effect of planetary waves, a 'symmetric vortex' state is described and calculated for one year. It is obtained from the observed state by deforming the contours of PV till they coincide with latitude circles, while maintaining thermal wind balance, and conserving the mass of each

air parcel. It is expected that a 2D model fed with an accurate estimate of the monthly-mean PV flux will lie closer to this symmetric vortex state than to the observed zonally-averaged state. The symmetric vortex is not blurred by the effect of travelling waves and reversible vortex vacillations, and so can reveal the effect of irreversible PV fluxes more clearly than can the zonally-averaged state.

Contents

1 Two-dimensional models - waves, circulations and coordinates **6**

1.1 Outline of thesis 6

1.2 The need for 2D models - and their problems 8

1.3 Meridional circulation versus eddies 9

1.4 Adiabatic waves, cancellations and new circulations 12

1.5 Eddy Flux Parametrisation - The Transport Matrix 15

1.5.1 The Transport and Residual Circulations 17

1.6 Isentropic Coordinates 18

1.7 Dissipation - Chemical eddy fluxes 21

1.8 Perspective of 2D models 21

1.8.1 Classical Eulerian mean circulation 21

1.8.2 Residual circulation models 22

1.8.3 Isentropic coordinates 23

1.9 Conclusions 24

2 Ertel's Potential Vorticity and its flux **26**

2.1 Some Equations 26

2.2 Ertel's Potential Vorticity 28

2.3 Eddy Forcing of the Zonal Wind - $\mathcal{F}$ 30

2.4 Potential vorticity fluxes and $\mathcal{F}$ 31

2.5 Ways in which planetary waves can give rise to PV fluxes 31

2.5.1 Vacillation of the vortex 32

2.5.2 Wave breaking 33

2.5.3 Dissipation of PV 34

2.6 Calculation of winds and PV 34

2.7 Summary 39

3 A Symmetric Vortex **41**

3.1 Introduction 41

3.2 The effect of vacillations and free Rossby waves 44

3.3 Defining the symmetric vortex 45

3.4 Finding the symmetric vortex 46

3.4.1 Test of the numerical method 48

3.5 Differences between $\tilde{u}$ and $\overline{u}$, $\tilde{T}$ and $\overline{T}$ 49

3.5.1 The northern hemisphere 49

3.5.2 The southern hemisphere 51

3.6 Discussion 53

3.7 Summary 55

4 Estimation of $\mathcal{F}$ from observed Inter-hemispheric differences **57**

4.1 Introduction 57

4.2 Deductions from the small summer IHD 59

4.3 The equations and their solution 60

4.4 Estimates of $\mathcal{F}^{D}$ and $\mathcal{F}$, and their sensitivity to uncertainties . . . 62

4.4.1 Sensitivity to approximations 65

4.5 The effects of $\mathcal{F}$ revealed by IHDs 66
4.5.1 Effect of $\mathcal{F}$ 67
4.5.2 Tropopause cooling and the cross-equator effect 69
4.6 IHDs and the symmetric vortex 71
4.7 Summary 72

5 The Model **73**

5.1 Formulation of the Model 73
5.2 Method of solution 76
5.2.1 Sigma Coordinates 76
5.2.2 The stream function as a means of finding $\overline{V}$ 78
5.2.3 Spatial finite differencing 79
5.2.4 Time-stepping 83
5.2.5 Boundary conditions 83
5.2.6 Ellipticity, and stability fluxes 83
5.3 Parametrisation of physical processes 85
5.3.1 Treatment of vertical eddy fluxes 85
5.3.2 Planetary waves in the stratosphere - PV flux 88
5.3.3 Friction and gravity waves in the mesosphere and stratosphere 89
5.3.4 Radiative Heating in the stratosphere and mesosphere 89
5.3.5 Variation of daylight hours with height 90
5.3.6 The Troposphere 93
5.3.7 Radiative Cooling in the troposphere 95
5.3.8 Convection 96
5.3.9 Surface Heating - sensibile and latent heat fluxes 96

5.3.10 $\mathcal{F}$ in the troposphere and Surface Friction 97

5.4 Summary 98

6 Control run and effect of changes in tropospheric parametrisations 100

6.1 Model run without stratospheric $\mathcal{F}$ 100

6.1.1 Comparison with observations 101

6.1.2 Heating rates, $\mathcal{F}$ and meridional circulation 106

6.2 Sensitivity to Tropospheric Changes 107

6.2.1 Reduced tropospheric friction 109

6.2.2 Increased Cooling 110

6.2.3 Increased Critical Lapse Rates 112

6.2.4 Decreased Vertical Diffusion 112

6.3 Summary and Conclusions 115

7 PV flux estimates and their effect on the model 117

7.1 Sources of uncertainty in the calculated PV flux 118

7.2 Calculation of PV flux from estimated winds 119

7.2.1 Features of PV flux 119

7.2.2 Comparison with Chapter 4 122

7.2.3 Effect on model 122

7.3 Comparison of model $\overline{u}$ and $\overline{T}$ with observations 126

7.4 Effect of $\mathcal{F}$ on the tropopause, and the cross-equator effect 130

7.5 Conclusions and Summary 133

8 PV flux due to dissipation 135

8.1 Introduction 135

8.2 Dissipation of PV 136

8.3 Estimation and parametrisation of dissipative flux 141

8.4 Previous estimates of the dissipative potential vorticity flux 142

8.5 Some sources of error in estimating the PV Flux 143

8.5.1 Lengthening of the thermal dissipation time-constant due to local geostrophic effects 144

8.6 The estimated dissipative flux 147

8.6.1 Comparison with Chapters 4 and 7. 150

8.7 Effect on the model 151

8.8 Summary and conclusions 154

9 Summary and Conclusions **157**

9.1 The isentropic model 157

9.2 Estimate of the real PV flux due to planetary waves 159

9.2.1 Using the IHD in the diabatic circulation - Chapter 4 160

9.2.2 Direct calculation using approximate winds - Chapter 7 . . . 160

9.2.3 Dissipation of zonal-asymmetries of PV - Chapter 8 161

9.2.4 Deduction of main features of real PV flux 163

9.2.5 Ways of improving the estimate of PV flux, or $\mathcal{F}$ 164

9.3 The symmetric vortex 165

A Appendix **167**

A.1 Table of Symbols 167

A.2 Note on approximate heights used in graphs 169

B Bibliography **171**

Chapter 1

Two-dimensional models - waves, circulations and coordinates

The atmosphere's not what it seems.
The fact that waves aren't zonal means
a headache - but, don't give up hope,
for look! here comes an isentrope!

1.1 Outline of thesis

The work of this thesis is directed towards the construction of a two-dimensional (2D) model which can be used to predict the climatological behaviour of the zonally-averaged chemical and dynamical state of the middle atmosphere. The need for this kind of model has become greater in recent years as man releases chemicals into the atmosphere, such as carbon-dioxide and CFCs, which will have a long-term and, as yet, undetermined effect. The choice of a 2D, rather than a 3D, model is governed by the prohibitive length of computer time needed to run a 3D model which includes all the necessary chemical reactions. 2D models provide a compromise in which both the chemistry and dynamics of the atmosphere can be modelled interactively.

The main problem facing zonally-averaged models is their need for a realistic parametrisation of the fluxes produced by zonally-asymmetric waves which, when viewed in non-isentropic coordinates, can transport matter both horizontally and vertically. In isentropic coordinates, however, this transport is almost completely horizontal, which leads to a simpler parametrisation. This problem, and the advantages offered by isentropic coordinates, are discussed in this chapter. Chapter 2

introduces the basic dynamical equations used in the thesis, and Ertel's potential vorticity (PV). In isentropic coordinates the horizontal flux of this quantity by the waves accounts for most of the direct wave-forcing of the zonal-mean dynamical state.

It is important therefore to estimate the size of this flux in the real atmosphere so that it can be parametrised in the model. An accurate estimate, however, is difficult to obtain. Chapters 4, 7 and 8 describe three different estimates of this PV flux. Chapter 7 describes a direct calculation using approximate winds derived from satellite data. Chapter 4 estimates the flux indirectly from the effect it has on observed winds and temperatures while Chapter 8 assumes the flux arises purely as a result of thermal dissipation of zonal asymmetries in PV. All these methods involve approximations, and produce different results, but comparison enables useful information on the real PV flux to be extracted.

In the hope that isentropic coordinates will be superior to non-isentropic coordinates in the construction of 2D models of the stratosphere, the thesis centres on the isentropic model described in Chapter 5. As described in that chapter, it does not include an interactive chemical scheme, nor does the PV flux depend on the zonal-mean state. Both of these interactions have been included in other models using isobaric coordinates, and will eventually need to be included in this model if it is to be used to predict the future state of the stratosphere. Even without these interactions, though, it can still provide information on the effect of planetary waves on the zonal-mean state.

Chapter 6 describes the performance of the model in the absence of any specified PV flux, and also describes its sensitivity to changes in the parametrisation of the physical processes in the troposphere. The effect on the model of the PV fluxes of Chapters 7 and 8 is described in those chapters and comparison with the observed atmosphere reveals several features attributable to a real PV flux.

Chapter 3 lies somewhat off the main track of the thesis, describing how a zonally symmetric reference state was obtained from any observed atmospheric state by 'straightening out' the contours of PV while conserving PV and maintaining thermal wind balance. It is suggested that 2D models should perhaps aim to reproduce this reference state rather than the monthly-mean zonal-mean observed state. The reference state also reveals the effect of irreversible PV fluxes on the atmosphere more clearly than does the unsymmetric state, a point which is touched on in

Chapters 4 and 7. It is also used in Chapter 8 to estimate the PV flux.

Chapter 9 gathers together the main conclusions of the thesis and suggests areas for future development.

The appendix contains a table of the symbols used in the thesis, along with a note on the vertical 'approximate height' scale used in some of the graphs.

In this introductory chapter the need for two-dimensional (2D) models (which represent only the zonal-mean state of the atmosphere) will be outlined, along with the problems that their simplified representation introduces. A review of the advances made in understanding the factors which influence the zonal-mean state of the atmosphere, the stratosphere in particular, and the effect these advances have had on 2D models will then be given. The development will centre round the complementary roles played by the zonal-mean meridional circulation and the eddy fluxes. One of the latest developments in the theory has suggested that isentropic coordinates might provide a more practical, and a conceptually simpler, framework on which to build a 2D model. Since the primary purpose of this thesis is the construction and development of such a model, the advantages offered by isentropic coordinates will be sketched.

1.2 The need for 2D models - and their problems

Recently, people have become aware that their activities do affect the atmosphere. In order to understand the current state of the atmosphere and predict its evolution in the face of anthropogenic perturbations, numerical models have been developed which try to include as many of the important atmospheric processes and chemical reactions as possible. For a full, 3-dimensional, model this would be very computationally expensive, and in order to reduce this problem, 2D models have been developed which represent only the averages around latitude circles and so reduce the three dimensions of the atmosphere to two (height and latitude).

The penalty to be paid for this reduction is that of parametrising the eddy fluxes. Eddy fluxes arise when zonally averaging the equation governing the evolution of any quantity χ. The un-averaged equation is

$$\chi_t + u\chi_x + v\chi_y + w\chi_z = S$$

where S is the source of χ, and u, v and w are the zonal (eastward), northward and vertical winds. A subscript x, y, z or t denotes differentiation by that coordinate. When χ, u, v and w are split into zonal-mean (i.e. the average around a latitude circle) and eddy parts (eg. $u = \overline{u} + u'$) and the equation is zonally-averaged, it is found, using the continuity equation $u_x + v_y + w_z = 0$, that

$$\overline{\chi}_t + (\overline{v}\,\overline{\chi})_y + (\overline{w}\,\overline{\chi})_z = \overline{S} - (\overline{v'\chi'})_y - (\overline{w'\chi'})_z \tag{1.1}$$

The mean meridional circulation is then $\overline{v}$ and $\overline{w}$ and the eddy fluxes are $\overline{v'\chi'}$ and $\overline{w'\chi'}$. A way in which eddies can transport heat and momentum meridionally is illustrated in figure 1.1. Figures 1.1(a) and (b) both illustrate a wind in the

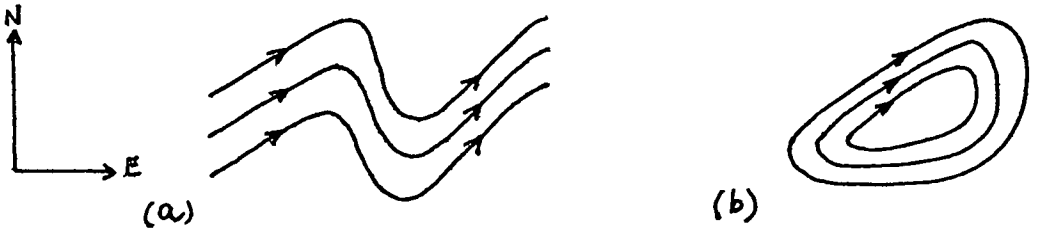


Figure 1.1: *Eddies typical of (a) the stratosphere (b) the troposphere*

zonal (east-west) direction which is more westerly when an air parcel is travelling northward than when it is travelling southward. There is therefore a northward momentum flux ($\overline{v'u'} > 0$). Also, for horizontal motion, if the air parcels are being diabatically heated while they are south of their mean position and cooled while they are north of it, they will be transporting heat northwards.

Since 2D models only represent the zonal-means, the eddy fluxes cannot be calculated explicitly from model variables and must be estimated in some other way. This is the main problem facing 2D models. Because both eddy fluxes and the mean circulation affect $\overline{\chi}_t$, it is necessary to estimate their relative contributions. The following section summarises the development of understanding of the wave-mean flow interaction.

1.3 Meridional circulation versus eddies

One of the first attempts at deducing the meridional circulation of the middle atmosphere was that of Brewer (1949), who suggested that the observed dryness of the stratosphere, compared with the troposphere, could be explained by a two-celled

circulation, with rising motion over the equator and sinking over the poles. The upward motion would cool the air enough for the water vapour to freeze and fall out before it enters the stratosphere. Dobson (1956) further pointed out that such a circulation would explain the observed anomalously high values of ozone in the polar lower stratosphere, far from the main photochemical source in the strongly illuminated tropical regions. A circulation similar to this 'Brewer-Dobson' circulation had been proposed in 1735 by Hadley to explain the surface winds. Above this two-celled circulation, Dutsch (1946) had earlier proposed a one-cell circulation from the summer to the winter pole in the upper stratosphere. Murgatroyd and Singleton (1961) calculated the mean meridional circulation which would be required to balance their calculations of the zonal-mean radiational heating rate (the adiabatic cooling due to rising motion balancing the diabatic heating). This so-called 'diabatic circulation' was qualitatively similar to the combined Brewer-Dobson and Dutsch circulations (see fig 1.2).

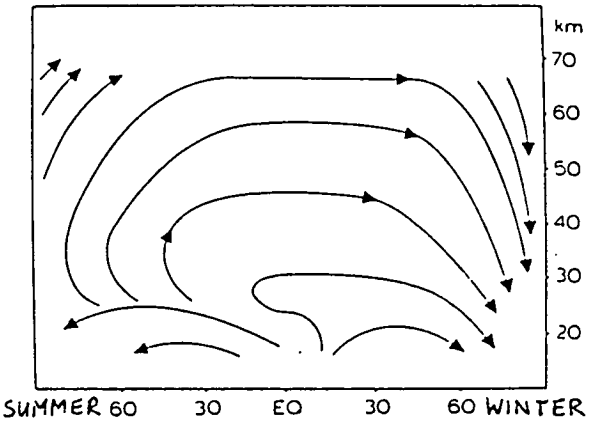


Figure 1.2: *The diabatic circulation (taken from Dunkerton 1978, who used Murgatroyd and Singleton's calculations.)*

This cannot, however, be the full picture, since the Brewer-Dobson circulation is continually transporting angular momentum from the earth's surface to the atmosphere, the surface easterlies slowing down, through friction, the earth's rotation while the poleward flow in the upper branches accelerates the zonal wind through the Coriolis torque. In fact, the theory fell out of favour when it was observed that in the northern hemisphere (NH) winter there was a zonal-mean rising motion in mid-latitudes and sinking in the sub-tropics. This indirect circulation is termed the Ferrel cell. It is driven by the eddy fluxes, which had previously been considered unimportant.

The Ferrel cell acts to balance the effect of the observed tropospheric and lower stratospheric eddy heat and momentum fluxes (see figure 1.3). The eddy heat

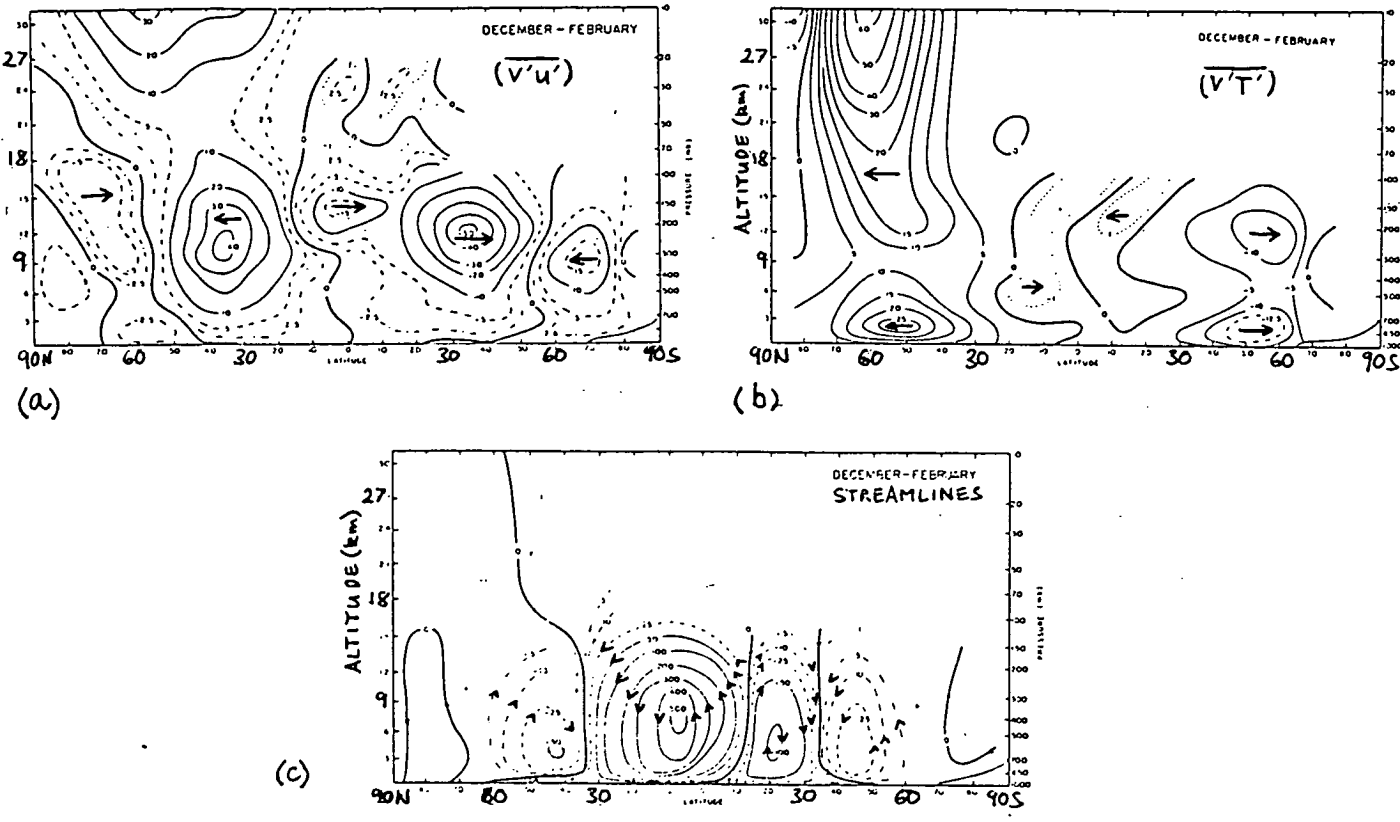


Figure 1.3: *(a) Eddy momentum fluxes, (b) heat fluxes and (c) the two-celled circulation for Dec-Feb (from Newell 1975)*

fluxes transport heat from about 30N but this cooling is balanced by the adiabatic heating of the sinking motion there. Similarly, the eddy-flux heating near the pole is opposed by the rising motion. The momentum flux convergence likewise opposes the Coriolis force due to the meridional circulation. For example, the convergence of the eddy fluxes at about 60N in the lower stratosphere is opposed by the equatorwards flow there. Mahlman (1969) noted that in the stratosphere also the heating effects of the mean meridional circulation and the eddy fluxes act in opposition. As pointed out by Matsuno (1980), though, there is a difference between the tropospheric Ferrel cell and its extension in the stratosphere. In the troposphere the cancellation between the eddy heat fluxes and the induced Ferrel cell cannot be large, since there must be a large transport of heat polewards by the eddies to balance the excess solar heating in the tropics. In the lower stratosphere the excess is smaller and there is a large cancellation between the eddy fluxes and the meridional circulation.

Hunt and Manabe (1968) pointed out that the eddy fluxes of tracers in their 3D general circulation model (GCM) opposed the effect of the mean meridional circulation. (Harwood and Pyle (1977) noted a similar cancellation for ozone in their two-dimensional model, though this was likely due to the diffusive parametrisation of the eddy fluxes). Recognition of this cancellation between the mean meridional circulation and eddy fluxes lead to, or was accompanied by, the theoretical developments described in the following section.

1.4 Adiabatic waves, cancellations and new circulations

Although Matsuno's (1980) work followed earlier developments, it is useful to consider it first, since it gives us a physical picture of how adiabatic, steady waves can produce the above kind of cancellation. He studied planetary-wave solutions to the atmospheric equations on a steady background zonal-mean state. The fact that the zonal-mean state was unchanging and yet there were waves present meant that the waves were having no net effect on the zonal-mean state. What makes this interesting is that the waves *were* producing eddy fluxes and yet these were not affecting the zonal-mean state since the waves were also the cause of a cancelling mean meridional circulation.

He considered a channel, bounded by two walls at different latitudes, with planetary waves whose amplitudes decreased to zero as the walls were approached but were independent of height. The zonal-mean potential temperature, $\overline{\theta}$, increased linearly with height, but was independent of latitude. The planetary waves were *adiabatic and steady* i.e. the θ of every air parcel was conserved, and the wave amplitudes were independent of time. The trajectories of parcels, when projected onto a latitude-height plane, were ellipses, but in our diagram are drawn as circles. (See Figure 1.4). All the waves have the same frequency, so to the right of plane P there will be an upward mean velocity, since along any line of latitude the upward-moving parcels will be moving faster on average than the downward-moving ones, and similarly to the left of P there will be a downward mean velocity. Along any horizontal plane there will be an eddy flux of θ to the right, since the parcels moving to the right will have come from above and will have a higher θ than the mean θ at that height. Due to the special conditions imposed on the motion there is no vertical eddy flux. Using the fact that $\overline{\theta}_y \equiv 0$, the equation of evolution for $\overline{\theta}$ is

$$\overline{\theta}_t = -\overline{w}\,\overline{\theta}_z - (\overline{v'\theta'})_y \qquad (1.2)$$

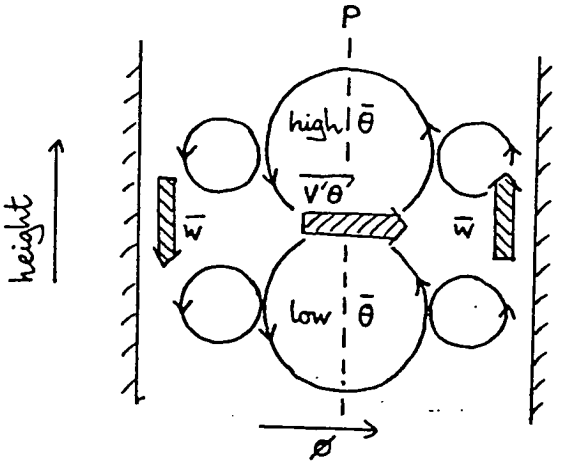


Figure 1.4: *Projections, on the meridional plane, of trajectories in planetary waves*

and since $\overline{\theta}_t \equiv 0$ we have

$$\overline{w} = -\left(\frac{\overline{v'\theta'}}{\overline{\theta}_z} \right)_y \tag{1.3}$$

This applies not only to θ, but also to any quantity with no source. Thus, it is possible that planetary waves can be the cause of both a mean circulation and of exactly cancelling eddy fluxes, so that there is no net effect on any quantity which does not have an eddy source. What would be useful is a circulation which says more about the time-averaged motion of individual parcels, without the complications caused by such planetary waves.

Various solutions have been proposed. Consideration of the above example suggests the 'residual circulation' $(\overline{v}^*, \overline{w}^*)$, first introduced by Andrews and McIntyre (1976).

$$\overline{v}^* = \overline{v} - \psi_z \text{ and } \overline{w}^* = \overline{w} + \psi_y, \text{ where } \psi = (\overline{v'\theta'})/\overline{\theta}_z \tag{1.4}$$

Note that the residual circulation also satisfies the continuity equation. In this case, $\overline{v}^* = \overline{w}^* = 0$. Alternatively, we could follow the motion of the average position of the string of parcels which in the absence of waves would have the same latitude and height. This defines the Generalised Lagrangian Mean (GLM) circulation, (see McIntyre 1980), which in this case would also be zero. Another alternative is to follow the motion in isentropic coordinates. A parcel's 'height' in isentropic coordinates is a function only of its entropy. Its entropy can change only if it is being diabatically heated or cooled, so the vertical velocity in isentropic coordinates is determined solely by the diabatic heating rate, and thus the circulation in isentropic coordinates is similar (though not identical) to the diabatic circulation of fig. 1.2. In the above example, since θ is conserved, the parcels remain on the

same isentropic surface and only oscillate horizontally about their mean position. Again, there is therefore no mean meridional circulation.

The relation between the residual, Lagrangian and diabatic circulations was elucidated by Dunkerton (1978), whose argument will be broadly followed here. He considered the zonal-mean thermodynamic equation in log-pressure coordinates at solstice so that he could neglect the time derivative of temperature, and further neglected vertical eddy fluxes and advection by $\overline{v}$. He could then write

$$\overline{\theta}_z \overline{w} = \overline{Q} \frac{\overline{\theta}}{\overline{T}} - (\overline{v'\theta'})_y, \tag{1.5}$$

where Q is the diabatic heating rate. In terms of the residual circulation defined above, this is

$$\overline{\theta}_z \overline{w}^* = \overline{Q} \frac{\overline{\theta}}{\overline{T}}, \tag{1.6}$$

where horizontal variations in $\overline{\theta}_z$ are neglected. This is actually the definition of the diabatic circulation of Murgatroyd and Singleton (1961), with the vertical motion balancing the diabatic heating. The residual circulation is therefore, under these approximations, equal to the diabatic circulation. Now the first-order approximation (for small waves) to the Lagrangian vertical velocity is

$$w^L \simeq \overline{w} + \overline{\eta' w'_y}, \tag{1.7}$$

where η is the meridional displacement of the air parcel. Using the eddy thermodynamic equation, defining velocities in terms of derivatives of displacements (eq.1.10 of the next section), assuming that $\overline{\theta}_t \ll \theta'_t$ and that the waves are steady and adiabatic Dunkerton showed that

$$\overline{v'\theta'} \simeq -\overline{\eta' w'}\, \overline{\theta}_z \tag{1.8}$$

On combining equations 1.4, 1.7 and 1.8 (and, like Dunkerton, neglecting η'_y) it is seen that the Lagrangian mean, residual and diabatic circulations are all equal under these approximations. Dunkerton thus concluded that the Lagrangian mean circulation, which gives us the mass transport, is approximately given by the diabatic circulation when waves are small, steady and are not being dissipated, and when $\overline{\theta}_t$ is small. Under these conditions also, the residual and diabatic circulations are approximately equal.

Thus the Brewer-Dobson circulation has now found favour again, provided that it is interpreted as an approximation to the Lagrangian mean circulation rather than the zonal-mean meridional circulation in isobaric coordinates. Dunkerton's work

has been used by many authors to justify the use of the diabatic circulation in 2D models (eg. Rogers and Pyle (1980), Ko et al (1985)).

A way will now be considered in which the eddy fluxes may be parametrised. Besides being important for modelling purposes, it offers an alternative view of the cancellation between eddy fluxes and mean circulation.

1.5 Eddy Flux Parametrisation - The Transport Matrix

For a quantity χ with no eddy source, and for small waves,

$$\chi'_t + \overline{u}\chi'_x + v'\overline{\chi}_y + w'\overline{\chi}_z = 0 \tag{1.9}$$

where $\overline{v}$ and $\overline{w}$ have been neglected. Introducing displacements η and ξ such that

$$v' = \eta_t + \overline{u}\eta_x \text{ and } w' = \xi_t + \overline{u}\xi_x \tag{1.10}$$

and assuming that $\overline{\chi}_t \ll \chi'_t$, leads to

$$\chi' + \eta\overline{\chi}_y + \xi\overline{\chi}_z = 0 \tag{1.11}$$

and the eddy fluxes can be expanded as

$$(\overline{v'\chi'}) = -\frac{1}{2}(\overline{\eta^2})_t\overline{\chi}_y - (\overline{v'\xi})\overline{\chi}_z \tag{1.12}$$

and

$$(\overline{w'\chi'}) = -(\overline{w'\xi})\overline{\chi}_y - \frac{1}{2}(\overline{\xi^2})_t\overline{\chi}_z \tag{1.13}$$

Defining the 'transport matrix', $\underline{\underline{\mathrm{K}}}$, by

$$\begin{bmatrix} \overline{v'\chi'} \\ \overline{w'\chi'} \end{bmatrix} \equiv -\underline{\underline{\mathrm{K}}} \begin{bmatrix} \overline{\chi}_y \\ \overline{\chi}_z \end{bmatrix}, \tag{1.14}$$

we deduce, from equations 1.12 and 1.13, that

$$\underline{\underline{\mathrm{K}}} = \begin{bmatrix} K_{yy} & K_{yz} \\ K_{zy} & K_{zz} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(\overline{\eta^2})_t & \overline{v'\xi} \\ \overline{w'\eta} & \frac{1}{2}(\overline{\xi^2})_t \end{bmatrix} \tag{1.15}$$

The first use of such a transport matrix was by Reed and German (1965), who assumed the effect of eddies to be diffusive and took $\underline{\underline{\mathrm{K}}}$ to be symmetric. Later, it

was recognised by Matsuno (1980) that $\underline{\underline{K}}$ need not be entirely symmetric and that in fact for steady, adiabatic planetary waves $\underline{\underline{K}}$ is completely anti-symmetric. The natures of the symmetric and anti-symmetric parts of $\underline{\underline{K}}$ will now be explained.

$\underline{\underline{K}}$ may be split into its symmetric and anti-symmetric parts :-

$$\underline{\underline{K}} = \underline{\underline{K}}_S + \underline{\underline{K}}_A = \begin{bmatrix} B & D \\ D & C \end{bmatrix} + \begin{bmatrix} 0 & A \\ -A & 0 \end{bmatrix} \tag{1.16}$$

Now, if $\underline{\underline{K}}$ were totally antisymmetric the eddy fluxes would be

$$\overline{v'\chi'} = -A\,\overline{\chi}_z \text{ and } \overline{w'\chi'} = A\,\overline{\chi}_y$$

Substituting this into equation 1.1 gives

$$\overline{\chi}_t + \overline{v}\,\overline{\chi}_y + \overline{w}\,\overline{\chi}_z = (A\,\overline{\chi}_z)_y - (A\,\overline{\chi}_y)_z$$

or

$$\overline{\chi}_t + (\overline{v} + A_z)\overline{\chi}_y + (\overline{w} - A_y)\overline{\chi}_z = 0$$

so the antisymmetric part of $\underline{\underline{K}}$ contributes to the mean circulation and will *advect* $\overline{\chi}$ along the contours of A.

The effect of the symmetric part of $\underline{\underline{K}}$ can be seen by rotating the y,z axes through an angle α, where

$$\tan 2\alpha = \frac{2D}{B - C}$$

With this choice of α, $\underline{\underline{K}}_S$ in the new coordinates is diagonal. (The new coordinates are defined by the transformation matrix

$$T = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

so that

$$\begin{bmatrix} \overline{v'\chi'} \\ \overline{w'\chi'} \end{bmatrix}^N = T \begin{bmatrix} \overline{v'\chi'} \\ \overline{w'\chi'} \end{bmatrix}^O = -T\,\underline{\underline{K}}^O \begin{bmatrix} \overline{v'\chi'} \\ \overline{w'\chi'} \end{bmatrix}^O = -T\,\underline{\underline{K}}^O\,T^{-1} \begin{bmatrix} \overline{\chi}_y \\ \overline{\chi}_z \end{bmatrix}^N$$

where the superscripts O and N denote quantities in the old and new coordinates. Comparison with eq.1.14 reveals that $\underline{\underline{K}}^N \equiv T\,\underline{\underline{K}}^O T^{-1}$. It is then straightforward to prove that $\underline{\underline{K}}^N$ is diagonal if $\underline{\underline{K}}^O$ is symmetric). With a diagonal transport matrix, the flux along either axis is proportional to the zonal-mean gradient along that axis only, which is characteristic of diffusion. In the un-rotated coordinates, then,

a symmetric $\underline{\underline{K}}$ leads to diffusion along axes which are inclined at an angle α to the coordinate axes.

It has thus been shown that the symmetric part of $\underline{\underline{K}}$ has a diffusive nature while the antisymmetric part contributes to the advection of $\overline{\chi}$. It is this advective nature of $\underline{\underline{K}}_A$which gives rise to the cancellation between mean circulation and eddy flux.

Although a parametrisation of the eddy fluxes has been derived in this section, there is no reason to suppose that it will accurately reflect the nature of the real eddy fluxes. In particular, it is not expected to reproduce the behaviour of the real atmosphere during a sudden warming, when the waves are not small and the rate of change of the zonal-mean state is not small compared with that of the eddies. Nevertheless, Plumb and Mahlman (1987) have shown that a 2D model using such a transport matrix can reproduce fairly well the behaviour of a three dimensional model.

Two redefinitions of the mean meridional circulation will now be examined which attempt to remove this 'advective' part of $\underline{\underline{K}}$, and hence also the large fluxes which lead to the above cancellation.

1.5.1 The Transport and Residual Circulations

Plumb and Mahlman (1987) defined the 'transport circulation'

$$\overline{v}^T = \overline{v} + A_z = \overline{v} + \frac{1}{2}(\overline{v'\xi'} - \overline{w'\eta'})_z$$

and

$$\overline{w}^T = \overline{w} - A_y = \overline{w} - \frac{1}{2}(\overline{v'\xi'} - \overline{w'\eta'})_y.$$

The advective flux due to the eddies is completely absorbed in the redefinition, and the effect of the eddies is now solely diffusive. The transport circulation cannot readily be calculated from real data however, since it requires knowledge of the displacements η and ξ. Plumb and Mahlman therefore used a 3D model to calculate the transport circulation.

The residual circulation arises as a more practical redefinition of the meridional circulation, by considering adiabatic eddies. These eddies conserve θ', so that, replacing χ by θ in 1.9, and following through the previous argument up to equation

1.12 we can write

$$\overline{v'\theta} = -\frac{1}{2}(\overline{\eta^2})_t\,\overline{\theta}_y - \overline{v'\xi}\,\overline{\theta}_z. \tag{1.17}$$

If the eddies are also stationary, so that

$$(\overline{\eta^2})_t = (\overline{\xi^2})_t = (\overline{\eta\xi})_t = 0, \tag{1.18}$$

then

$$\underline{\underline{K}} = \begin{bmatrix} 0 & -(\overline{v'\theta})/\overline{\theta}_z \\ (\overline{v'\theta})/\overline{\theta}_z & 0 \end{bmatrix}, \tag{1.19}$$

using Equations 1.15 and 1.17. The transport matrix can therefore be completely absorbed by a redefinition of the meridional circulation

$$\overline{\mathrm{v}}^* = \overline{\mathrm{v}} - \left(\frac{\overline{(\mathrm{v}'\theta')}}{\overline{\theta}_z}\right)_z$$

and

$$\overline{\mathrm{w}}^* = \overline{\mathrm{w}} + \left(\frac{\overline{(\mathrm{v}'\theta')}}{\overline{\theta}_z}\right)_y$$

which is the residual circulation mentioned before.

The transport matrix can be redefined in the residual circulation as

$$\underline{\underline{K}}_R \equiv \underline{\underline{K}} - \begin{bmatrix} 0 & -(\overline{v'\theta})/\overline{\theta}_z \\ (\overline{v'\theta})/\overline{\theta}_z & 0 \end{bmatrix},$$

so that the remaining eddy flux is now that due solely to diabatic, transient or large waves. Therefore, for steady, small, adiabatic waves the eddy flux of any quantity is completely absorbed in the redefinition of the mean meridional circulation.

1.6 Isentropic Coordinates

The use of isentropic coordinates for stratospheric modelling purposes has been pioneered mainly by Tung (1982, 1986, 1987). He showed that in isentropic coordinates unsteady, adiabatic waves produce a $\underline{\underline{K}}$ with K_{yy} the only non-zero element, since the vertical eddy velocity is zero. The usefulness of isentropic coordinates, then, depends on how adiabatic eddies are in the real atmosphere and so how justified we are in neglecting vertical eddy fluxes. Pawson and Harwood (1990) have calculated vertical eddy fluxes of ozone in isentropic coordinates and found them to be small in comparison to the transport by the mean meridional circulation in the upper stratosphere.

Because, as noted above, the diabatic and isentropic circulations are similar and hence also similar to the Lagrangian mean circulation, the circulation in isentropic coordinates should be a good indicator of the mean transport of matter. The simplicity of the thermodynamic equation in isentropic coordinates (which relates the vertical velocity to the diabatic heating) compared with that in isobaric coordinates (which includes the time derivative of temperature, horizontal advection and eddy fluxes) leads to conceptual advantages (eg. Andrews (1986) and section 4.5.1 of this thesis).

Tung (1986) also showed (as did Andrews 1986) that the eddy forcing of the zonal-mean zonal wind for adiabatic waves is given by the eddy flux of Ertel's potential vorticity. He further showed that even for diabatic waves this still holds, to a very good approximation. It was known that in isobaric coordinates the eddy forcing of $\overline{u}$ is given by the eddy flux of quasi-geostrophic potential vorticity (e.g. review by Andrews 1987), but that this only held for adiabatic, *small amplitude, geostrophic* waves. As Tung (1986) points out, quasi-geostrophic potential vorticity is in general less well conserved than Ertel's potential vorticity and the geostrophic approximation is rather restrictive. Note that for small, steady, adiabatic waves the eddy flux of Ertel's PV will be zero (since $K_{yy} = 0$) which is a direct proof of the EP flux theorem in isentropic coordinates. This theorem states (see for example Andrews et al. 1984) that steady, small, adiabatic waves cannot force a change in $\overline{u}$.

For adiabatic, non-steady waves in isentropic coordinates $\underline{\underline{K}}$ has no antisymmetric part, since K_{yy} is the only non-zero element, and therefore the eddy fluxes are entirely diffusive (see section 1.4). The partial cancellation in isobaric coordinates, due to the presence of adiabatic waves, of the mean circulation flux by the anti-symmetric part of $\underline{\underline{K}}$ is therefore not possible in isentropic coordinates. This applies to the eddy flux of any quantity. If the eddies are strongly non-adiabatic then even in isentropic coordinates a cancellation can occur through $\underline{\underline{K}}$ having an antisymmetric part.

To illustrate this latter point we consider a situation similar to Matsuno's, but now with diabatic eddies. Consider air parcels which are cooled as they move polewards, and heated as they move equatorwards, but in such a way that they follow elliptical orbits in the meridional plane (see figure 1.5 in isentropic coordinates (i.e. the total heating during one north-south oscillation is zero). There will thus be a mean sinking motion at the pole and mean rising near the equator, with mean poleward

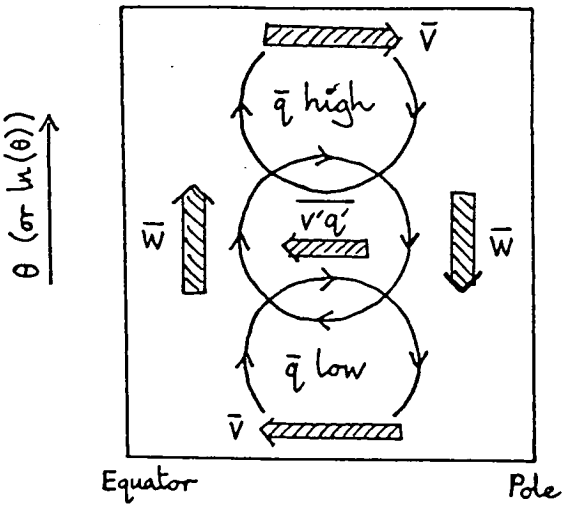


Figure 1.5: *Diabatic waves but no net transport (schematic projection of parcel trajectories in the meridional plane)*

and equatorward flow at the top and bottom of the region of waves. If the zonal mean mixing ratio of a tracer, q, has a positive vertical gradient then there will be an equatorwards eddy flux of q inside the region of waves, but this will be balanced by the vertical advection of $\overline{q}$ on the side boundaries of the region. Thus the diabatic waves are the cause of eddy fluxes and a cancelling mean meridional circulation.

However, the advantage of isentropic coordinates is that there are no eddy fluxes in the thermodynamic equation in isentropic coordinates, even if the eddies are not adiabatic. The residual circulation was defined so that it removed the eddy fluxes from the thermodynamic equation, but only when the waves were adiabatic, steady and small. (It is easy to show that for adiabatic, steady, small waves,

$$\overline{v'\theta'}\,\overline{\theta}_y = -\overline{w'\theta'}\,\overline{\theta}_z.$$

The zonal-mean thermodynamic equation can then be rewritten as

$$\begin{aligned}\overline{\theta}_t + \overline{v}\,\overline{\theta}_y + \overline{w}\,\overline{\theta}_z &= -(\overline{v'\theta'})_y - (\overline{w'\theta'})_z + \overline{Q} \\ &= -\left(\frac{\overline{v'\theta'}}{\overline{\theta}_z}\right)_y + \left(\frac{\overline{v'\theta'}}{\overline{\theta}_z}\right)_z + \overline{Q}\end{aligned}$$

The eddy fluxes are thus absorbed in the definition of the residual circulation. Therefore for non-adiabatic, non-steady or large waves, the complication of having eddy fluxes in both the thermodynamic and momentum equations arises even when using the residual circulation, while in isentropic coordinates the eddy forcing still occurs only in the momentum equation. This forcing can be well approximated by

the eddy flux of Ertel's PV. As yet, there have been few publications of estimates of this flux, and a large part of this thesis will be devoted to obtaining an estimate of it in three different ways.

1.7 Dissipation - Chemical eddy fluxes

In section 1.5 we did not consider the situation where χ has an eddy source. Eddy fluxes would arise in the same way as the heat fluxes described in the text pertaining to figure 1.1, with the χ of an air parcel increasing if it is lower than the zonal-mean χ, and decreasing if it is higher. The transport matrix corresponding to a small dissipative source, with dissipative time constant α, is given by (see eg. Smith et al. 1988)

$$\underline{\underline{K}} = \alpha \begin{bmatrix} \overline{\eta'^2} & \overline{\eta'\xi'} \\ \overline{\eta'\xi'} & \overline{\xi'^2} \end{bmatrix}$$

This symmetric $\underline{\underline{K}}$ leads to a down-gradient eddy flux. Other forms of $\underline{\underline{K}}$ which are not entirely symmetric have been derived (see for example Matsuno 1980, Rogers and Pyle 1980, Smith et al. 1988). The PV flux due to dissipation will be estimated in Chapter 8.

1.8 Perspective of 2D models

In this section a brief review of some of the main 2D models will be given in order to put the model of this thesis into perspective. A wider assesment is given in WMO (1986) Chapter 12. The models will be divided into three groups depending on whether they use an Eulerian or a residual mean circulation, or isentropic coordinates. The model of this thesis will be included in this last group. An equally useful classification might be based on the number of interactions accounted for in the models.

1.8.1 Classical Eulerian mean circulation

A model which has been used extensively to study the dynamics and chemistry of the middle atmosphere is that of Harwood and Pyle (1977, 1980). Though its chemistry and dynamics have been updated, current versions still use an Eulerian

mean circulation and a log-pressure vertical coordinate. The horizontal momentum fluxes are specified from observations and heat and chemical eddy fluxes are parametrised using a diffusive transport matrix. The heating rate, and therefore the circulation, is calculated interactively from the modelled temperature and ozone. It produced fairly realistic ozone columns and displayed the observed partial cancellation between eddy and mean fluxes of ozone. This is not too surprising given the down gradient, diffusive, nature of the eddy fluxes. What is perhaps surprising is that the good results indicate that the small difference between the eddy and mean fluxes was modelled well. It is worth pointing out that the cancellation due to stationary, adiabatic, planetary waves is between advection by the mean circulation and the *advective* part of the eddy flux (see section 1.4). The nature of the cancellation in this model is therefore wrong since the eddy fluxes are entirely diffusive.

The model has been used, with improved chemical and radiative heating schemes, to study perturbations to the ozone layer by CFCs (Pyle 1980, Haigh and Pyle 1982), the semi-annual oscillation (Gray and Pyle 1987) and lower-stratospheric radiation schemes (Haigh 1984). It also provides some features for the design of the isentropic model of this thesis.

1.8.2 Residual circulation models

Following Dunkerton's (1978) elucidation of the similarity of the residual circulation to the Lagrangian mean circulation when the planetary waves are approximately stationary and adiabatic, many modellers have neglected the planetary-wave eddy fluxes in the thermodynamic equation and considered the resulting circulation as an approximation to the Lagrangian mean circulation. Effects of wave transience and dissipation are sometimes implicitly included by calculating the heating rates from observed temperatures (e.g. Jackman et al. 1988) or from the temperature fields of a model in whose dynamical equations planetary wave fluxes were included (Rogers and Pyle 1984). Other models (e.g. Schneider et al. 1989, Garcia and Solomon 1983, Hitchman and Brasseur 1988) calculate the heating rates interactively from the modelled temperatures, dynamical planetary wave fluxes being either ignored (Garcia and Solomon) or parametrised. Schneider et al. parametrise the potential vorticity flux using an idealised K_{yy} while Hitchman and Brasseur use a sophisticated technique to find $\nabla.F$ which allows an interaction between $\nabla.F$ and the zonal mean state.

The transport matrix used for parametrising the eddy flux of chemicals in these models is derived in a number of different ways. Hitchman and Brasseur divide the $\nabla.F$ calculated in their model (which is approximately equal to the eddy flux of quasi-geostrophic potential vorticity (QPV)) by the latitudinal gradient of QPV to obtain a K_{yy} which they then use for all chemicals. This implicitly assumes that the dissipative time-constant for QPV is the same as those of the chemicals. More recently, Smith and Brasseur (1990) attempted to take into account the variation of life times. As noted by Smith and Brasseur, QPV is transported only horizontally along isobaric surfaces, which enables them to deduce one element, K_{yy}, of the transport matrix, but does not allow them to deduce the other elements needed to parametrise the eddy flux of chemicals, without further assumptions. We note here that the use of isentropic coordinates would alleviate this problem since *all* eddy fluxes would be almost horizontal (see Tung 1982). Jackman et al. deduce the QPV flux required to balance the Coriolis torque of their residual circulation and, like Hitchman and Brasseur, divide this by the gradient of QPV to obtain a K_{yy}. Rogers and Pyle use modelled wave amplitudes to calculate the species-dependent chemical eddy fluxes, which they find to be significant and capable of producing realistic model results, while Garcia and Solomon use observed winter wave amplitudes.

1.8.3 Isentropic coordinates

Little modelling work has been done using isentropic coordinates. Ko et al. (1984) used an isentropic model with a mean circulation derived from heating rates calculated from observations to study the effect of the single diffusion coefficient, K_{yy}, on the distribution of gases which were effectively tracers. They found a good agreement with observations on using a value for K_{yy} which was small when compared with the 'chemical eddy' diffusion coefficients of Rogers and Pyle (1984). This suggests that the eddy fluxes due to wave transience and breaking are small compared with the chemical eddy fluxes of species with short dissipation time-constants.

Tung and Yang (1988) calculate the heating rate from four years of temperature observations and use the mean circulation derived from this to advect chemicals in their isentropic model. They derive a K_{yy} in a way similar to Newman et al. (1988) (used by Jackman et al. (1988)) by dividing the PV flux needed to balance the momentum equation by the latitudinal gradient of PV. The derivation of this single element of the diffusion matrix is all that is needed to parametrise eddy

fluxes by almost-adiabatic waves in isentropic coordinates, and avoids the problem faced by Smith and Brasseur (1990) mentioned in the previous section. However, Tung and Yang use the same K_{yy} for all chemicals, which neglects the variation in the dissipative time-constant among chemicals.

Because the rates of change of temperature and zonal wind are specified in Tung and Yang's model it is not interactive in the sense that, for example, Schneider et al's (1989) model is, where temperature and heating rate depend on each other. The model described in this thesis is the first such interactive model using isentropic coordinates. (As described, though, it has the same number of 'degrees of freedom' as Tung and Yang's model, PV flux being specified instead of temperature). As described in this thesis it does not have interactive chemistry, but a more up-to-date version does. A possible extension, to make it more fully interactive, so that PV fluxes need no longer be externally specified, could be along the lines of Hitchman and Brasseur's model. The advantage of isentropic coordinates in parametrising the effects of almost-adiabatic waves could then be used to overcome the problem faced by Smith and Brasseur in parametrising the species-dependent chemical eddies.

1.9 Conclusions

In this chapter were described some of the problems faced by 2D modellers, arising from the need to parametrise eddy fluxes. It was shown how the net transport can be the small residual of the effect of eddy fluxes and a cancelling mean meridional circulation. This cancellation can be greatly reduced by the redefinition of the mean meridional circulation as the residual or transport circulation. The transport circulation is impossible to deduce from current observations, and using the residual circulation the cancellation still occurs if the eddies are non-adiabatic, non-steady or large. There is therefore great need for an accurate parametrisation and observational estimate of the eddy fluxes.

The advantages of isentropic coordinates were described, which arise from the simple form of the thermodynamic equation, and the assumption that the eddies are almost adiabatic. Also, unlike the situation in isobaric coordinates, even using the residual circulation, the eddy forcing of winds and temperature occurs only in the momentum equation, and is mainly due to an eddy flux of Ertel's PV. Because eddy-forcing induces a mean meridional circulation and hence affects the

distribution of atmospheric constituents, it is important to know the extent of the eddy forcing of the dynamics of the real atmosphere, and in isentropic coordinates this means the eddy flux of Ertel's PV.

There exists a place in the current set of 2D models for a fully interactive isentropic model. The model to be described in this thesis includes the interaction between temperature and heating and is capable of being upgraded to include the interactions with chemistry and between PV flux and the zonal-mean state.

Chapter 2

Ertel's Potential Vorticity and its flux

Though confusion may reign
in this world of duplicity,
don't doubt the potential
of Ertel's vorticity.

In this chapter some of the basic dynamical equations used in this thesis will be introduced. It will be explained how the mean meridional flux of Ertel's potential vorticity can affect the zonal wind, and ways in which planetary waves can effect such a flux will be described. Finally, our method of calculating the winds and Ertel's PV from satellite data will be described.

2.1 Some Equations

In this section some equations used in the course of the thesis will be derived. Much use is made of the work of Tung (1986). A subscript x, y, z or t will denote differentiation with respect to that coordinate. The symbols used are listed in the appendix.

In isentropic coordinates, the momentum equations are identical in form to those in isobaric coordinates, except that geopotential, $\Phi = gz^*$, is replaced by Montgomery potential, $M = gz^* + c_pT$, where z^* is geopotential height. This is because in converting x and y derivatives from constant pressure surfaces to surfaces of constant potential temperature,

$$(\Phi_x)_{p=const} = (\Phi_x)_{\theta=const} - \Phi_p(p_x)_{\theta=const},$$

(equation (1.22) of Holton(1979)) is used to find that

$$(\Phi_x)_{p=const} = (\Phi_x + c_p T_x)_{\theta=const} = (M_x)_{\theta=const},$$

where we have used the definition of potential temperature,

$$\theta \equiv T\left(\frac{p_0}{p}\right)^{\kappa}, \tag{2.1}$$

the hydrostatic approximation,

$$\frac{\partial p}{\partial z^*} = \frac{-g\,p}{R\,T}, \tag{2.2}$$

and the identity

$$c_p = \frac{R}{\kappa}.$$

The momentum equations, then, are

$$\frac{Du}{Dt} - v\left(\frac{u\tan\phi}{a} + f\right) = -M_x \tag{2.3}$$

$$\frac{Dv}{Dt} + u\left(\frac{u\tan\phi}{a} + f\right) = -M_y \tag{2.4}$$

where the material derivative, $\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}$, and the horizontal space derivatives are performed on isentropic surfaces.

The vertical coordinate, z, can be any one-to-one function of θ, and in this thesis it is chosen to be

$$z = \ln\left(\frac{\theta}{250K}\right)$$

(except in the troposphere of our model where a 'sigma coordinate' is used - see section 5.2). In an isothermal atmosphere, with T = 250K and p = 1000mb at the surface, z would be proportional to the geometric height,

$$z = \kappa\left(\frac{z^*}{7.3km}\right)$$

so that a 'scale height' (which is about 7km in pressure coordinates) is approximately κ units in our coordinates. In the figures in this thesis plotted using isentropic coordinates the approximate height of the vertical axis is given by (z/κ) times 7km.

We will denote the 'density' in these isentropic coordinates by

$$\rho \equiv \frac{-1}{g} p_z, \tag{2.5}$$

so that, if the atmosphere is in hydrostatic equilibrium, the mass per unit area between two isentropic levels separated by an infinitesimal Δz will be given by $\rho \Delta z$.

The hydrostatic approximation, eq 2.2, can be written in terms of the Montgomery potential as

$$M_z = c_p T \tag{2.6}$$

which will be used in deriving the thermal wind equation in section 5.1.

Equation 2.3 can be rewritten in terms of angular momentum per unit mass

$$\tau = a\cos\phi(u + a\Omega\cos\phi) \tag{2.7}$$

as

$$\frac{D\tau}{Dt} = -a\cos\phi\, M_x \tag{2.8}$$

The equation of continuity is

$$\cos\phi\rho_t + \cos\phi(u\rho)_x + (\cos\phi\, v\rho)_y + \cos\phi(w\rho)_z = 0 \tag{2.9}$$

and the equation for the evolution of a mixing ratio χ is

$$\cos\phi(\chi\rho)_t + \cos\phi(\chi u\rho)_x + (\cos\phi\, \chi v\rho)_y + \cos\phi(\chi w\rho)_z = \rho S\cos\phi \tag{2.10}$$

where S is the source of χ.

2.2 Ertel's Potential Vorticity

Since Ertel's potential vorticity (hereafter refered to as PV) will play an important role in this thesis, it is useful to have a clear idea of what it is and what its properties are. The framework of isentropic coordinates offers the simplest definition and the clearest way of understanding PV. In these coordinates it has the definition

$$\Pi \equiv \frac{\zeta + f}{\rho}$$

where ζ is the relative vorticity, which, on a beta plane is $(v_x - u_y)$ and on a sphere is
$v_x - (u\cos\phi)_y/\cos\phi$. f is the planetary vorticity and ρ is the density in isentropic coordinates.

Usually Π is defined in coordinates where z=θ (see eg. Andrews et al. 1987), so that $\rho = \frac{-1}{g} p_\theta$. Our definition, then, differs by a factor of θ.

One of the main properties of PV, which will be made use of several times in this thesis, is that in the absence of friction or diabatic heating an air parcel will conserve its value of PV. (With regard to the note in the previous paragraph, an extra factor of θ in the definition of PV does not alter this property since in the absence of diabatic heating the θ of an air parcel is conserved). This property is most easily and clearly shown with the beta-plane equations, and the proof is included here to show how PV is conserved through the opposing effects of horizontal divergence on absolute vorticity and density. For a proof on a sphere see for example Andrews, Holton and Leovy (1987). The frictionless, adiabatic, momentum equations on a beta plane are

$$u_t + uu_x + vu_y - vf = -M_x \tag{2.11}$$

$$v_t + uv_x + vv_y + uf = -M_y, \tag{2.12}$$

where M is the Montgomery potential. Differentiating eq(2.12) w.r.t. x and eq(2.11) w.r.t. y and subtracting gives

$$(\zeta + f)_t + u(\zeta + f)_x + v(\zeta + f)_y = -(u_x + v_y)(\zeta + f). \tag{2.13}$$

Introducing the material derivative operator

$$D_t \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y},$$

which is the rate of change following an air parcel, eq(2.13) can be written as

$$D_t(\zeta + f) = -(u_x + v_y)(\zeta + f). \tag{2.14}$$

The continuity equation $\rho_t = -(\rho u)_x - (\rho v)_y$ can be written as

$$D_t\rho = -(u_x + v_y)\rho. \tag{2.15}$$

Thus divergence of the wind field changes the total vorticity of an air parcel, but it also changes its density. However, combining equations 2.14 and 2.15 gives

$$D_t(\log(\zeta + f)) = D_t(\log(\rho)),$$

so that $D_t(\log(\Pi)) = 0$ and hence Π is conserved following an air parcel.

Because the stratosphere appears to be fairly friction-free (gravity-wave breaking, the main source of 'friction', seems to occur mainly in the mesosphere, above about

60km, (see Marks(1989)), or at tropopause level (see Tanaka(1986)) and adiabatic over short (5-10 days) time scales (see the discussion of dissipative time-constants in Chapter 8), conservation of PV seems to be a reasonable first-order approximation for time-scales of about 5 days. However, as will be shown in Chapter 8, over the course of a month dissipation of PV anomalies may account for a large part of the PV flux.

Another useful property of PV is that, given suitable boundary conditions and a balance equation (such as thermal wind balance) it can be inverted to give the winds and temperatures (see eg. Hoskins et al. 1985).

The observed PV distribution in the stratosphere has predominantly positive values in the northern hemisphere (NH), and negative values in the SH, with a pool of large PV values near the pole (the vortex). PV values are strongest in winter, when there is also a definite edge to the vortex, equatorwards of which is the 'surf zone' where the PV gradient is small and planetary waves tend to break (see McIntyre 1982).

2.3 Eddy Forcing of the Zonal Wind - $\mathcal{F}$

An expression will now be derived for the eddy forcing of the zonal wind involving the meridional eddy flux of PV, broadly following Tung (1986).

Taking the zonal average of (2.8), gives

$$\overline{\tau}_t + \overline{\mathrm{v}\,\tau_y} + \overline{\mathrm{w}\,\tau_z} = 0$$

which can be expanded as

$$\overline{\tau}_t + \hat{\mathrm{v}}\overline{\tau}_y + \overline{\mathrm{v}^*\tau_y} + \hat{\mathrm{w}}\overline{\tau}_z + \overline{\mathrm{w}^*\tau_z} = 0$$

where $v^* = v - \hat{v}$ and $\hat{v}$ is the density-weighted zonal mean $\overline{v\rho}/\overline{\rho}$. Use of the zonal average of the continuity equation (2.9),

$$\overline{\rho}_t \cos\phi + \overline{\mathrm{V}}_\phi + \overline{\mathrm{W}}_z = 0 \tag{2.16}$$

where $\overline{\mathrm{V}} = \overline{v\rho}\frac{\cos\phi}{a}$ and $\overline{\mathrm{W}} = \overline{w\rho}\cos\phi$, gives

$$\cos\phi(\overline{\rho}\,\overline{\tau})_t + (\overline{\mathrm{V}}\,\overline{\tau})_\phi + (\overline{\mathrm{W}}\,\overline{\tau})_z = -\overline{\rho}\cos\phi[(\overline{\mathrm{v}^*\tau_y}) + (\overline{\mathrm{w}^*\tau_z})] \equiv \mathrm{a}\cos^2\phi\,\overline{\rho}\,\mathcal{F} \tag{2.17}$$

where $\mathcal{F}$ is defined as

$$\mathcal{F} \equiv \frac{-1}{a\cos\phi}(\overline{v^*\tau_y} + \overline{w^*\tau_z}).$$

Now, $\rho\Pi = v_x - \frac{(u\cos\phi)_y}{\cos\phi} + f = v_x^* - \frac{\tau_y}{\cos\phi}$ so

$$\overline{v^*\rho\Pi} = \overline{v^*v_x^*} - \frac{\overline{v^*\tau_y}}{\cos\phi} = -\frac{\overline{v^*\tau_y}}{a\cos\phi}$$

and so

$$\mathcal{F} = \overline{v^*\rho\Pi} - \frac{\overline{w^*\tau_z}}{a\cos\phi} \tag{2.18}$$

which relates $\mathcal{F}$, the eddy forcing of the zonal-mean zonal wind, to the sum of the eddy flux of PV and a term which is usually small (see next section).

2.4 Potential vorticity fluxes and $\mathcal{F}$

It shall be assumed here that the wave-forcing of the zonal wind, $\mathcal{F}$, can be well approximated in the stratosphere by the horizontal flux of PV due to large-scale waves. This is a reasonable stance to take for the following reasons. First, it is well accepted that gravity waves break mainly in the mesosphere and it is believed that they do not play a significant role in the stratosphere, although it is possible that gravity waves break near the tropopause and that Kelvin waves are important at all heights near the equator. Therefore, at least at middle and high latitudes, planetary waves are likely to be dominant in the stratosphere. Looking at equation (2.18) we see that whether $\mathcal{F}$ can be approximated by $\overline{v^*\rho\Pi}$ depends on the size of the vertical eddy-flux term. Tung (1986) has estimated it to be two orders of magnitude smaller than the horizontal flux term. However, although adiabatic waves imply that $\overline{w^*\tau_z} = 0$, neglecting $\overline{w^*\tau_z}$ does not imply that the waves are nearly adiabatic, but instead that the main effect of diabatic waves is to produce a *horizontal* flux of PV rather than a vertical eddy flux of τ.

2.5 Ways in which planetary waves can give rise to PV fluxes

A horizontal flux of PV can occur in one of two basically different ways. One is a physical movement of air parcels north and south, carrying with them their PV which will be conserved if there is no diabatic heating or frictional dissipation

occuring in the parcel. The other can occur even if the waves are steady, and is due to the presence of sources of PV which, if it tends to relax the parcel's PV towards some background, zonally-symmetric, value, will produce an eddy flux of PV down the gradient of the background distribution.

There is also a third way (discussed by McIntyre 1982) which, although it is a combination of the above two ways, merits being treated as a separate category. It is planetary wave-breaking, which involves a dispersion of air parcels north and south which then acquire the PV value of their new surroundings through dissipation. Large-scale motions tend to be more reversible than small-scale ones (thermal dissipation occurs more rapidly for small-scale anomalies of PV - see Haynes and McIntyre (1987), Gille and Lyjak (1986)), so a movement of the vortex away from and towards the pole should result in little permanent change to the vortex while the breaking-off of small parcels from the main vortex by an anticyclonic vortex (see for example O'Neil and Pope 1987) and their subsequent dispersal towards the equator should lead to a more irreversible flux of PV. We will now consider each of these three cases in more detail.

2.5.1 Vacillation of the vortex

When the vortex moves away from the pole there will be a flux of PV away from the pole, at least in the vicinity of the pole. This is fairly obvious, since if a vortex of positive PV, initially with its maximum value over the north pole, moves southward off the pole, (or a similar vortex of negative PV moves away from the south pole) the zonal-mean PV near the pole must decrease, and since the total PV is conserved this implies that the zonal-mean PV in some region nearer the equator must increase, and therefore there must be a negative flux of PV at mid-latitudes. A large part of this flux is likely to be an eddy flux since a movement of the vortex away from the pole is a zonally asymmetric event. This negative eddy PV flux will tend to decelerate $\overline{u}$ at mid-latitudes and, as will be shown in Section 4.2, a *polewards* mean meridional velocity will be induced which will lead to a mean flux of PV in opposition to the eddy flux. The deceleration of $\overline{u}$ at mid-latitudes and acceleration near the pole (due to the induced polewards meridional velocity) will shift the jet in $\overline{u}$ polewards, and from McIntyre's (1982) discussion this is likely to produce a region of positive refractive index near the pole, encouraging the convergence of the flux of planetary-wave activity there. This positive feedback could be responsible for the stratospheric sudden warmings.

The correlation between vacillations in PV flux and $\overline{u}$ is obvious when it is seen to be due to a vacillation of the polar vortex. The correlation between $\nabla.F$ and $\overline{u}_t$ has been observed and modelled by many authors (for example Hirooka and Hirota(1985), Salby and Garcia(1987)) and has been described in terms of the interference between stationary and travelling waves. Shiotani and Hirota(1985) show the correlation between $\overline{u}_t$ and the $\nabla.F$ in the NH winter (see Figure 2.1)

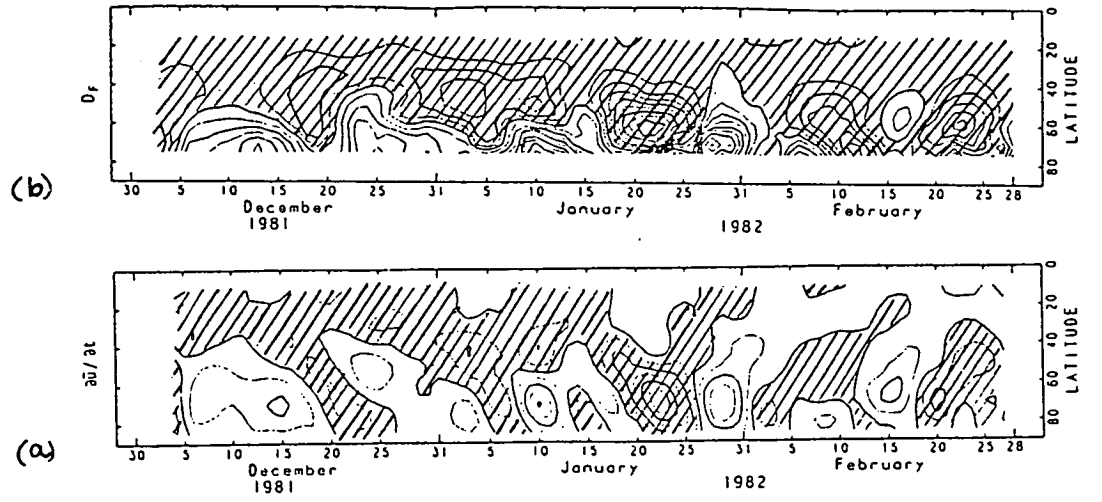


Figure 2.1: *(a)* $\overline{u}_t$ *and (b)* $\nabla.F$ *in NH winter at 5mb, negative regions shaded. (from Shiotani and Hirota(1985)*

2.5.2 Wave breaking

Planetary wave breaking, as first described by McIntyre (1982), is now accepted to be a frequent occurence in the NH winter stratosphere, (e.g. O'Neill and Pope (1984), McIntyre and Palmer (1983)) and reproduced even in a model which represents only wavenumber-1 (Kohno 1984). It seems to require the presence of an anticyclonic vortex, along with the main cyclonic vortex, (large wavenumber-1) which draws low-latitude (low-PV) air polewards and tears high-PV air from the edge of the cyclonic vortex which is then mixed with low-PV air at low latitudes. Plumb and Mahlman (1987) found that the horizontal dispersion in their GCM in winter, due mainly to wave breaking, was strongest in the regions where $\overline{u}$ was small. This ties in with 'critical layer' theory (see eg. McIntyre 1982) which recognises that since Rossby waves with zero phase speed cannot propogate into regions where $\overline{u}$ is easterly they must either be dissipated or reflected on encountering a 'critical line', where $\overline{u}$ is zero. The meridional dispersion of wave parcels leads directly to a down-gradient flux of PV, if PV is approximately conserved inside

the air parcels. Fig. 2.6 shows a typical breaking planetary wave in the NH winter, revealed by the distribution of PV which acts almost like a tracer over the short time-scales involved.

2.5.3 Dissipation of PV

Consider a parcel in a steady planetary wave, oscillating north and south about its mean latitude, ϕ_0, as it travels around the earth. If, while it is south of ϕ_0, it loses PV, and acquires it while it is north of ϕ_0, it must be producing a negative flux of PV since it is effectively transporting PV from north to south. If the waves are steady in the sense that the mean latitude of a parcel does not change then the mean meridional velocity, and hence the mean flux of PV, must be zero and so the flux of PV must be an eddy flux.

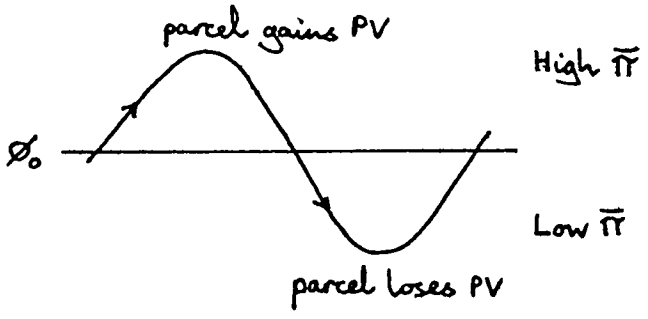


Figure 2.2: *How dissipation can lead to an eddy flux*

A way in which this situation may be described mathematically is as follows. The eddy source of Π is defined as $-\alpha\,\Pi'$, where α is a positive time-constant. Since $\overline{\Pi}_y$ is generally positive, if a parcel is displaced north of its mean position it will have a negative Π' and hence will gain PV, and if it is displaced south it will lose PV, thus we have the situation described in the previous paragraph. This definition of the eddy source has been used by many authors (eg. Smith et al (1988), Garcia and Solomon (1983), Matsuno (1980)) to deduce these dissipative fluxes. This kind of PV flux will be more fully discussed in Chapter 8.

2.6 Calculation of winds and PV

Having now established the importance of the eddy flux of PV, much of the rest of the thesis will be spent estimating its value in a number of different ways. All of these ways required values of horizontal winds on isentropic surfaces. The rest

of this chapter describes the way in which an estimate of these horizontal winds was derived from satellite data.

The data used came from the stratospheric sounding unit (SSU) on the NOAA-7 satellite (see Pick and Brownscombe 1981 and, for example, Clough et al. 1985). The data was supplied as values of geopotential height on pressure surfaces at points on a regular latitude-longitude grid, with a 5° spacing in both latitude and longitude. There were data points over the north and south poles, so there were 37 points from pole to pole, and 72 points around a circle of latitude. The pressure levels were 1000, 850, 500, 300, 200, 100, 50, 20, 10, 5, 2 and 1mb.

These geopotential heights were then interpolated onto a polar stereographic grid (see Haltiner and Williams (1980)) at each pressure level. Points spaced evenly on a polar stereographic grid are not clustered round the poles as in a latitude-longitude grid, and so the finite differencing used in deriving the wind fields leads to much smoother results. The relation between the two grids is shown in fig 2.3. A horizontal line in the figure crossing the pole, would have 37 points along it in

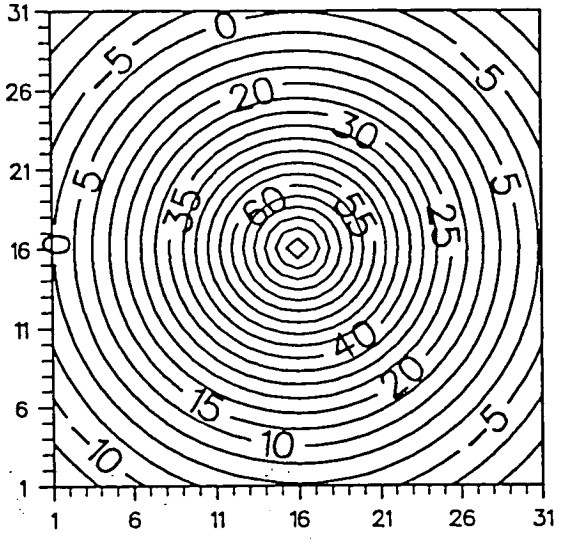


Figure 2.3: *Projection of latitude circles on the polar stereographic grid used.*

the lat-lon grid, and 31 in the polar-stereographic. The resolution, in this sense, is thus almost equal for the two grids although the average density of points in the lat-lon grid is about twice that of the polar-stereographic. However, it was found that increasing the resolution of the polar-stereographic grid much further resulted in increasingly noisy fields of PV.

In order to estimate the winds on isentropic surfaces we need to know the Mont-

gomery potential, $M = gz^* + c_pT$, on those surfaces. The isentropic levels used here were those at which $\ln(\theta/250)$ was equal to $(0.25 + J/2)\kappa$, for J=3 to 13. In an atmosphere in which the temperature is 250K everywhere, these levels would be at about 12, 15.5, 19,..., 47km. Danielson (1959) pointed out that interpolating geopotential and temperature separately to isentropic levels violates the hydrostatic relation, leading to large errors in any subsequent analysis. He therefore suggested using the hydrostatic relation, which in our coordinates is

$$M_z = c_pT$$

to build up M from a base level. We therefore evaluated M on the 850mb level, and the entropy on every pressure level in a column, and used the above relation to find M on every pressure level. We then found M on the chosen isentropic surfaces using first-order interpolation. Cubic spline interpolation was found to lead to problems since near the tropopause where there was a rapid change in isentropic density.

In order to estimate the horizontal winds we used an approximation similar to Newman et al.'s(1988) (which was in turn based on Randel's (1987) work) in which they neglected vertical advection and time derivatives and used the geostrophic winds to estimate the non-linear terms in the momentum equations. Randel used a higher-order approximation than the geostrophic one in a partially successful attempt to reduce the size of the regions of positive $\nabla.F$ found on using geostrophic winds. However, Li (1991) has investigated several higher-order approximations and found that, while they are an improvement over the geostrophic approximation when the Rossby number is small, they all are in considerable error in regions where the Rossby number is large. Our method differs from those of the above studies in being carried out on isentropic rather than isobaric surfaces, using a polar-stereographic rather than a lat-long grid and in the treatment of the non-linear terms due to spherical geometry. Our equations were

$$-V\left(f - \frac{xV^g - yU^g}{2a^2}\right) = -m(U^gU^g_x + V^gU^g_y) - V^gf$$

and

$$U\left(f - \frac{xV^g - yU^g}{2a^2}\right) = -m(U^gV^g_x + V^gV^g_y) + U^gf$$

where

$$m = \frac{2}{1 + \sin\phi}$$

$$x = a\,m\cos\phi\cos\lambda, \quad y = a\,m\cos\phi\sin\lambda$$

$$U = -u\sin\lambda - v\cos\lambda, \quad V = u\cos\lambda - v\sin\lambda$$

$$U^g = \frac{-m\,M_y}{f}, \quad V^g = \frac{-m\,M_x}{f}$$

and λ is longitude. The relative vorticity was then calculated on the polar stereographic grid, using the definition

$$\zeta = (mV)_x - (mU)_y$$

and divided by the isentropic density to give the PV, which was then interpolated back onto the lat-long grid, along with v.

In figure 2.4 we show the derived $\bar{u}$ fields for the mid-season months. Yang et al.

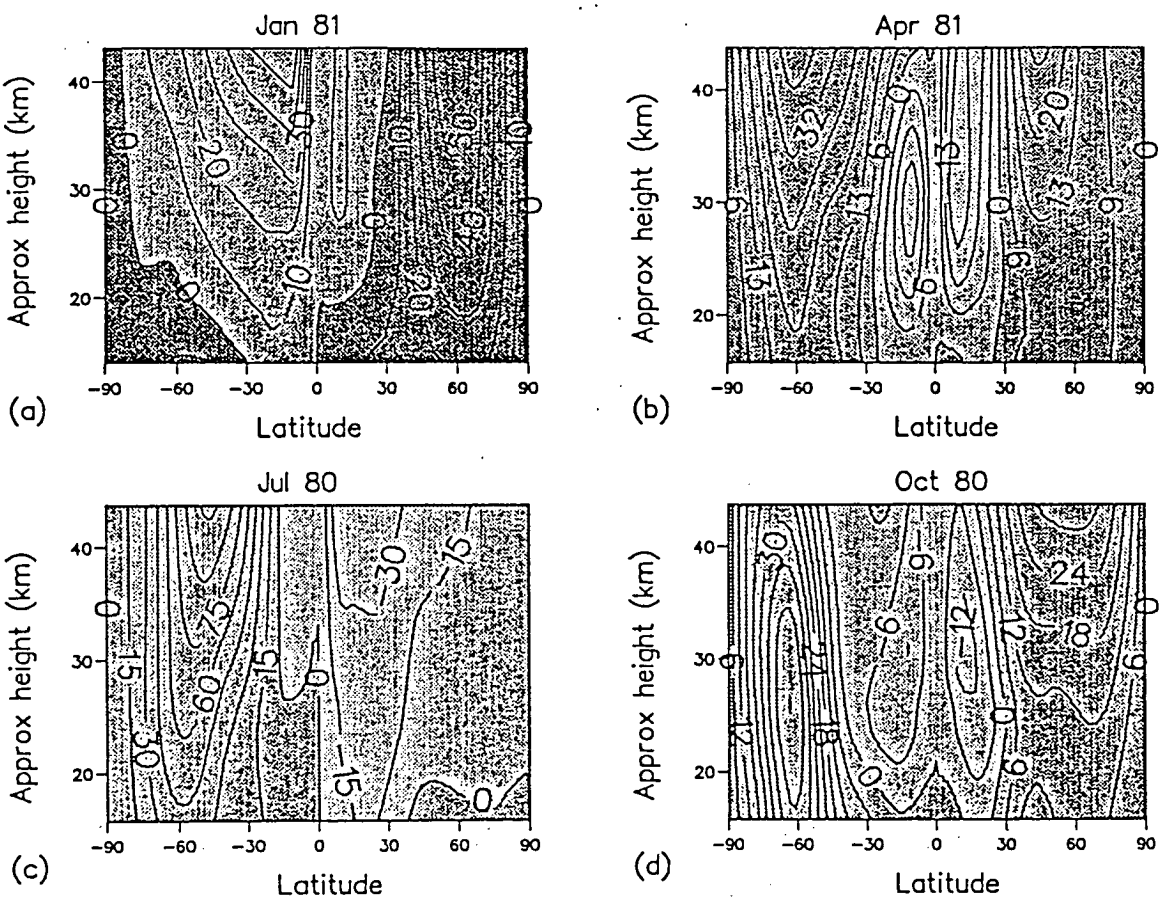


Figure 2.4: $\bar{u}$ *for (a) Jan 81 (b) Apr 81 (c) Jul 80 (d) Oct 80*

(1990) have calculated similar fields on isentropic surfaces from NMC data and our results are very similar to theirs, though theirs are continued above and below ours (see figure 2.5).

As a check on our PV calculations we show polar-stereographic plots of PV for Feb 1st 1981 on the 850K isentropic surface in the NH in figure 2.6. O'Neill and Pope (1987) studied the NH 80/81 winter and used geostrophic winds derived from

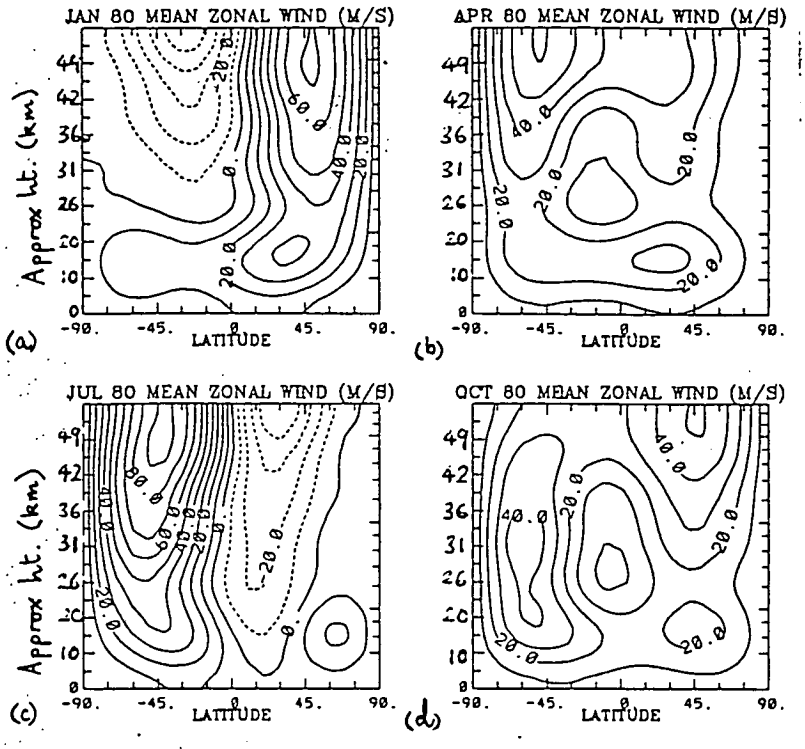


Figure 2.5: $\bar{u}$ *derived by Yang et al. (1990) for (a) Jan 80, (b) Apr 80, (c) Jul 80 and (d) Oct 80.*

SSU data to calculate PV on isentropic surfaces. Our calculation using geostrophic winds is very similar. We will not compare values since we used a log-theta vertical coordinate and defined our density and hence PV accordingly, while they used theta as their vertical coordinate. Our plots are scaled so that the maximum value is 11. O'Neill and Pope calculated geostrophic winds (from geopotentials which were truncated at wavenumber 12 in the zonal and meridional directions - see Clough et al. 1985) and vorticity on pressure surfaces and then interpolated to isentropic surfaces. This could account for some of the discrepancy, but on the whole the patterns are similar enough to give us confidence in our calculations. The PV values used in this thesis were not, however, calculated from geostrophic winds, but from the 'corrected' winds described above. The PV field calculated using these winds is then also shown in figure 2.6. Due to the scaling the values appear to be the same, but in fact the corrected winds lead to a maximum PV value which is about 10% smaller than that of the geostrophic PV. It also appears to lead to a smoother field. The overestimation of PV using geostrophic winds was noted by Clough et al. (1985). The spurious peaks south of about 20N are due to the interpolation of Montgomery potential onto the isentropic surfaces and the sensitivity of the geostrophic wind to this when the Coriolis parameter is small. Values equatorwards of 20N and 20S were afterwards extrapolated from 20N and

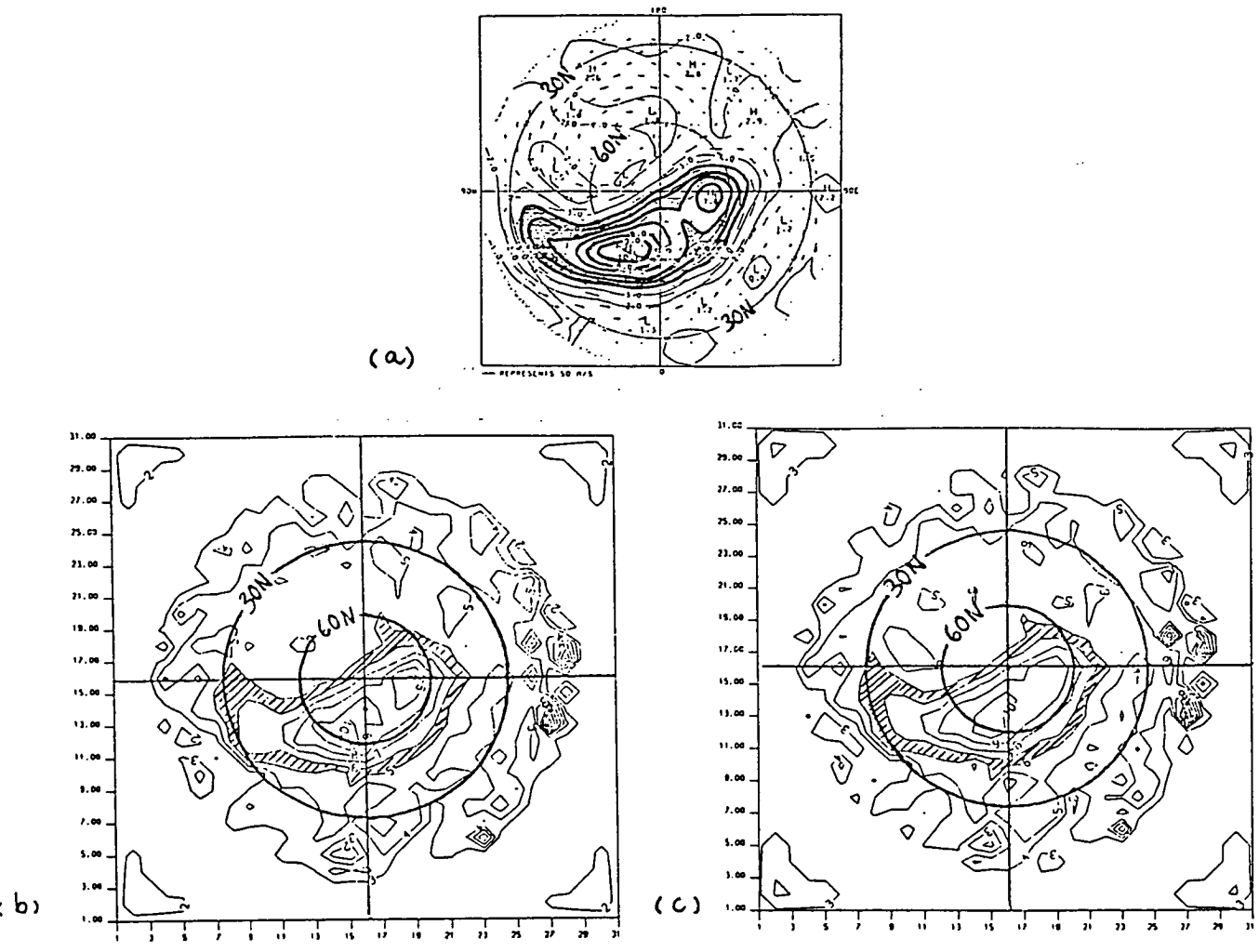


Figure 2.6: *PV on 1st Feb 81 in NH on 850K surface (a) from O'Neill and Pope (1987) (b) our calculations using geostrophic winds (c) using 'corrected' winds.*

20S to become zero over the equator.

These values of winds and PV will be used in Chapter 3 to calculate the 'symmetric vortex' state, in Chapter 4 when investigating inter-hemispheric differences, in Chapter 7 to calculate the eddy flux of PV, and in Chapter 8 to estimate the part of the PV flux due to dissipation of PV.

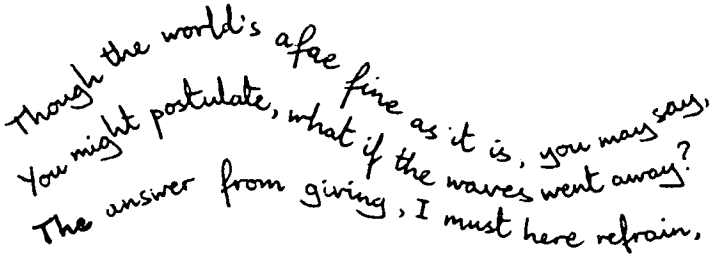
2.7 Summary

The basic dynamical equations have been introduced, along with a description of Ertel's potential vorticity (PV) and some of its properties. The work of Tung (1986) was referred to, in which he estimates that the horizontal eddy flux of PV

is the main part of the wave-driving of $\overline{u}$. The three main ways (treating planetary wave breaking as a separate case) in which planetary waves are believed to produce a horizontal flux of PV are described, all of which should produce a negative (or zero) flux in the long-term average (over about a month). Finally, the method of using satellite geopotential height data to estimate the horizontal winds was described. These winds will be used throughout the thesis in various methods of estimating the PV flux.

Chapter 3

A Symmetric Vortex



yet it, like the hills, will at last be made plain.

3.1 Introduction

This chapter will describe the way in which a zonally-symmetric state was associated with any 3-dimensional atmospheric state by deforming the contours of PV on an isentropic surface into latitude circles, while conserving the mass and the PV of each parcel. This zonally-symmetric state will be shown, in chapter 4, to reveal the effect of a stratospheric $\mathcal{F}$ more clearly than the zonally-averaged state. It will also be used in chapter 8 as the reference state used to define large parcel displacements.

The motivation for this work came from two directions. In a paper on stratospheric sudden warmings, McIntyre (1982) noted that, because the vortex is highly distorted during sudden warmings, the strong vorticity gradients on the edge of the vortex are blurred in the Eulerian-mean picture. As a way of revealing more clearly these vorticity gradients, which influence the propogation of planetary waves, he suggested associating with the distorted vortex a zonally symmertic state which would be constructed by deforming the contours of quasi-geostrophic potential vorticity on each isobaric surface back onto latitude circles while preserving the area enclosed by each contour (in effect conserving the quasi-geostrophic potential

vorticity of each parcel). Because of the invertibility principle (see for example

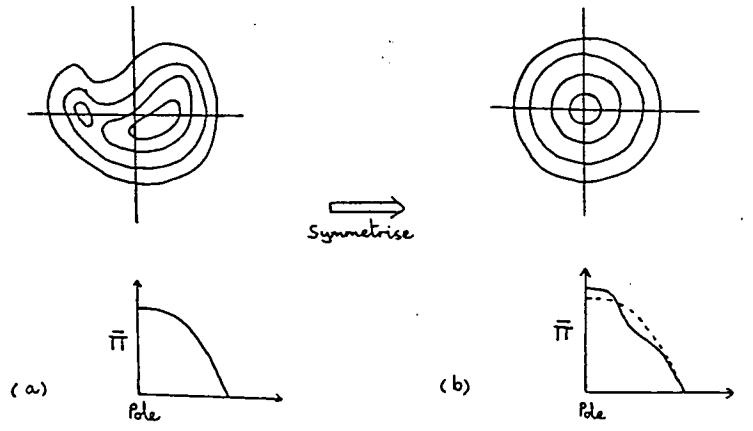


Figure 3.1: *PV on an isentropic surface (above) and zonally-averaged PV on that surface (below) for (a) the distorted and (b) the symmetric vortex.*

Hoskins et al. 1985) it would then be possible, with suitable boundary conditions and a balance equation (such as thermal wind balance) to find $\overline{u}$ and $\overline{T}$ for this 'resymmetrised' state. Discussion of wave-mean flow interactions, he suggested, would then make more sense if this were taken as the mean state. In Chapter 4 of this thesis we will see how the effect of $\mathcal{F}$ appears more clearly in the evolution of the symmetric vortex state, similar to McIntyre's one, to be described in this chapter.

The other incentive arose out of the need to parametrise eddy fluxes, and the estimation of the diffusion coefficient K_{yy} (see Chapter 1). In isentropic coordinates we can assume that K_{yy} is the dominant element of the diffusion matrix (see Tung 1987). This enables us to write (see Tung 1986)

$$\overline{v^*\rho\Pi^*} \simeq -K_{yy}\hat{\Pi}_y$$

(where the 'hat' denotes the density-weighted zonal mean -see p30)

and hence estimate K_{yy} from the observed flux and gradient of PV. However, the meridional gradient of $\hat{\Pi}$, like the gradient of quasi-geostrophic potential vorticity, though generally positive, sometimes changes sign and so leads to undefined, or negative K_{yy} (see Newmann et al. 1987, who estimated K_{yy} in a similar way). A negative $\hat{\Pi}_y$ may actually occur in the real atmosphere, but a negative K_{yy} is unacceptable for use in 2D models since an up-gradient eddy flux, which tends to increase the mean gradient and which increases as the gradient increases, would obviously lead to instabilities. Since distortion of the vortex can cause such a reversal of the gradient a solution to this problem is to use a symmetric vortex state, similar to the one described above, in place of the zonal-mean state, defined

so that its meridional gradient is always positive. However, K_{yy} was not actually calculated in this thesis, the main reason being that, because of the work reported in Chapter 8, where the PV flux is estimated to be due mainly to dissipation of PV, a K_{yy} suitable for parametrising PV flux using the above relation would not be suitable for parametrising the eddy flux of a quantity with a different dissipative time constant (see Chapter 1). The symmetric vortex was, however, used in Chapter 8 to calculate displacements used in estimating the dissipative PV flux.

The resymmetrised state described in this chapter is similar to McIntyre's but is defined in terms of Ertel's PV on isentropic surfaces. This frees us from the constraints of quasi-geostrophic theory but introduces the complication that, in resymmetrising the distorted vortex, we cannot assume that the area enclosed by every contour of PV is preserved. This is because the density in isentropic coordinates, unlike that in isobaric coordinates, is not constant on a horizontal coordinate surface. We must allow the density of the parcel, and hence its area, (so that its mass will be conserved) to change as we move it to its position in the symmetric vortex. Combining the thermal wind equation and the definition of PV we arrive at a 2nd order PDE, whose dependent variable is the pressure of the isentropic levels, which we can then solve after imposing suitable boundary conditions in the troposphere, mesosphere, tropics and at the poles, and taking into account the mass conservation just mentioned.

Another possible use of the symmetric vortex is in removing the effect that reversible displacements of the vortex due to free travelling Rossby waves have on the evolution of the monthly-mean zonal-mean state, so that the symmetric vortex evolves only under the influence of irreversible fluxes of PV and the diabatic circulation. An attempt is made in Chapter 4 to estimate these irreversible PV fluxes from the evolution of the symmetric vortex.

The effect of vacillations on the monthly-mean zonal-mean state will be explained in the next section. We will then define the symmetric vortex, describe the method of calculating it and compare it with the zonal-mean state.

3.2 The effect of vacillations and free Rossby waves

Travelling Rossby waves, by introducing zonal asymmetries into the PV field, are likely to reduce the monthly-mean $\overline{u}$ below the value it would have in the absence of these waves. (A displacement of the vortex from the pole is likely to reduce the value of PV near the pole, and hence also reduce $\overline{u}$). In the absence of a stationary wave, travelling Rossby waves will not induce a PV flux, provided their amplitude is steady, neither will they show up in the monthly-mean geopotential field if their period of rotation is less than half a month or so. If a stationary wave is present, travelling waves will cause the vortex to vacillate about the pole (see section 2.5.1). The negative flux of PV as the vortex moves away from the pole will be accompanied by a deceleration of $\overline{u}$, while the return of the vortex to the pole will be marked by a positive PV flux and an acceleration of $\overline{u}$. Again, the net PV flux is thus likely to be small, while the effect on $\overline{u}$ is to reduce its monthly-mean value below what it would be in the absence of travelling waves.

Travelling waves differ from stationary waves in that while stationary waves cannot propogate upwards into summer easterlies or strong westerlies, travelling waves appear to be fairly indifferent to $\overline{u}$, and even seem to be slightly stronger in the summer than in the winter (Rodgers (1976) and Geisler and Dickinson (1976)).

This has the following implication for 2D models. If the amplitude of travelling waves is fairly constant throughout the year and they produce little PV flux there would be little change in the PV flux if the travelling waves were to be absent. Hovever, these waves reduce the monthly-mean $\overline{u}$ below the value it would have in their absence. There is therefore little reason to expect a 2D model driven by the observed PV flux to reproduce the observed monthly-mean $\overline{u}$ if the travelling waves introduce sufficient asymmetry in the vortex to significantly reduce the monthly-mean $\overline{u}$. A 2D model driven by an accurate estimate of the real PV flux should therefore be closer to the monthly-mean symmetric vortex, which is not so affected by travelling waves, than to the monthly-mean zonal-mean state.

If, in re-symmetrising the vortex, the PV, the mass and the potential temperature of every air parcel are conserved, then the PV distribution of the symmetric vortex will evolve only as a result of processes which change the PV of an air parcel, or its potential temperature, or both. These processes are frictional dissipation and diabatic heating. The zonal wind of this symmetric vortex, then, evolves solely

under the influence of the monthly-mean circulation and irreversible PV fluxes, and so reflects more clearly than the monthly-mean zonal wind the effect of wave-breaking and dissipation. In this respect it is similar to the area diagnostic used by Butchart and Remsburg (1986), but does not rely on the assumption that the area of an air parcel is conserved by adiabatic processes.

3.3 Defining the symmetric vortex

The constraints placed on the symmetric vortex in order to define it uniquely are as follows:

(1) It was said above that resymmetrising the vortex should conserve the PV, potential temperature and mass of every air parcel. From this can be deduced the more useful statement that the mass of air between two isentropic levels whose PV exceeds a certain value Π_o (where Π_o can take any value) is the same for the symmetric vortex as it is for the disturbed vortex.

(2) The PV values in the symmetric vortex should decrease monotonically from north pole to south pole.

(3) The temperature, $\tilde{T}$, and zonal-wind, $\tilde{u}$, of the symmetric vortex should satisfy the thermal wind equation

$$(\tilde{u} + a\,\Omega\cos\phi)\,\tilde{u}_z = F\,\tilde{T}_\phi \quad , \text{ where } F = -\frac{c_p}{2\tan\phi} \tag{3.1}$$

This equation is derived from equation 2.4, noting that $\tilde{u}$ is zonally-symmetric, by ignoring the material derivative of v and finally taking the vertical derivative. It is also the thermal wind relation used in many 2D models (see for example Tung 1986 and Chapter 5).

(4) Equatorwards of 15N and 15S, and above and below the isentropic levels which lie at about 45 and 10km (z=0.576 and z=1.872), we shall not solve for the symmetric vortex (since PV values calculated from winds estimated from Montgomery potentials cannot be considered reliable near the equator where geostrophic scaling is invalid and because our data does not extend much above 45km, while the isentropic level which lies at about 10km near the pole intersects the ground near the equator). On these boundaries we set the pressure of the symmetric vortex equal to the zonal-mean of the pressure of the disturbed vortex ($\tilde{p} = \overline{p}$). At the poles, we see from 3.1 that $\tilde{T}_\phi = 0$ (since $\tilde{u}$ must be zero at the poles) and hence $\tilde{p}_\phi = 0$ also (recalling the definition of potential temperature and remembering that the meridional derivative is along an isentropic surface). This boundary condition is applied at the poles (which is preferable to setting $\tilde{p} = \overline{p}$ there since it is near the

poles that we would expect the differences between the symmetric vortex and the disturbed states to be greatest). Thus, either $\tilde{p}$ or $\tilde{p}_\phi$ has been constrained around the boundary of the two independent regions where the 2nd order PDE for $\tilde{p}$ will be solved. This is sufficient to determine $\tilde{p}$, provided the PDE is elliptic.

3.4 Finding the symmetric vortex

The method by which $\tilde{u}$ and $\tilde{T}$ are found will now be described. It is well known (see for example Hoskins et al. 1985) that, given only the PV distribution, a balance condition (such as thermal wind balance) and suitable boundary conditions, it is possible to determine the winds and temperatures uniquely. Our situation is not quite as simple as this, since until the pressures of the isentropic surfaces of the symmetric vortex are known its PV distribution cannot be determined (see the following paragraph). The way out of this impasse is to iterate towards a solution, first estimating the PV distribution, $\tilde{\Pi}_e$, then finding the pressure distribution, $\tilde{p}_e$, which maintains thermal wind balance, and using this pressure distribution to re-estimate the PV distribution, and so on. There are actually two ways in which this procedure could fail to reach a solution. The first occurs if the iterative procedure above fails to converge, and the second occurs if the thermal wind equation (or rather the 2nd order PDE which we shall derive from it) cannot be solved. In practice, a solution was reached for each of the days investigated. Both of the steps mentioned above (estimating PV from pressure; solving the thermal wind equation for pressure) require further explanation, and this will now be given.

In order to see how the PV distribution depends on the pressure of the isentropic levels, we must first use condition (1) above. It allows us to calculate, for the disturbed vortex, a function $M(\Pi_o, z_1, z_2)$ which is the mass of air between the isentropic levels z_1 and z_2 whose PV exceeds Π_o, and it then says that this function, calculated using the disturbed vortex, also applies to the symmetric vortex. Using condition (2) we can then further interpret $M(\Pi_o, z_1, z_2)$ as the mass of air between the levels z_1 and z_2 in the symmetric vortex which lies north of the latitude at which $\tilde{\Pi} = \Pi_o$. Now, using the value of $\tilde{p}_e$ (the zonally-symmetric current estimate of $\tilde{p}$) along the isentropic levels we can work out the mass of air between levels z_1 and z_2 which lies north of a certain latitude, and using the *inverse* of $M(\Pi_o, z_1, z_2)$ we can find $\tilde{\Pi}_e$ at that latitude, and thus at every latitude. Using this $\tilde{\Pi}_e$ the thermal wind equation is then solved to find $\tilde{p}_e$ for the next iteration. We arrive at the 2nd order PDE in the following manner. Starting from (3.1) and after re-arranging

and taking the derivative with respect to ϕ we have

$$\{\tilde{u}_z \cos\phi\}_\phi = \left(\frac{F \tilde{T}_\phi \cos\phi}{\tilde{u} + a\,\Omega \cos\phi} \right)_\phi \tag{3.2}$$

From the definition of PV we have

$$\tilde{\rho}\,\tilde{\Pi} = -\frac{(\tilde{u}\cos\phi)_\phi}{a\cos\phi} \tag{3.3}$$

so that combining the vertical derivative of (3.3) with (3.2) we have

$$(\tilde{\rho}\tilde{\Pi})_z = \frac{-1}{a\cos\phi} \left(\frac{F \tilde{T}_\phi \cos\phi}{\tilde{u} + a\,\Omega \cos\phi} \right)_\phi \tag{3.4}$$

Since

$$\tilde{T} = \theta \left(\frac{\tilde{p}}{p_o} \right)^\kappa \quad \text{, and} \quad \tilde{\rho} = -\frac{\tilde{p}_z}{g}$$

we reach finally our 2nd order PDE for $\tilde{p}$

$$\{\tilde{p}_z \tilde{\Pi}\}_z + G \left(\frac{\tilde{p}_\phi^\kappa \cos\phi}{\tan\phi(\tilde{u} + a\,\Omega\cos\phi)} \right)_\phi = 0 \quad \text{where } G = \frac{g\,c_p\,\theta}{2\,a\,p_o^\kappa \cos\phi} \tag{3.5}$$

Eq. 3.5 is solved by writing it in finite difference form on the grid shown in figure 3.2. The boxes are $\kappa/2$ units high (which is 3.5 km in a 239K isothermal atmosphere - but see appendix) and 5° wide. p is held at their corners and π at the middle of their edges. The method of successive over-relaxation

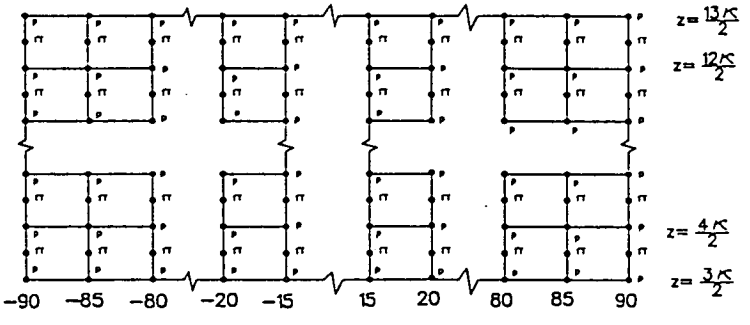


Figure 3.2: *Grid used for solving PDE*

(see Haltiner and Williams 1976) is then used to solve for $\tilde{p}$. Whether equation 3.5 will converge or not is difficult to determine - if (3.5) is not elliptic then it certainly will not - but it is enough to remark that for every day analysed it did. Actually, (3.5) will be elliptic if $\tilde{\Pi}\tilde{\tau}\sin\phi > 0$, which is the same ellipticity condition as found for the stream function of the isentropic model of Chapter 5. It will therefore be satisfied by a statically and inertially stable atmosphere.

$\bar{p}$ is used as the first estimate of $\tilde{p}_e$ from which $\tilde{\Pi}_e$ can be found, along with the function M. The two steps in the iterative process are
(1) Find $\tilde{\Pi}_e$ from $\tilde{p}_e$ and M
(2) Update $\tilde{p}_e$ with one successive over-relaxation iteration. (Only one iteration is used because of the outer iterative loop).
Iteration is continued until $\tilde{p}_e$ is in close enough balance (using a criterion explained below) with $\tilde{\Pi}_e$. This takes between about 40 and 100 iterations, and by that time $\tilde{\Pi}_e$ is very close to $\tilde{\Pi}$. As a check, another iteration is performed, so that $\tilde{\Pi}_e$ is updated once more, and if $\tilde{p}_e$ is still in balance with $\tilde{\Pi}_e$ then we stop iterating since we can then be sure that $\tilde{\Pi}$ and $\tilde{p}$ have been found. If not, the iterations are continued.

The criterion used to end the iterations will now be described. Let ϵ be the left-hand side of 3.5 when $\tilde{\Pi}_e$ and $\tilde{p}_e$ are inserted in place of $\tilde{\Pi}$ and $\tilde{p}$. The smallness of ϵ is a measure of closeness of the balance between $\tilde{p}_e$ and $\tilde{\Pi}_e$. We iterate until ϵ is smaller than $0.02/(a\Delta z\Delta\phi)$ everywhere, where Δz and $\Delta\phi$ are the grid spacings. From equations 3.5 and 3.3 this implies that, to a first approximation, the horizontal and vertical differences of the zonal wind field across a grid box differ from those of that zonal wind field which would balance the pressure field, by about 0.02 m/s. Since the grid is 9 boxes high and 15 boxes wide in each hemisphere, the zonal wind can be, at most, about $(9+15)\times 0.02 \simeq 0.5$m/s out of balance.

3.4.1 Test of the numerical method

Before analysing our results we will describe a test on the method of finding $\tilde{u}$. If we start with a zonally-symmetric geopotential field we might expect $\tilde{u}$ and $\bar{u}$ to be identical, since the u and PV fields before they are 'symmetrised' should also be zonally symmetric and, as noted in Section 3.3, the thermal wind equation satisfied by $\tilde{u}$ is the same as that for a zonally-symmetric u. In this case, the material derivative of v in eq.(2.4) is zero since a zonally symmetric Montgomery potential leads to a zonally symmetric u and a zero v field, even after the non-geostrophic correction to v described in section 2.6. Equation (2.4) is therefore the same as eq.(3.1), in this case. Figure 3.3 shows ($\tilde{u}$-$\bar{u}$) and ($\tilde{T}$-$\overline{T}$) for the zonally-averaged geopotential field of Dec 15 (ie: the initial longitudinally-varying geopotential field was replaced by its zonal average). We see that the difference between $\tilde{u}$ and $\bar{u}$ is not in fact zero, but exceeds 6 m/s at about 20S. Comparing

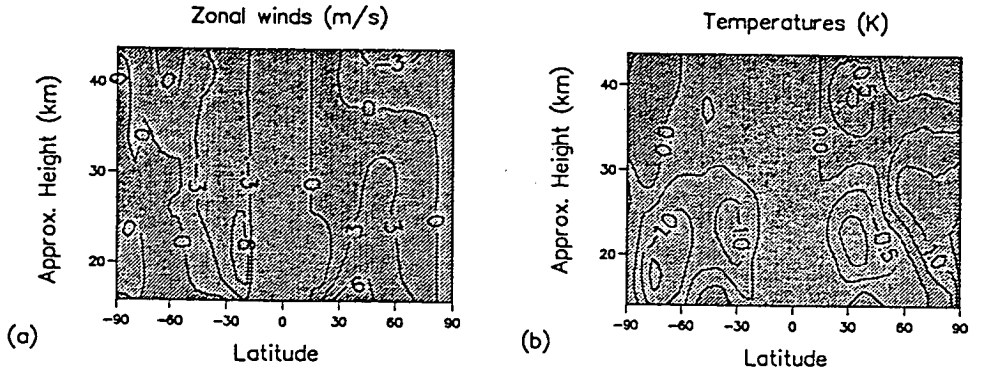


Figure 3.3: *(a)* $\tilde{u}$-$\overline{u}$ *and (b)* ($\tilde{T}$-$\overline{T}$) *for a zonally-symmetric geopotential (set to zero outside the regions (see Section 3.3) where* $\tilde{u}$ *is found).*

fig. 3.3 with the $\overline{u}$ field for December (which is similar to January, shown in fig 2.4(a)) we see that there appears to be some correlation between large $\tilde{u}$-$\overline{u}$ and strong zonal winds, but that the sign of $\tilde{u}$-$\overline{u}$ is not correlated with the sign of $\overline{u}$. The difference between $\tilde{T}$ and $\overline{T}$, shown in fig.3.3(b), reaches 1.5K in places and can have either sign. Since, as just noted, there is no theoretical reason for $\tilde{u}$ and $\overline{u}$ to differ, we conclude that the cause of the difference lies in the numerical scheme.

To recap, this involves interpolation of the geopotential heights from the lat-long to the polar stereographic grid. Horizontal winds and PV are then estimated on isentropic surfaces, and the function M calculated, all on the polar stereographic grid. M, and the zonally-averaged pressure on the grid boundary, are then used to find $\tilde{u}$. It seems plausible, then, that working on the polar stereographic grid (which has no symmetry in the zonal direction) can lead to this discrepancy between $\tilde{u}$ and $\overline{u}$, especially in the regions where the meridional gradient of Montgomery potential (and hence $\overline{u}$) is large.

A numerical effect of this magnitude is not big enough, however, to affect the conclusions drawn in the following sections.

3.5 Differences between $\tilde{u}$ and $\overline{u}$, $\tilde{T}$ and $\overline{T}$

3.5.1 The northern hemisphere

The NH winter of 1981 is analysed by O'Neil and Pope (1987) using hemispheric projections of PV and the 'area diagnostic' of Butchart and Remsberg (1986).

They show that the vortex is reasonably symmetric in January and March but very disturbed in February, with an anticyclonic vortex (the Aleutian high) tearing tongues of high-PV air from the main vortex and irreversibly dispersing them towards the equator. They attribute the build up and resymmetrising of the main vortex again during March to radiative effects reducing the Aleutian anticyclone.

If we look at time-latitude projections of $\overline{u}$ and $\tilde{u}$ at about 38km (see Figure 3.4) we see that $\overline{u}$ fluctuates greatly through early February while $\tilde{u}$ changes much more

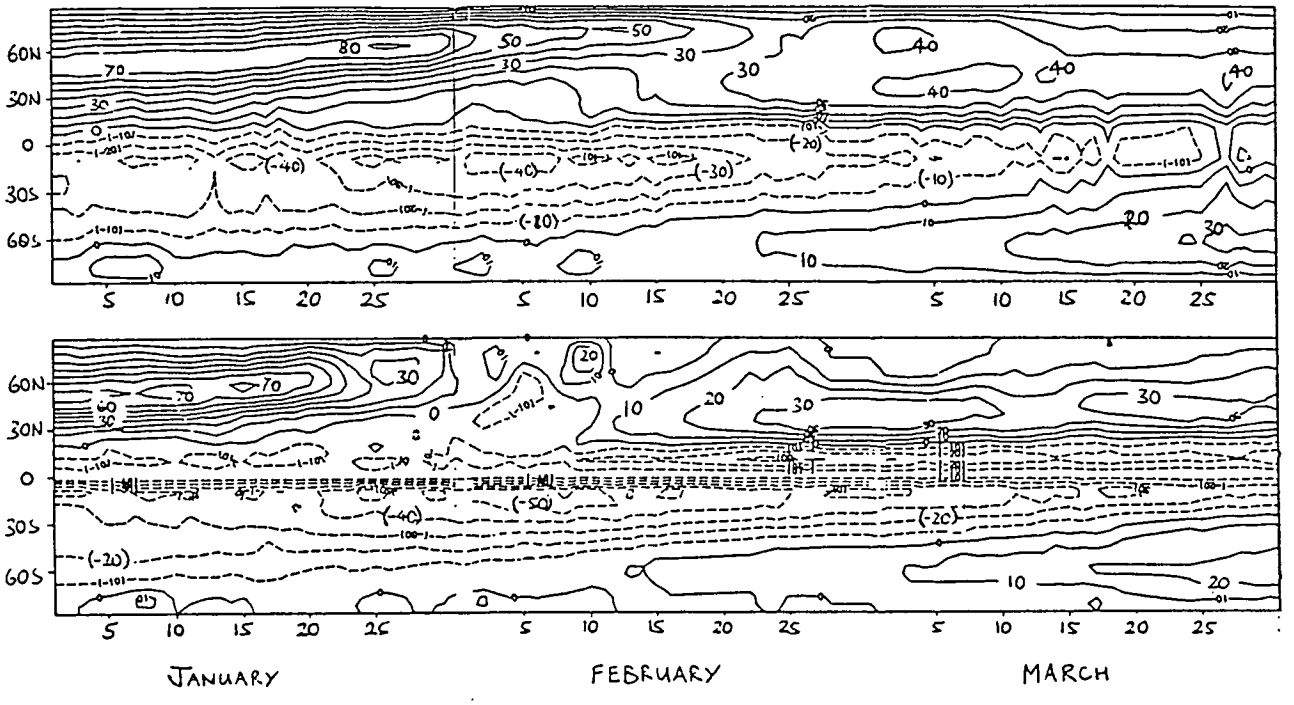


Figure 3.4: $\tilde{u}$ *(top) and* $\overline{u}$ *(bottom) at about 38km from January to March 1981*

steadily. The jet in $\tilde{u}$ moves poleward through late January/early February while in late February the wave breaking noted by O'Neil and Pope has a strong effect on $\tilde{u}$. Note that in early and mid Feb, $\overline{u}$ has large regions where it is easterly although the westerly jet reappears in March, while $\tilde{u}$ is hardly affected, but continues to decrease slowly. This shows the effect a distortion of the polar vortex can have on the zonal winds, disguising the truly irreversible effects of dissipating waves.

We will now consider the monthly-averaged $\tilde{u}$ and $\tilde{T}$ in the NH. From late spring till early autumn, over which period $\overline{u}$ is small (less than 20 m/s everywhere) or easterly, the difference between $\tilde{u}$ and $\overline{u}$ is small, typically less than 5 m/s and similarly ($\tilde{T}$-$\overline{T}$) is small, of the order of 1°C. In figure 3.5 we show some fields for June 1981. We note that the easterlies of $\tilde{u}$ are slightly weaker than those of $\overline{u}$ and do not extend as far down, but still reach to the north pole. We see from

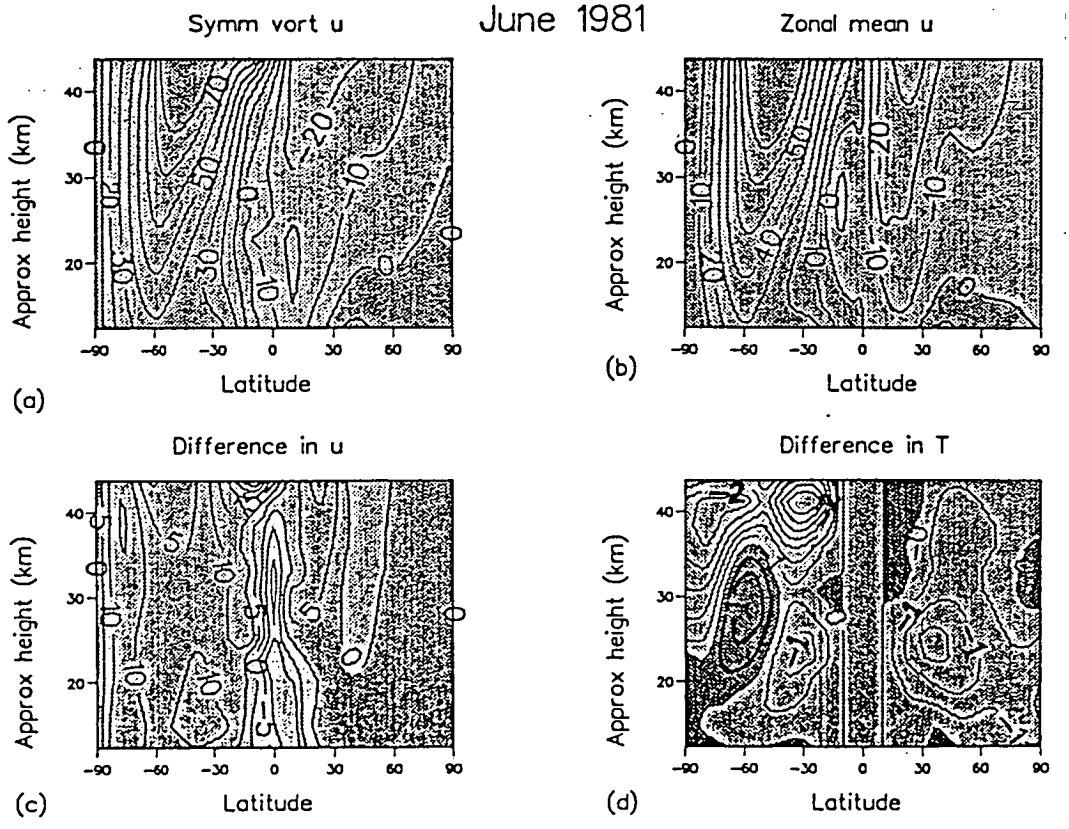


Figure 3.5: *(a)* $\tilde{u}$ *(b)* $\overline{u}$ *(c)* $(\tilde{u}\text{-}\overline{u})$ *and (d)* $(\tilde{T}\text{-}\overline{T})$ *for June 81*

fig.3.5(c) that $\tilde{u}$ is generally greater than $\overline{u}$, but by less than 5 m/s, while from fig.3.5(d) we see that $\tilde{T}$ is slightly colder than $\overline{T}$. The SH will be remarked upon in the next subsection. The differences between the symmetric vortex and the zonal-mean state increase from late autumn to reach a peak in February (see figure 3.6). Here (fig.3.6(a) and (b)) we see that the jet, which has almost disappeared in $\overline{u}$ still remains in $\tilde{u}$, though it is much closer to the pole than it was in January. Corresponding to the large vertical gradient of $(\tilde{u}\text{-}\overline{u})$ there is a large difference between $\tilde{T}$ and $\overline{T}$, with $\tilde{T}$ being over 7 degrees colder than $\overline{T}$. The westerly jet reappears in the $\overline{u}$ field for March (not shown). The peak value in the westerly jet in $\overline{u}$ from late autumn till early spring differs by less than 15m/s from that of $\tilde{u}$ except in January and February, when the jet in $\tilde{u}$ is 20 and 30 m/s stronger than the jet in $\overline{u}$. The westerly jet in $\tilde{u}$ is generally broader than that in $\overline{u}$.

3.5.2 The southern hemisphere

In the SH, the peak in $(\tilde{u}\text{-}\overline{u})$ occurs in December (see figure 3.7). We see that the easterlies in $\overline{u}$ near the pole (which are stronger than their opposites in the NH

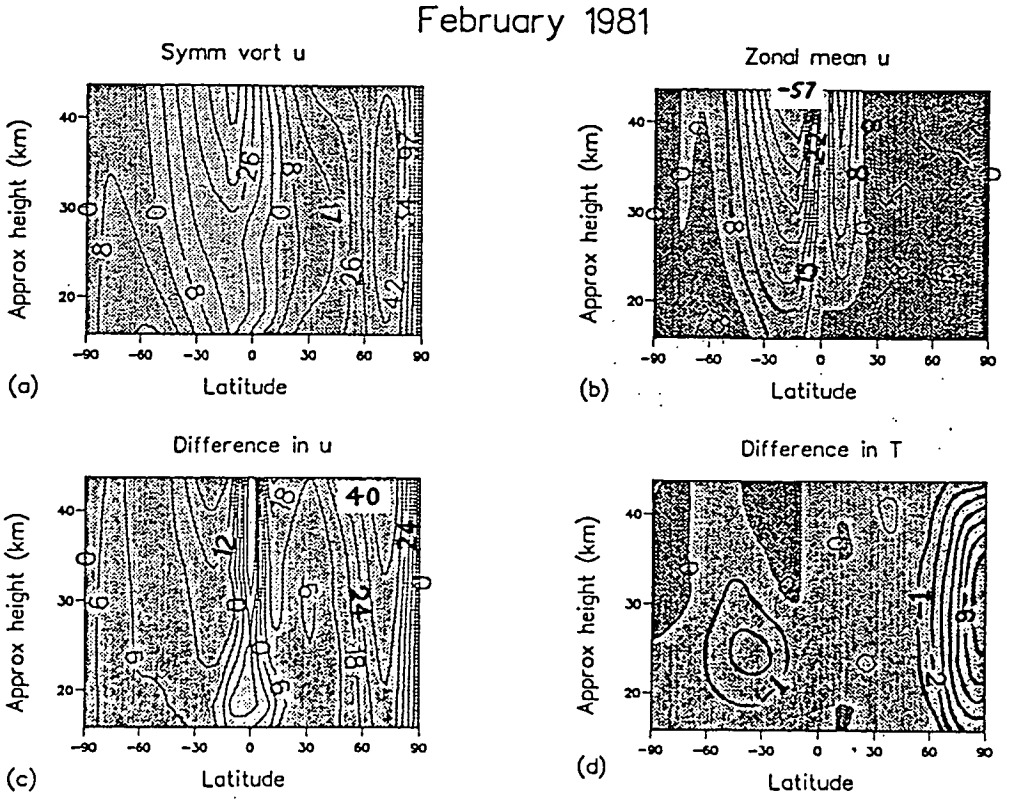


Figure 3.6: *(a)* $\tilde{u}$ *(b)* $\overline{u}$ *(c)* $(\tilde{u}\text{-}\overline{u})$ *and (d)* $(\tilde{T}\text{-}\overline{T})$ *for Feb 81*

in June) have been replaced by westerlies in $\tilde{u}$ and the winter westerly jet is still present in the stratosphere. From (c) we see that $\tilde{u}$ exceeds $\overline{u}$ by over 15 m/s, polewards of about 50S, from the bottom to the top of our observations. It is perhaps worth noting that even though $(\tilde{u}\text{-}\overline{u})$ is large, its vertical gradient is small and hence, from thermal wind balance, it is not accompanied by a large $(\tilde{T}\text{-}\overline{T})$. We note the fact that $(\tilde{u}\text{-}\overline{u})$ is much bigger in the SH in summer than it is in the NH summer. In fact, it is generally bigger in the SH over spring and early summer than in the SH winter, exceeding 15 m/s in Sep, Oct, Nov and Dec. It also exceeds 15 m/s in April, while the peak value over winter never exceeds 10 m/s. From figure 3.5 we see that the winter jet in $\tilde{u}$ is broader than that in $\overline{u}$, although the peak value is virtually the same.

Although $(\tilde{u}\text{-}\overline{u})$ peaks in December, the peak value of $(\tilde{T}\text{-}\overline{T})$ occurs in October, when it is less than -4 °C at the pole, and is similar in shape to that in the NH in February. In general, except during the winter, the differences between $\tilde{u}$ and $\overline{u}$, and $\tilde{T}$ and $\overline{T}$, are greater in the SH than in the NH. A simple test, which involved reflecting the geopotential heights about the equator and checking that the $\tilde{u}$ calculated from these geopotential heights was also exactly reflected, proved

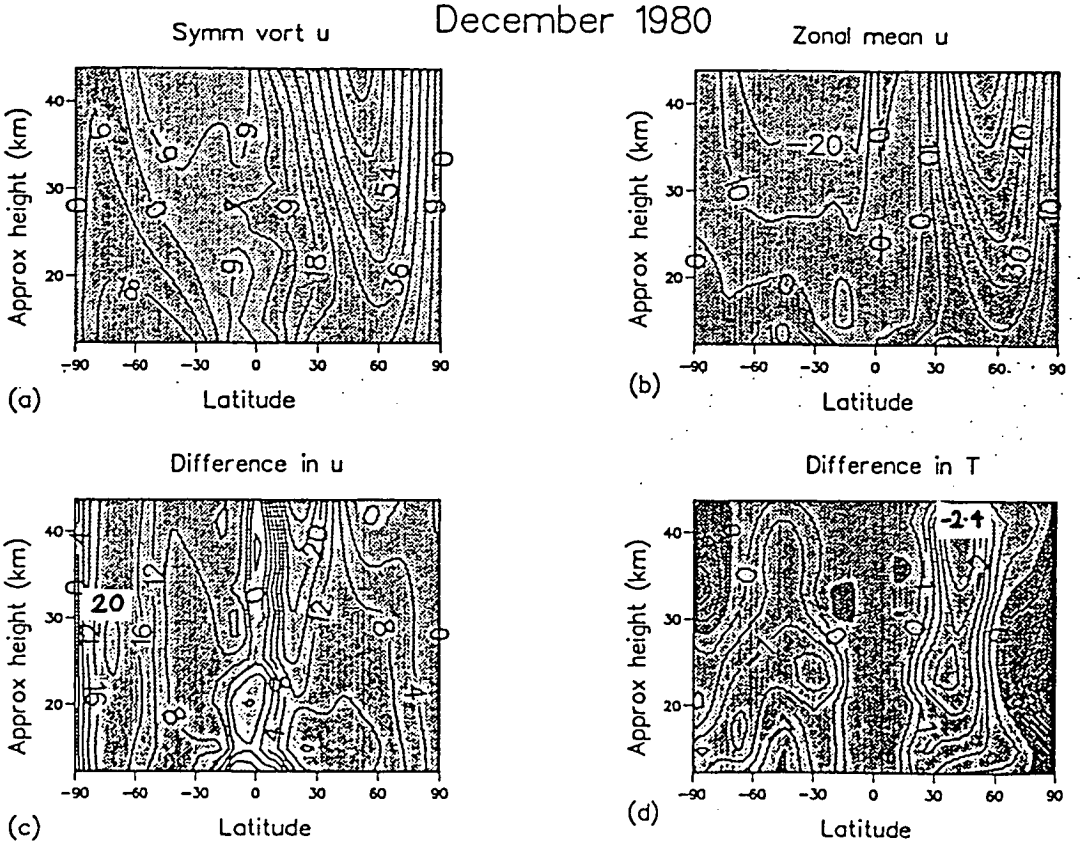


Figure 3.7: *(a)* $\tilde{u}$ *(b)* $\overline{u}$ *(c)* $(\tilde{u}\text{-}\overline{u})$ *and (d)* $(\tilde{T}\text{-}\overline{T})$ *for Dec 80*

that the program did not contain a bug which introduced an inter-hemispheric difference.

3.6 Discussion

In summary of the last section, we can say that the winter westerly jet is more affected by resymmetrisation in the NH than in the SH. However, the summer easterlies of the SH are reduced more (and even become westerly over the pole) than those of the NH, on resymmetrising the vortex.

Because the NH in winter is known to be more disturbed than the SH in winter, or even in early spring, we can understand the first statement made above. However, since we might expect the NH to be generally more disturbed than the SH, we are prompted to inquire why $(\tilde{u}\text{-}\overline{u})$ in the SH summer is so much greater than it is in the NH summer? The answer must be that the NH summer stratosphere is more zonally symmetric than the SH summer stratosphere. Looking at polar-stereographic plots of Montgomery potential on the (850K) isentropic sur-

face (at about 38km) for June 15 in the NH (fig 3.8(a)) and December 15 in the SH (fig 3.8(b)) we see that the SH is indeed more disturbed than the NH. (The plots we will show in this section will all be scaled to have a maximum value of 11, and we will ignore the scaling factors since we will be concerned only with the patterns). As a check that the asymmetries are not due to a fault in our derivation

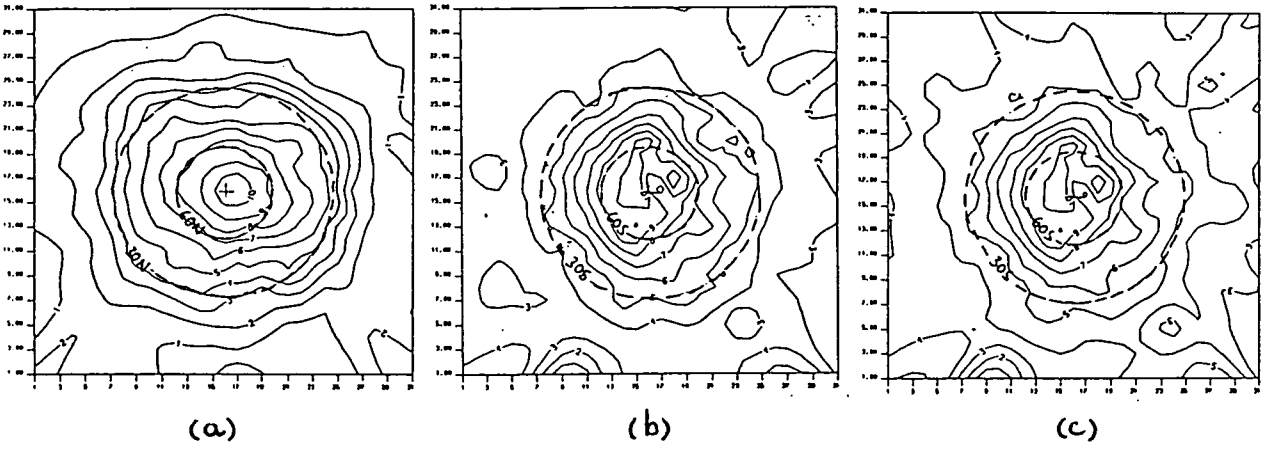


Figure 3.8: *Montgomery potential for (a) NH June 15, 1981 (b) SH Dec 15, 1980 and (c) geopotential for SH Dec 15, 1980.*

of the Montgomery potential from geopotential, we see from fig 3.8(c) that the geopotential on the pressure level at about the same height has an almost identical shape to that of the Montgomery potential. The asymmetries therefore originate in the geopotentials, and so are either real or are introduced in their derivation from the radiances measured by the satellite (see Chapter 7). The asymmetries are present in the SH throughout December, while the NH in June remains fairly symmetric. The work of Hirooka and Hirota (1985) on travelling waves does not provide supporting evidence for this claim that the SH summer is more disturbed than the NH summer, although it does not contradict it either. They found that travelling wave amplitudes were generally larger in the NH than in the SH, and were small in the summer. They were looking only for travelling waves, however, and not just at the asymmetry of the stratosphere.

Another factor which will contribute to zonal asymmetries are spurious values in PV, especially near the equator. In figure 3.9 we show polar-stereographic plots of PV on the 850K surface for the same days as before. Again, the NH seems to be more symmetric, and the highest PV values appear near the pole. However, in the SH the highest PV values are near the equator. These high tropical values are probably due to a failure of the quasi-geostrophic procedure

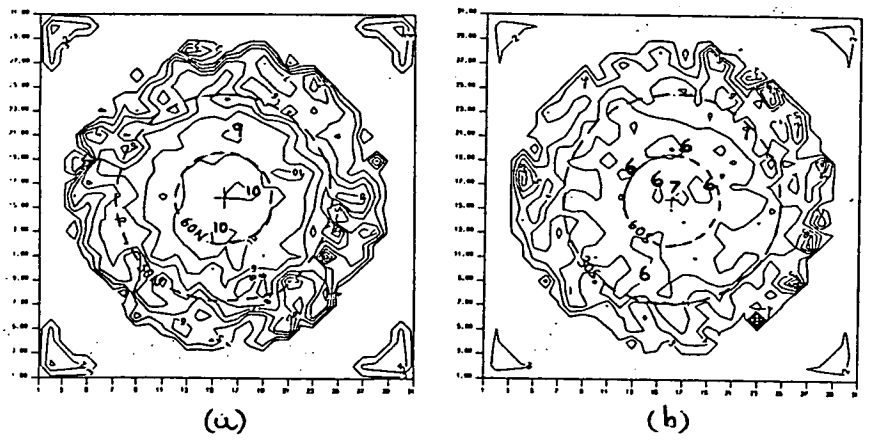


Figure 3.9: *PV for (a) NH June 15, 1981 (b) SH Dec 15, 1980*

used to estimate the winds near the equator, but they will still affect our procedure for finding $\tilde{u}$ since in calculating the function M (see section 3.3) we have included all air parcels polewards of 10 degrees. The effect they have on $\tilde{u}$ was investigated by recalculating $\tilde{u}$ for December but including only those air parcels which are polewards of 30 degrees, thus excluding the anomalous tropical values. The $\tilde{u}$ obtained was negligibly different (less than 5 m/s) from the $\tilde{u}$ of figure 3.7(b). Also, the errors in $\tilde{u}$ described at the start of section 3.5, believed to be due to interpolation between grids, are much smaller than the values in fig.3.7(a) and hence we can reasonably claim that the main reason for the difference between ($\tilde{u}$-$\overline{u}$) in the NH and SH in summer is due to the genuinely greater asymmetry of the SH geopotential heights.

3.7 Summary

In this chapter a 'symmetric vortex' state has been described. It is constructed from the observed atmospheric state by rearranging the air parcels, while conserving each parcel's value of PV, into a zonally-symmetric arrangement with PV values decreasing monotonically from the north pole to the south pole. The requirement that the symmetric state be in thermal wind balance, along with suitable boundary conditions, ensures a unique solution. It is noted that interpolation from a latitude-longitude to a polar stereographic grid introduces spurious zonal asymmetries which can result in uncertainties of typically 5 m/s in the deduced symmetric vortex zonal wind, $\tilde{u}$, and 1K in the temperature, $\tilde{T}$.

This symmetric vortex evolves only as a result of diabatic heating and irreversible wave fluxes of PV, and is not affected by the reversible distortions of the vortex which 'blur' the zonally-averaged state. It should therefore reveal, more clearly than the zonally-averaged state, the effect of irreversible wave fluxes of PV, a point which will be illustrated in Chapter 4. It will find a further application in Chapter 8 when it is used as the zonally-symmetric reference state from which air parcels are assumed to have been displaced. In this way the parcel displacements, needed to estimate the PV flux due to the dissipation of PV in planetary waves, will be estimated.

The symmetric state should be relatively unaffected by travelling Rossby waves and vortex vacillations, as should the monthly-mean PV flux, while the monthly-mean $\overline{u}$ should be reduced. It is therefore argued that the fields of a 2D model fed with a correct estimate of the monthly-mean PV flux should be closer to the symmetric vortex state than the observed monthly-mean state.

The smoother evolution, and the longer persistence, of the westerly jet in $\tilde{u}$, compared with $\overline{u}$, is illustrated for the NH winter when large planetary waves disrupt the polar vortex. The biggest difference between the symmetric vortex and zonally-averaged states occurs during winter in the NH, when the westerly jet in $\tilde{u}$ is over 35 m/s stronger than the one in $\overline{u}$, and $\tilde{T}$ is 7K cooler than $\overline{T}$. In the NH summer, there is little difference between $\tilde{u}$ and $\overline{u}$. In the SH summer, however, $\tilde{u}$ can exceed $\overline{u}$ by over 15 m/s, indicating that the SH summer is more disturbed than the NH summer. This greater asymmetry might be real, or it might be due to errors introduced when deriving the geopotential heights from satellite data (eg. using a base level geopotential height which is inaccurate due to lack of sufficient surface pressure readings in the SH). The biggest difference between the zonally-averaged and symmetric states in the SH occurs from early spring till early summer.

Chapter 4

Estimation of $\mathcal{F}$ from observed Inter-hemispheric differences

Few facts I knew, when I was small - with confidence I'd shout.
But when I grew, I knew it all, my mind was full of doubt.
Now old I've got, forgot a lot, and yet I'm just as muddled.
I would have thought, the *more* forgot, the *less* I'd be befuddled!

4.1 Introduction

Since the planetary-wave forcing of the stratospheric circulation, whether $\nabla.\mathrm{F}$, $\mathcal{F}$ or PV flux, is notoriously difficult to estimate accurately, it is useful to have as many different ways of estimating it as possible so that common features can be identified.

In this chapter a method of estimating $\mathcal{F}$ will be introduced which analyses the inter-hemispheric differences (IHDs) in $\overline{\mathrm{u}}$ and $\overline{\mathrm{T}}$ to estimate the IHD in $\mathcal{F}$. By 'IHD' is meant the difference between one month's data and the data, reflected about the equator, of the month half a year later. An IHD shall be denoted by a superscript 'D'. An estimation of $\mathcal{F}^{\mathrm{D}}$, though not as useful as $\mathcal{F}$ itself, is still valuable since in winter and spring $\mathcal{F}$ is expected to be large in one hemisphere (the northern in winter, the southern in spring) and smaller in the other (by a factor of about 5 - see Chapter 7), so that from $\mathcal{F}^{\mathrm{D}}$ one can obtain a useful estimate of $\mathcal{F}$.

One assumption, that diabatic effects relax $\overline{\mathrm{T}}^{\mathrm{D}}$ towards zero, discussed more fully in the following section, will allow us to estimate $\overline{\mathrm{V}}^{\mathrm{D}}$ from $\overline{\mathrm{T}}^{\mathrm{D}}$, $\overline{\mathrm{T}}^{\mathrm{D}}{}_{t}$ and the equation of continuity. $\mathcal{F}^{\mathrm{D}}$ is then estimated as the residual between $\overline{\mathrm{u}}^{\mathrm{D}}{}_{t}$ and the

forcing of $\overline{u}^D$ by the meridional circulation just estimated. The method is therefore similar to those of previous studies (e.g. Tung and Yang 1988, Shine 1989), the difference being that they estimated the full diabatic circulation, and not its IHD. The uncertainty in calculations of the full diabatic circulation, even using a sophisticated radiation scheme, is large enough to make even the sign of the deduced stratospheric $\mathcal{F}$ uncertain (Shine 1989). The motivation for estimating only the IHD in the heating rate, Q^D, arises from the following consideration. The observed summer $\overline{T}^D$ is small, and therefore it is likely that Q calculated using a scheme such as Shine's will also have a small IHD. However, if there is a systematic error in Q (such as too much cooling in the lower stratosphere), there is also likely to be a systematic error in the estimation of $\mathcal{F}$. Taking inter-hemispheric differences would remove these systematic errors. If the systematic error in Q persists throughout the year then it should affect Q^D, and hence $\mathcal{F}^D$, less than Q and $\mathcal{F}$. Thus in considering only $\mathcal{F}^D$ and not $\mathcal{F}$ some information would be lost but also, hopefully, a large part of the uncertainty would be removed.

The approximation of Q^D by a term proportional to $\overline{T}^D$ might, however, outweigh this advantage, although it does appear to lead to a robust and credible estimation of $\mathcal{F}^D$.

With hindsight, a better procedure, perhaps worth further study, would be to repeat Shine's calculations, but to analyse only the IHD in the deduced $\mathcal{F}$, and not its absolute value. Nevertheless, the work done in this chapter, with its simple heating parametrisation, is a useful first step.

The chapter will end with a description, using isentropic coordinates, of the way in which $\mathcal{F}$ induces changes in the $\overline{u}$, $\overline{T}$ and meridional circulation of an atmosphere in which thermal wind balance is maintained. This description is thus along the lines of that of Eliassen (1951). This mechanism will be used to explain the annual temperature cycles over the equator near the tropopause and the stratopause, observed by previous authors.

It will then be shown how the observed $\overline{u}^D$ and $\overline{T}^D$ reveal the presence of a stratospheric $\mathcal{F}$ in a way that $\overline{u}$ and $\overline{T}$ do not, since in taking the inter-hemispheric difference the effect of the large meridional circulation which is approximately independent of $\mathcal{F}^D$ is effectively cancelled out. $\overline{T}^D$ and $\overline{u}^D$ therefore evolve under the influence of $\mathcal{F}^D$ and the meridional circulation induced by it.

4.2 Deductions from the small summer IHD

In this section some remarks on features of the $\overline{T}^D$ and $\overline{T}^D{}_t$ fields, and some plausible deductions about the processes at work in the stratosphere, will be made. The main deduction, on which hinges the method of estimating $\mathcal{F}^D$, is that there is a process acting throughout the year to reduce $\overline{T}^D$, whose rate is proportional to the size of $\overline{T}^D$.

The validity of this approximation is limited by the fact, noted by Barnett (1974), among others, that the SH summer stratosphere is about 4K warmer than that of the NH summer. This was suggested to be due to the ellipticity of the earth's orbit, the earth being about 3% closer to the sun in January than in July. Therefore , even in the absence of a $\mathcal{F}$, radiative effects are only likely to reduce the magnitude of $\overline{T}^D$ to about 4K.

First, the nomenclature for the IHDs will be explained. Figure 4.1 shows $\overline{T}^D$ for Sep/Mar. This is the monthly-averaged September $\overline{T}$, minus the monthly-averaged $\overline{T}$ for March reflected about the equator. On the left hand side of the figure is the

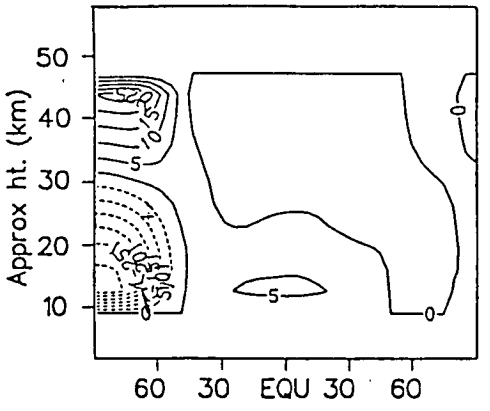


Figure 4.1: *The IHD in* $\overline{T}$ *for Sep/Mar. Dashed contours indicate negative values.*

SH September $\overline{T}$ minus the NH $\overline{T}$ for March. Time-derivative fields will also be referred to, which were calculated as differences between monthly-averaged fields. Thus the $\overline{T}^D{}_t$ field for Sep/Mar is the $\overline{T}^D$ field for Oct/Apr minus $\overline{T}^D$ for Sep/Mar, divided by 30 to give approximately the rate of change per day. The values shown in figures 4.1, 4.2, 4.4, 4.5, 4.9 and 4.10 above about 45km and below about 10km have been set to zero because of lack of data (see Chapter 2). The vertical coordinate used is isentropic and the approximate height scale is described in the appendix.

Looking at figure 4.1, it is seen that $\overline{T}^D$ in autumn (the RHS of the figure) is less than 5K (except near the tropical tropopause) while in spring its peak value exceeds 30K. Graphs of $\overline{T}^D$ throughout the year reveal that $\overline{T}^D$ is less than 5K in amplitude from mid-summer (July NH/Jan SH) until early autumn (Sep/Mar) and only exceeds ±5K (but not ±10K) in small regions from early summer (June/Dec) until late autumn (Nov/May). Over the same seasonal period $\overline{u}^D$ is also small, rarely exceeding ±10 m/s.

Therefore, from the smallness of the IHD over the summer and autumn, despite the large IHD over winter and spring, it seems reasonable to deduce that some process (probably radiative heating) is acting to reduce the IHD in $\overline{T}$ and $\overline{u}$. In the summer it succeeds because of the lack of planetary wave activity but it is assumed here that it acts throughout the year, even in the presence of a stratospheric $\mathcal{F}$. Observational support for this hypothesis comes from figure 4.2 which shows (a) $\overline{T}^D$ and (b) $\overline{T}^D{}_t$ for Nov/May. On the left side of the figures it is late spring, a time

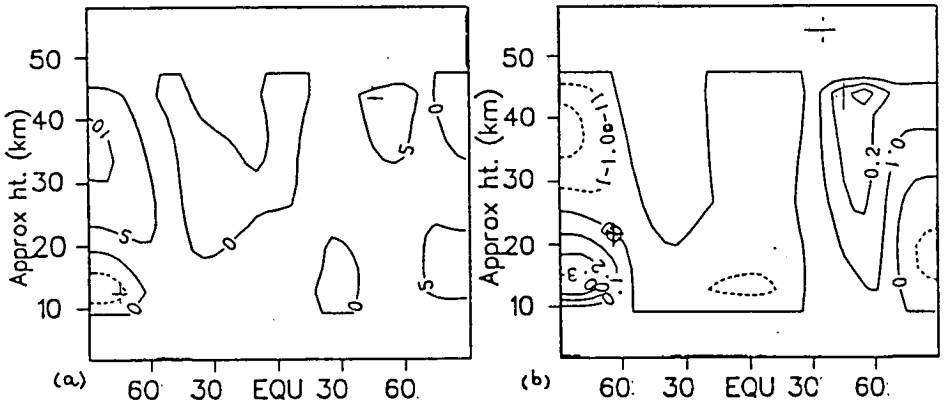


Figure 4.2: *(a)* $\overline{T}^D$ *and (b)* $\overline{T}^D{}_t$ *for Nov/May. Dashed contours indicate negative values.*

when $\mathcal{F}$ is expected to be small in both hemispheres, and yet the effects of the winter and spring $\mathcal{F}$ are still present. It can be seen that there are large regions where $\overline{T}^D{}_t$ and $\overline{T}^D$ have opposite signs.

4.3 The equations and their solution

In this section the two equations which are used to estimate $\mathcal{F}^D$ will be introduced. They are derived from the continuity and momentum equations in isentropic coordinates, but in order to write them in terms of IHDs certain approximations must be made. The simple numerical method of solution will also be described.

The 'difference' continuity equation can be immediately written down since all the terms are linear.

$$\overline{\rho}_t^D \cos\phi + \overline{V}_\phi^D + \overline{W}_z^D = 0 \tag{4.1}$$

Using the assumption made above that there is a process acting to remove the IHD in $\overline{u}$ and $\overline{T}$, and hence the IHD in $\overline{\rho}$, and assuming further that it is solely diabatic, $\overline{W}_z^D$ is parametrised as $-\alpha\overline{\rho}^D\cos\phi$. Equation 4.1 can then be written as

$$(\overline{\rho}_t^D - \alpha\overline{\rho}^D)\cos\phi + \overline{V}_\phi^D = 0 \tag{4.2}$$

The difference between the two momentum equations (a superscript 'a' denoting the 'present' atmosphere and a 'b' denoting the atmosphere six months later, reflected about the equator) can be written

$$\begin{aligned} \overline{\tau}_t^D \cos\phi \equiv (\overline{\tau}_t^a - \overline{\tau}_t^b)\cos\phi = \quad & a\cos^2\phi\mathcal{F}^D \\ & - \left(\frac{\overline{V}^a\overline{\tau}_\phi^a}{\overline{\rho}^a} - \frac{\overline{V}^b\overline{\tau}_\phi^b}{\overline{\rho}^b}\right) \\ & - \left(\frac{\overline{W}^a\overline{\tau}_z^a}{\overline{\rho}^a} - \frac{\overline{W}^b\overline{\tau}_z^b}{\overline{\rho}^b}\right) \end{aligned} \tag{4.3}$$

There are two problems here. The first is the 'non-linearity' of the $\frac{\overline{V}\overline{\tau}_\phi}{\overline{\rho}}$ and $\frac{\overline{W}\overline{\tau}_z}{\overline{\rho}}$ terms, which cannot be expressed entirely in terms of IHDs, and the second is that in order to find $\overline{W}^D$ we would have to integrate $-\alpha\overline{\rho}^D$ (since this is the parametrisation of $\overline{W}_z^D$) vertically with a suitable boundary condition, which is likely to introduce a large error into $\overline{W}^D$. These problems are circumvented by replacing $\overline{\tau}$ on the RHS of 4.3 with the planetary angular momentum (which is constant with height so that $\overline{W}\overline{\tau}_z$ vanishes), and replacing $\overline{\rho}^a$ and $\overline{\rho}^b$ by the average density, $\overline{\rho}^0 \equiv (\overline{\rho}^a + \overline{\rho}^b)/2$. The approximate 'difference' momentum equation can now be written as

$$\overline{u}^D{}_t = \mathcal{F}^D + \frac{\overline{V}^D f a}{\cos\phi\,\overline{\rho}^0} \tag{4.4}$$

The effect of approximating $\overline{\tau}_\phi/\overline{\rho}$ by $-f a\cos\phi/\overline{\rho}^0$ can be estimated from the ratio $X \equiv \left(\frac{-f a\cos\phi}{\overline{\rho}^0}\right) / \left(\frac{\overline{\tau}_\phi}{\overline{\rho}}\right)$. The effect of this approximation will be discussed in the next section.

Equations 4.2 and 4.4 are solved, on a grid with a horizontal spacing of 5° and a vertical spacing of $\kappa/2$ (approximately 3.5 km), in the following way. First, equation 4.2 is integrated, using observed values of $\overline{\rho}^D$ and $\overline{\rho}_t^D$. Integration was started at the poles, thus ensuring that $\overline{V}^D = 0$ at the poles. The value at the

equator turned out to be small, and anyway was not considered to be significant because of the geostrophic approximation used in deriving 4.4. Equation 4.4 was then solved using the $\overline{V}^D$ just derived, and the observed $\overline{u}^D$ and $\overline{\rho}^0$, to find $\mathcal{F}^D$.

4.4 Estimates of $\mathcal{F}^D$ and $\mathcal{F}$, and their sensitivity to uncertainties

Before looking at $\mathcal{F}^D$ it will be helpful to look at figure 4.3 to see when it is reasonable to deduce from $\mathcal{F}^D$ the $\mathcal{F}$ of one hemisphere. That is, taking the wave

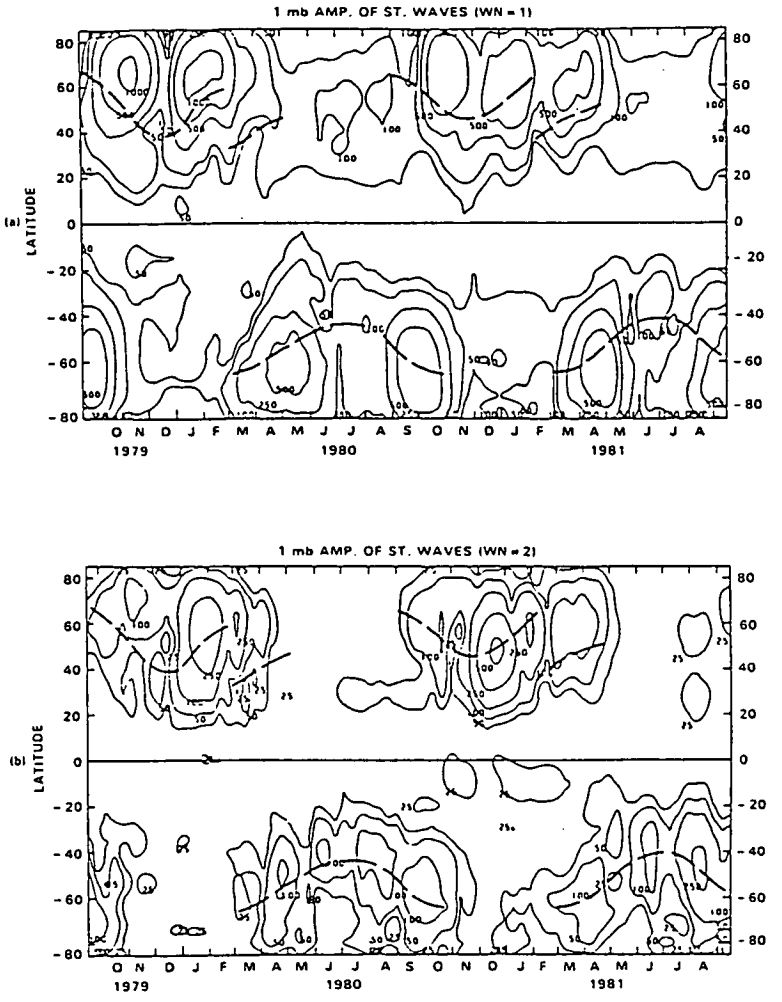


Figure 4.3: *Amplitudes of quasi-stationary wavenumber (a) 1 and (b) 2 waves. (After Hirota et al. 1983)*

amplitude as an indicator of $\mathcal{F}$ (large waves being likely to produce a $\mathcal{F}$ both through dissipation and breaking) we look for the seasons when $\mathcal{F}$ is large in one

hemisphere and small in the other. In the summer and early autumn the waves are small. However, as autumn progresses waves build up in both hemispheres (see figure 4.3). Over winter in the NH the waves persist or intensify while in the SH they die down, especially wavenumber 1 and wavenumber 2 polewards of about 45S. They are prevented from propogating upwards into the stratosphere by the strong westerlies in the SH winter - presumably the SH westerlies are allowed to become more intense than those in the NH because of a weaker $\mathcal{F}$ in the SH in autumn (a positive feedback between waves and mean flow). Therefore, over the winter we expect $\mathcal{F}^{D}$ to reflect the NH $\mathcal{F}$. In spring the change in the solar heating causes the SH westerlies to abate, to the point where planetary waves are again allowed into the stratosphere and in turn help to reduce the westerlies. The NH is more variable from year to year than the SH, and it can be seen from figure 4.3 that in March and April of 1980 the planetary waves were much smaller than in the winter while in March and April of 1981 they were still comparable to their winter values, and in fact similar in size to those in the SH in September and October. Since the data used throughout the thesis spans July 1980 to June 1981 it is not possible to deduce with much confidence the SH $\mathcal{F}$ in spring from $\mathcal{F}^{D}$, but if the previous year had been chosen it should have been more feasible. Nevertheless, in summary we can say that it is possible to estimate from $\mathcal{F}^{D}$ the main occurrences of $\mathcal{F}$ in the NH and (though with less confidence) in the SH, $\mathcal{F}^{D}$ being mainly due to the NH $\mathcal{F}$ in winter and the SH $\mathcal{F}$ in spring.

Figure 4.4 contains $\mathcal{F}^{D}$ from July 1980 to June 1981. There seems to be a distinction between the (generally) large values above about 42km (which we will assume are due mainly to breaking gravity waves) and the smaller ones below (which we will assume are mainly due to planetary waves). Note that the NH winter $\mathcal{F}$ contributes, with its sign reversed, to the left side of fig.(4.4(a)).

The values of $\mathcal{F}^{D}$ over winter are consistent with a negative $\mathcal{F}$ in the NH, and do not suggest any regions of positive $\mathcal{F}$. It is also expected (see Chapter 2) that the PV flux due to a dissipating or breaking planetary wave should be negative in the time-average. This is an important point since positive regions of $\nabla.F$ have appeared in calculations using geostrophic winds, and these were reduced or eliminated by using winds which were a closer approximation to the balance wind (see Randel 1987). For the same reason the calculations presented in Chapter 7 also at times had positive regions. In Shine's (1989) study significant positive regions appeared in the stratosphere though, as already noted, he put little confidence in his results in the stratosphere. In daily values of $\mathcal{F}$ positive regions almost

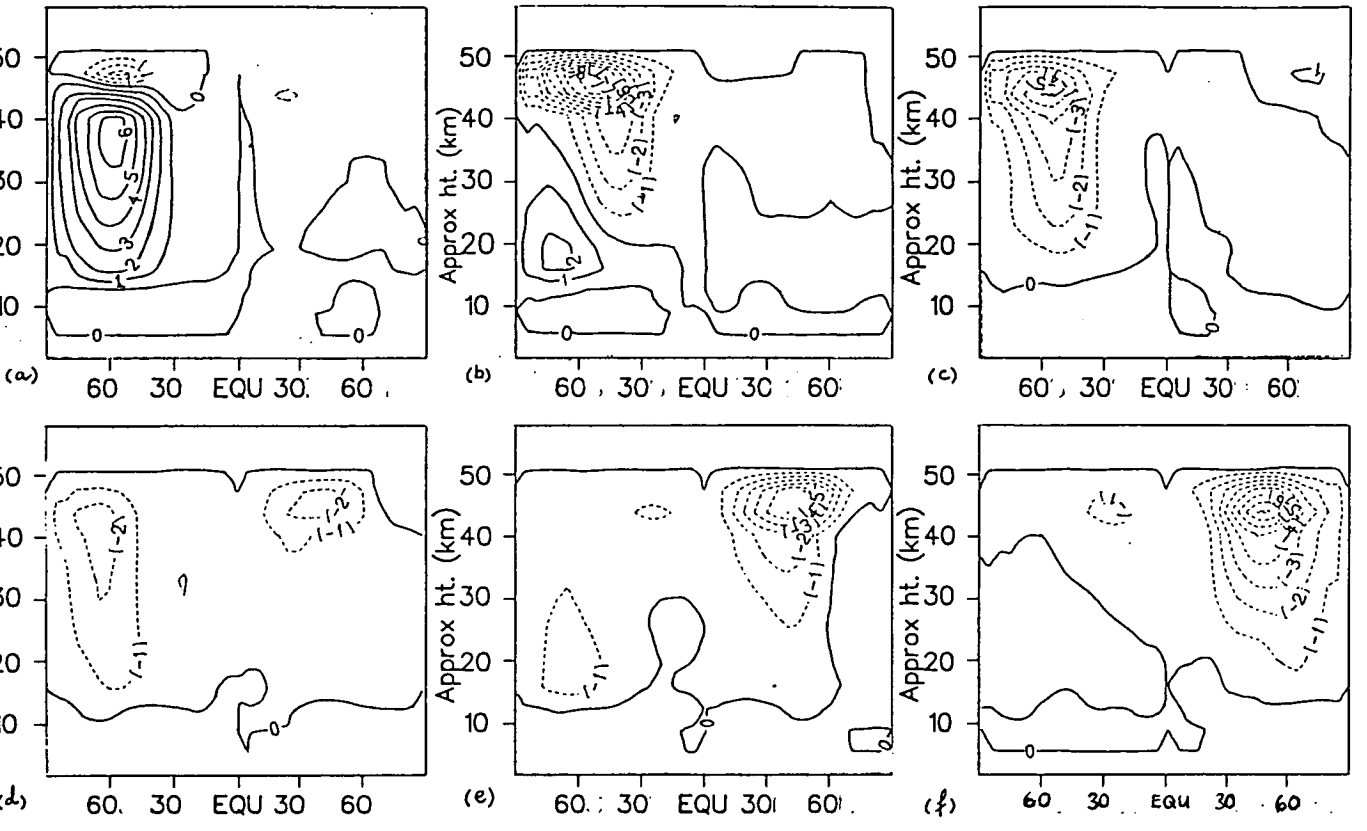


Figure 4.4: $\mathcal{F}^D$ *in* $ms^{-1}day^{-1}$ *for (a) Jul/Jan (b) Aug/Feb (c) Sep/Mar (d) Oct/Apr (e) Nov/May (f) Dec/Jun. Dashed contours indicate negative values.*

certainly appear due to vacillating Rossby waves (see Chapter 2) but in the long time average these positive regions should disappear. We will therefore assume that below about 42km $\mathcal{F}$ is, in general, negative.

With this assumption, we deduce the following about $\mathcal{F}$ in the stratosphere (below about 42km) from fig.(4.4).
For the NH :-

(i) In late autumn/early winter (fig.(e)) $\mathcal{F}$ is significant (being over $2ms^{-1}day^{-1}$ in magnitude) at about 40N. It might in fact be larger than this if there is a negative $\mathcal{F}$ in the SH in May/Jun, as suggested by fig.(4.3).

(ii) The magnitude of $\mathcal{F}$ increases throughout winter, rising to $4ms^{-1}day^{-1}$ (fig.(f)) and over $6ms^{-1}day^{-1}$ (fig.(a)). The peak moves polewards, from 45N to 60N. These values are probably not masked by a $\mathcal{F}$ in the SH since the SH planetary waves are small in winter (fig.(4.3)).

(iii) in late winter (fig.(b)) $\mathcal{F}$ is still at least $2ms^{-1}day^{-1}$ in magnitude in the

lower stratosphere, though its value in the upper stratosphere is less negative than that in the SH.

For the SH :-

(iv) $\mathcal{F}$ is at least $3ms^{-1}day^{-1}$ in magnitude in late winter and early spring (figs.(b) and (c)), the actual value being uncertain because of a significant NH planetary wave amplitude until April (see fig.(4.3)).

(v) In late spring, (fig.(d)) $\mathcal{F}$ is about $-2ms^{-1}day^{-1}$, the peak having moved polewards from about 50S to 60S.

The values of $\mathcal{F}^D$ above 42km are in general much larger than those below. If they are indeed due to gravity wave breaking, which occurs throughout the year in both hemispheres, then $\mathcal{F}$ cannot be deduced from $\mathcal{F}^D$ as was done below 42km in winter and spring. However, from figures 4.4(e) and (f) it can be seen that $\mathcal{F}$ in the NH in early winter is over $5ms^{-1}day^{-1}$ more negative than $\mathcal{F}$ in the SH. In late winter and early spring, figs 4.4(a) to (c), the situation is reversed and the SH $\mathcal{F}$ can be up to $8ms^{-1}day^{-1}$ more negative than the NH $\mathcal{F}$. (Note that the apparent fall off in $\mathcal{F}^D$ above about 45km is due simply to the absence of data there). This agrees with the results of Shine (1989) who found that at about 60km $\mathcal{F}$, which he attributed mainly to gravity wave breaking, in the NH peaked at about $-40ms^{-1}day^{-1}$ in Nov/Dec before declining to about $-20ms^{-1}day^{-1}$ in Jan/Feb while in the SH $\mathcal{F}$ maintained a minimum of about $-30ms^{-1}day^{-1}$ throughout the winter. The fact that this same behaviour occurs at about 45km in our results supports the suggestion that the values of $\mathcal{F}^D$ above 42km are mainly due to breaking gravity waves.

4.4.1 Sensitivity to approximations

In figure 4.5, X (see section 4.3) is plotted for August. The pattern is typical of the other months too, with X in the summer hemisphere being close to unity, rarely exceeding 1.2, while in winter X is close to unity at about 45 degrees, falling to typically 0.7 closer to the pole (the August value of about 0.5 is the extreme) and rising, sometimes to over 2.0, at low latitudes. These large values of X would have only a small effect on the estimate of $\mathcal{F}^D$ as they occur equatorwards of 30 degrees, where the Coriolis parameter is small (see equation 4.4). The effect on $\mathcal{F}^D$ of a

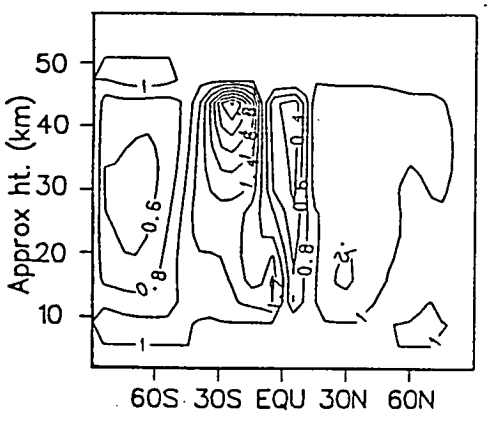


Figure 4.5: *X, the quantity approximated by unity, for August 1980*

30% variation in the final term in eq. 4.4 was estimated by multiplying this term by a factor of 1.3 before recalculating $\mathcal{F}^D$. The result was simply an increase in the magnitude of $\mathcal{F}^D$ by between 20 and 30%, with the pattern of the contours being largely unchanged. This reflects the fact that the dominant terms in eq. 4.4 are on the RHS, with $\overline{u}_t$ generally being the small residual. As noted, though, X is not independent of latitude, particularly in winter, and hence in reality this approximation (of $X \simeq 1$) would probably affect the shape of $\mathcal{F}^D$ more strongly.

The other important approximation made in deriving values for $\mathcal{F}^D$ was in the use of, and the values assigned to, the time constant $1/\alpha$. Gille and Lyjak (1986) and Pawson et al. (1991) have estimated time constants and found values of less than 10 days above 40km (their results can be seen in Chapter 8) while the values used here were rather higher (see Chapter 7). The effect on $\mathcal{F}^D$ of a 50% reduction in the time constant, (and hence a 100% increase in α) to bring our values in line with theirs, depended strongly on height. The change in $\mathcal{F}^D$ below about 30km was small (less than $1 ms^{-1}day^{-1}$)while above about 42km $\mathcal{F}^D$ increased by up to 100%. The increase at about 40km was about 50%, the peak value in fig.4.4(a), for example, increasing to about $9 ms^{-1}day^{-1}$. Thus the values of $\mathcal{F}^D$ in the lower stratosphere are much less sensitive to α than those in the upper stratosphere and mesosphere. In general, $\mathcal{F}^D$ increased in magnitude but changed little in shape.

4.5 The effects of $\mathcal{F}$ revealed by IHDs

In this section it is proposed that taking the IHD is roughly equivalent to removing the effect on $\overline{u}$ and $\overline{T}$ of the large meridional circulation common to both hemispheres. This circulation is probably mainly due to mesospheric gravity waves

and a seasonally varying solar heating. It is therefore assumed that the circulation remaining after this seasonally-symmetric circulation has been removed is due mainly to the stratospheric $\mathcal{F}^D$.

A qualitative description of the way in which $\mathcal{F}$ would affect the $\overline{u}$ and $\overline{T}$ fields, in an atmosphere with no other generator of a meridional circulation, will be given here. This description will be used in later chapters when analysing the results from model runs where it will be found to be very useful. The strong resemblance between these $\overline{u}$ and $\overline{T}$ fields and the observed IHDs in $\overline{u}$ and $\overline{T}$ will then be illustrated. Two of the effects of $\mathcal{F}$ revealed by the IHDs are the relatively cool tropical tropopause from January till March, and the deceleration of $\overline{u}$ in one hemisphere due to a negative $\mathcal{F}$ in the opposite hemisphere.

4.5.1 Effect of $\mathcal{F}$

Let us consider first the northern hemisphere and assume that the mean meridional circulation is initially zero. This is similar to studying the IHD in an atmosphere in which $\mathcal{F}$ is initially zero since, from our previous assumption, taking the inter-hemispheric difference is equivalent to cancelling out the meridional circulation which is not due to a $\mathcal{F}^D$The following discussion arrives at a conclusion similar to that of Eliassen (1951), at least in the qualitative details, but in isentropic coordinates.

If there is a region of negative $\mathcal{F}$ then $\overline{u}_z$ will tend to increase above it and decrease below it. If thermal wind balance is to be maintained then $\overline{T}_\phi$ must decrease above and increase below. In isentropic coordinates, a falling temperature corresponds to a rising isentrope, so a decrease in $\overline{T}_\phi$ along an isentrope requires it to rise on its poleward side and fall on its equatorward side. Thus, polewards of the region of negative $\mathcal{F}$, the separation of the isentropes (and hence the isentropic density) will be increasing and on the equatorward side the density will be decreasing. $\overline{W}$ will at first be small since the change in $\overline{T}$ is initially small. Continuity then requires a *poleward* mass flux to balance the changing densities. The Coriolis force due to this poleward $\overline{V}$ opposes the negative $\mathcal{F}$ and, after a time of the order of the radiative time constant, approximately balances it. Over the same time span the temperature change is likely to induce a balancing change in $\overline{W}$ (warming the adiabatically cooling regions and cooling the adiabatically warming regions) so there will then be a two-celled circulation (see Figure 4.6). The equatorward

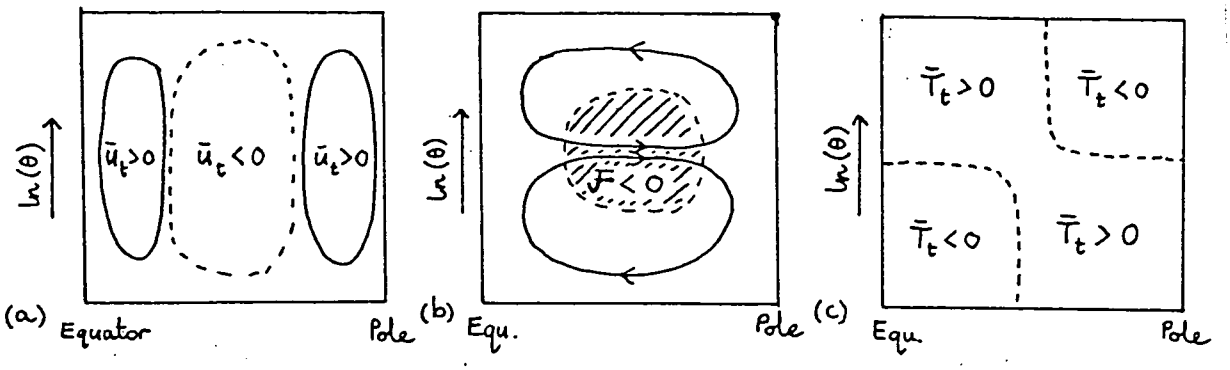


Figure 4.6: *Effect of region of negative $\mathcal{F}$ on (a) $\bar{u}$ (b) the meridional circulation and (c) $\bar{T}$*

return flow above and below acts to spread the deceleration of $\bar{u}$ vertically, while on both sides of the region of negative $\mathcal{F}$ the poleward flow accelerates $\bar{u}$ there. The same conclusions are reached for the southern hemisphere.

In the troposphere there is a region of positive $\mathcal{F}$ below a comparable region of negative $\mathcal{F}$ (estimated values are discussed in Chapter 5). Following a similar argument as above, an equatorwards mass flux will be induced through a region of positive $\mathcal{F}$. The poleward mass flux above is thus partially balanced by the equatorward flux below and there will be a largely one-celled circulation, with heating near the pole and cooling near the equator. (See Figure 4.7). This will be made use of in Chapter 6 when discussing the behaviour of the model troposphere.

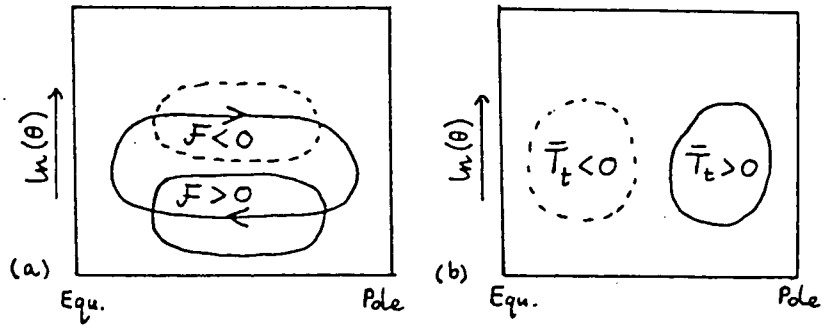


Figure 4.7: *Effect of regions of positive and negative $\mathcal{F}$*

In both cases the temperature rise near the pole would be expected to be larger than the fall near the equator. This is because the mass of a zonally-symmetric tube of air is $(2\pi a \cos\phi \bar{\rho}\, dy\, dz)$, so if, in the region of positive $\mathcal{F}$ we have an equatorward flux of mass we would expect $\bar{\rho}$ to decrease more at the pole than it

would increase at the equator (due to the cosϕ factor), and hence for there to be a greater pressure and temperature change at the pole.

The effect of eddies is simple to describe in isentropic coordinates, mainly because of the lack of eddy heat fluxes in the thermodynamic equation, so that all the eddy-forcing is in the momentum equation. The picture in log-pressure coordinates, where eddy heat fluxes enter the thermodynamic equation, can however be simplified by absorbing their effect in a redefinition of the meridional circulation as the residual circulation. Figure(4.8) is from Dunkerton et al. (1981) and shows the polewards residual circulation induced by the shaded region where ∇.F is large and negative in their 3D model.

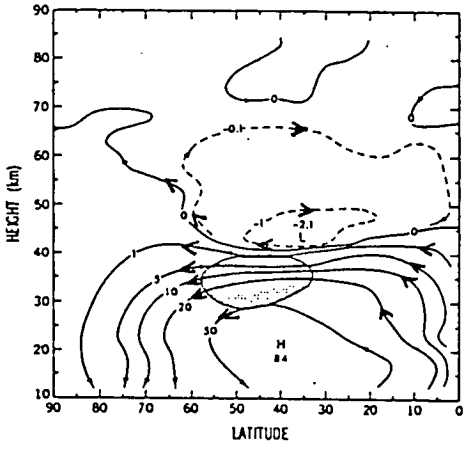


Figure 4.8: *Residual circulation due to shaded region of large negative* ∇.F *(after Dunkerton et al. 1981).*

4.5.2 Tropopause cooling and the cross-equator effect

These ideas will now be used when interpreting the IHDs in $\overline{u}$ and $\overline{T}$. The left side of figure(4.1)is similar to fig.(4.6(c)) with the signs reversed. This reveals the action over several months of a region of negative $\mathcal{F}$ in the NH stratosphere which, because of taking the IHD, appears on the LH side of fig.(4.1) with the signs of its effects reversed. As expected, the changes in $\overline{T}$ are greater nearer the poles.

Also from fig. 4.1 it can be seen that the equatorial tropopause is over 5K cooler in March than it is in September. This statement applies also to January and February compared with July and August. This is in agreement with fig. 4.6(c) where it is seen that a negative $\mathcal{F}$ in the stratosphere leads to a cooling of the

equatorial lower stratosphere and tropopause. This holds whichever hemisphere $\mathcal{F}$ is in and since it is likely that $\mathcal{F}$ is negative in one hemisphere from September till February, while from May till August $\mathcal{F}$ is small in both hemispheres, this would appear to explain the cold equatorial tropopause from January till March. This same annual cycle, with the equatorial tropopause being about 4K cooler in February than in August, has been observed by Barnett (1974), among others. He also noted the annual cycle near the stratopause, which had a similar amplitude but the opposite phase. He said there was no obvious explanation for these cycles, but mentioned Reed and Vlcek's suggested cause of an asymmetry in the tropospheric circulation. This would not, however, explain the cycle near the stratopause. The mechanism proposed here, of a stratospheric $\mathcal{F}$ cooling the equatorial tropopause and warming the equatorial stratopause (see fig.4.6(c)), would explain both cycles. This will be further confirmed using the model in Chapter 7.

Figure 4.9 shows $\overline{u}^D{}_t$ for Sep/Mar and Dec/Jun. The left side of fig.(a) (due mainly to the SH spring $\mathcal{F}$) and the right side of fig.(b) (due to the NH winter $\mathcal{F}$) resemble fig.4.6(a) in that there is a large negative region in mid-latitudes with a positive region nearer the equator. The polar positive region of fig.4.6(a)

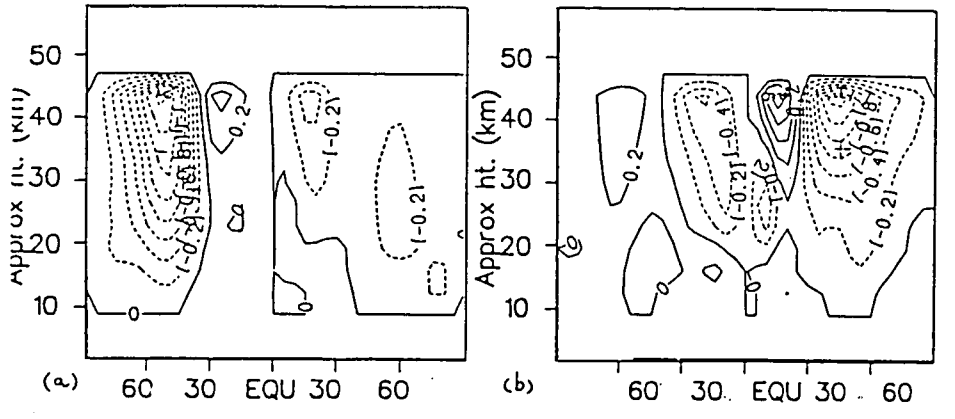


Figure 4.9: $\overline{u}^D{}_t$ *for (a) Sep/Mar and (b) Dec/Jun. Dashed contours indicate negative values.*

is, however, missing from figures 4.9(a) and (b), but as we shall see in the next section, it appears in the $\bar{u}^D_t$ field of the symmetric vortex.

The other feature apparent in both figs.4.9(a) and (b) is the negative region at low latitudes on the *right* side of fig.4.9(a) (which is during autumn) and on the *left* side of fig.4.9(b) (during summer). Over autumn and summer $\mathcal{F}$ is small, so these negative regions of $\overline{u}^D{}_t$ are unlikely to be due to the $\mathcal{F}$ of those hemispheres. It is more likely that they are due to the meridional circulation induced by $\mathcal{F}$

extending across the equator from the spring and winter hemispheres and advecting air with relatively low angular momentum equatorwards, in the same way that the complementary positive regions in the spring and winter are due to polewards advection of air with high angular momentum. Again, this suggestion is supported by a model experiment described in Chapter 7.

4.6 IHDs and the symmetric vortex

When $\mathcal{F}^D$ was estimated using the $\tilde{u}$ and $\tilde{T}$ fields of the symmetric vortex instead of the $\overline{u}$ and $\overline{T}$ fields, the peak values in winter and spring were found to be between 1 and 2 $ms^{-1}day^{-1}$ smaller in magnitude. Graphs of $\tilde{u}^D$ and $\overline{u}^D$ (not shown) reveal that symmetrising the vortex in general reduces the IHDs, and hence should also reduce the $\mathcal{F}^D$ deduced from them, as found. Thus, a vortex which vacillates and is distorted from a symmetrical shape will require an extra -1 to -2 $ms^{-1}day^{-1}$ of $\mathcal{F}$, if it is to be reproduced in a 2D model, compared with a symmetric vortex.

Inspection of the $\tilde{u}^D$ fields of the symmetric vortex reveal the positive region near the pole, expected from fig.4.6(a) but not found in the fields of $\overline{u}^D{}_t$. For example,

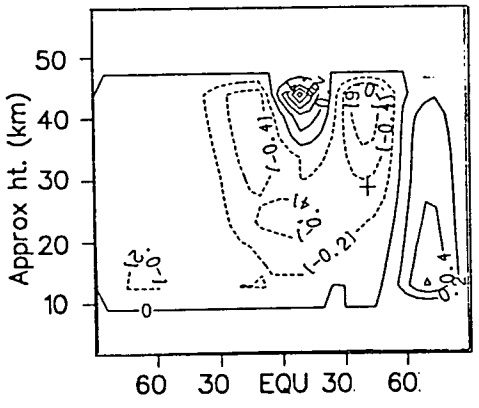


Figure 4.10: $\tilde{u}^D_t$ *for Dec/Jun. Dashed contours indicate negative values.*

in fig.4.10 is $\tilde{u}^D_t$ for Dec/Jun, which should be compared with fig.4.9(b). Thus it appears that symmetrising the vortex removes the blurring effect of asymmetries near the pole and reveals more clearly the effect of irreversible PV fluxes, or $\mathcal{F}$.

4.7 Summary

In summary of this chapter, then, we can say that the NH winter values of $\mathcal{F}$ below about 42km can be deduced fairly confidently from $\mathcal{F}^D$ and are estimated to reach a peak of about -6ms^{-1}day^{-1} in Jan/Feb, and to be significant from November until February or March. The early spring values of $\mathcal{F}$ in the SH are less confidently deduced, the NH planetary wave amplitude still being large in the spring though varying from year to year, but are estimated to be at least -3ms^{-1}day^{-1} in August and September and to still be significant (-2ms^{-1}day^{-1}) in October. The above values increase by about 50% if shorter diabatic time constants, similar to those estimated by Gille and Lyjak(1986) are used. In both hemispheres a polewards movement of the peak in $\mathcal{F}$ over its lifetime is also deduced.

Uncertainties due to approximations used in deriving the difference momentum equation contribute 20 to 30% uncertainty to $\mathcal{F}^D$. Uncertainties in $\mathcal{F}^D$ due to the time constant are strongly dependent on height, a halving of the time constant producing a 50% increase in $\mathcal{F}^D$ at 40km and a doubling at 49km. The effects of gravity wave breaking are observed above about 42km. IHD fields allow the effects of $\mathcal{F}$ to be separated out from the effect due to the part of the meridional circulation which is the same for both hemispheres, and are revealed even more clearly in the symmetric vortex fields, free from the blurring due to polar asymmetries. IHDs also reveal the influence of $\mathcal{F}$ on the equatorial tropopause and on the opposite hemisphere due to an induced cross-hemispheric flow.

Chapter 5

The Model

God's done not a bad job with the atmosphere
but a little tweak there and a little nudge here
would for Him, I am sure, be no more than a doddle
and then it would look, really quite like my model.

This chapter will describe a two-dimensional model whose vertical coordinate above the tropopause is isentropic while below the tropopause sigma-type coordinates are used to provide a smooth interface with the ground. The model is intended eventually to be used as a tool for chemical studies, but its use in this thesis is limited to studying the effect of stratospheric planetary waves on the dynamics of the atmosphere, since it was felt that until the effect of stratospheric planetary waves could be correctly included in the model its use in such perturbation studies was limited. It does not currently represent any chemicals as dependent variables. The gases used in the stratospheric and mesospheric radiative heating calculations are specified independently of the model state. The tropospheric processes are crudely parametrised, but allow an interaction between the stratosphere and troposphere, and so allow us to estimate the effect of stratospheric planetary waves on the troposphere and the effect on the stratosphere of changing the tropospheric parameters.

5.1 Formulation of the Model

The dynamical equations used in the model are the zonally-averaged continuity equation (2.16)

$$\overline{\rho}_t \cos\phi + \overline{\mathrm{V}}_\phi + \overline{\mathrm{W}}_z = 0 \tag{5.1}$$

and the zonal momentum equation (2.17)

$$\cos\phi(\overline{\rho}\,\overline{\tau})_t + (\overline{V}\,\overline{\tau})_\phi + (\overline{W}\,\overline{\tau})_z = a\cos^2\phi\,\overline{\rho}\,\mathcal{F}. \tag{5.2}$$

A balance is maintained between $\overline{\tau}$ and $\overline{\rho}$ by the thermal wind equation, derived as follows. From eqn.(2.4),

$$u_z\left(2u\frac{\tan\phi}{a} + f\right) = -M_{yz} \tag{5.3}$$

where the material derivative of v has been neglected but the $\tan\phi$ term has been kept for energetic consistency with (5.2) (see White (1978)). Use of the definition of angular momentum (2.7) and the hydrostatic approximation (2.6), taking the zonal mean and neglecting $\overline{\tau'_z\tau'}$ compared with $\overline{\tau}_z\overline{\tau}$ gives

$$2\tan\phi\overline{\tau}_z\overline{\tau} = -c_p\overline{T}_\phi a^2\cos^2\phi$$

From (2.1) $\overline{T}$ can be expressed as $\theta p_0^\kappa\overline{p^\kappa}$ and if the difference between $\overline{p}^\kappa$ and $\overline{p^\kappa}$ is neglected the thermal wind equation used in the model,

$$2\tan\phi\overline{\tau}_z\overline{\tau} = -c_p\kappa\theta\left(\frac{\overline{p}}{p_0}\right)^\kappa\left(\frac{\overline{p}_\phi}{\overline{p}}\right)a^2\cos^2\phi, \tag{5.4}$$

is obtained.

The two approximations made in deriving the above equation need some discussion. Neglecting the terms is deemed necessary at present, since there seems to be no simple way to deal with them properly. A proper treatment might be required in future if further study reveals them to be important. Their maximum values will now be estimated.

First the relative importance of the momentum eddy term will be considered, by estimating

$$X = \frac{\overline{\tau'_z\tau'}}{\overline{\tau}_z\overline{\tau}}.$$

Both u' and $\overline{u}$ are assumed to reach their maximum values, of u_1 and u_0, near the stratopause, and are small near the tropopause. This latter assumption might not be valid for u', but it should overestimate, and hence put an upper bound on, u'_z. In the stratosphere, then,

$$X \simeq \frac{u_1^2}{2\,u_0(u_0 + 463\cos\phi)},$$

(assuming that $\overline{u'^2} \simeq \frac{1}{2}u_1^2$). For a wave-1 disturbance, u_1 can be large (say 40 m/s) while at the same time u_0 is small (say 10 m/s). This gives, at mid- to high-latitudes ($\cos\phi \simeq 0.5$), $X \simeq 1600/4800 \simeq 0.3$. This is still perhaps an overestimate,

and X falls off rapidly as u_0 increases (0.08 for $u_0 = 20$ m/s), so it does not seem unreasonable to neglect the eddy term.

An estimate of the importance of the second approximation is given by the ratio $Y = \overline{p^\kappa}/\overline{p}^\kappa$. A typical wintertime amplitude of $p'/\overline{p}$ is 0.18 (corresponding to $T'/\overline{T} = 0.05$) and if therefore, for simplicity, it is assumed that $p=1.18\overline{p}$ over half the zonal integral, and $p=0.82\overline{p}$ for the other half then

$$Y = (1.18^\kappa + 0.82^\kappa)/2 \simeq 0.997.$$

Even if $p'=0.5\overline{p}$, Y is still only 0.97, so this approximation seems to be well justified.

The thermodynamic equation $dQ = TdS$, where Q is the internal energy and $S(\equiv c_p \ln\theta)$ is the entropy, leads to the expression

$$wT = H, \tag{5.5}$$

which relates the vertical velocity in isentropic coordinates, $w = \frac{d\ln\theta}{dt}$ to the diabatic heating rate $H = \frac{1}{c_p}\frac{dQ}{dt}$.

Taking the zonal average and ignoring $\overline{w'T'}$, this becomes

$$\overline{W} = \left(\frac{\overline{H}}{\overline{T}}\overline{\rho} + \overline{w'\rho'}\right)\cos\phi \tag{5.6}$$

Again, the eddy term $\overline{w'T'}$ might need to be included later, but we note that for a typical winter amplitude for $(T'/\overline{T})$ of 0.05 and an amplitude of 4 for $(w'/\overline{w})$, $\overline{w'T'}$ is still only about one tenth of $\overline{w}\,\overline{T}$.

The four equations, then, which must be solved simultaneously in the model are 5.1, 5.2, 5.4 and 5.6 (although they will be modified by a change of variable in the troposphere - see below). These equations are basically the same as those of Tung and Yang (1988). The difference between this model and theirs is that while $\overline{u}$ (or $\overline{T}$) is the independent variable in this model, in theirs it is $\mathcal{F}$. The numerical method used here to solve these equations will now be described.

5.2 Method of solution

5.2.1 Sigma Coordinates

In order that the model equations can be solved conveniently, a change of variables is made in the model troposphere. The need for this arises because the air at the earth's surface is not at a constant potential temperature and hence some tropospheric isentropic levels will intersect the ground (see figure 5.1). Lorentz

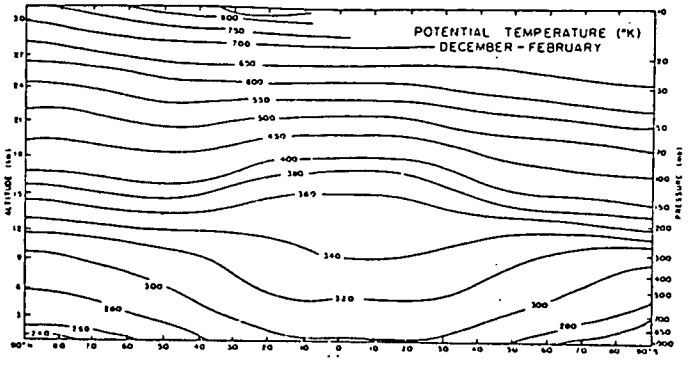


Figure 5.1: *Height of tropospheric isentropic surfaces. From Dutton (1976)*

(1955) proposed 'underground isentropes' which were used by Andrews (1983) in deriving an E-P flux theorem, but the method eventually adopted here was that of Tung and Yang (1988), in which the isentropic coordinates are replaced by sigma coordinates in the troposphere.

An earlier version of the model used isentropic coordinates in the troposphere, and so had 'underground' model boxes. This led to difficulties in maintaining thermal wind balance between two boxes when one of them was below the ground. However, since tropospheric waves are much less adiabatic than those in the stratosphere, there seemed to be little point, for the purposes of this study, to use isentropic coordinates near the ground.

The hybrid vertical coordinate is defined as

$$\eta = F\frac{ln(\theta/\theta_s)}{ln(350/\theta_s)} \tag{5.7}$$

where $F = ln(350/250)$, and for $\theta < 350$K, θ_s is the surface θ, while for $\theta > 350$K, $\theta_s = 250$K. Therefore, $\eta = 0$ at the ground, and $\eta = ln(\theta/250)$ for $\theta > 350$K. The $\theta = 350$K surface roughly corresponds to the tropopause, so above the tropopause the coordinates are isentropic while below, the surfaces are evenly spaced in $\ln\theta$, but

with a spacing which depends on latitude, till they meet the ground. The horizontal coordinate surfaces are shown in fig.5.2 (using ln(p) as a vertical coordinate). The 350K isentropic surface lies at the top of the 4th layer of boxes. Note that in the

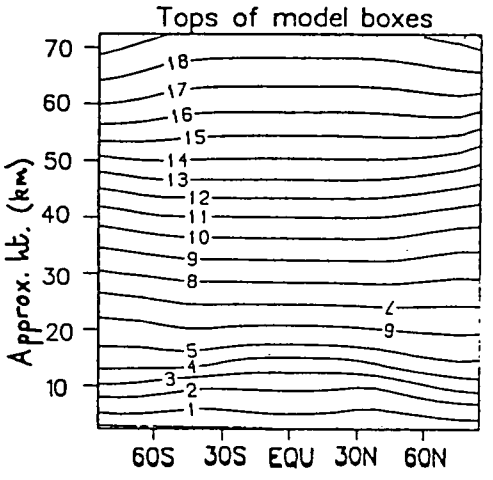


Figure 5.2: *Heights of tops of model boxes for the model September*

troposphere the surfaces of constant η are not isentropic surfaces.

The use of sigma coordinates affects the definition of the vertical velocity, and the thermal wind equation. Since θ at the surface can vary with time, the vertical velocity must be redefined as

$$w = \frac{d\eta}{dt} = \frac{F}{ln(350/\theta_s)}\left[\frac{H}{T} + \frac{d(ln\theta_s)}{dt}\left(\frac{\eta}{F} - 1\right)\right] \tag{5.8}$$

which reduces to 5.5 for $\theta > 350$K. The thermal wind equation is then rewritten, using the relation

$$\left(\frac{\partial T}{\partial \phi}\right)_{\eta=\text{const}} = \left(\frac{\partial T}{\partial \phi}\right)_{z=\text{const}} + \frac{\partial \eta}{\partial z}\frac{\partial T}{\partial \eta}\left(\frac{\partial z}{\partial \phi}\right)_{\eta=\text{const}},$$

as

$$\overline{\tau}_\eta \overline{\tau} = \frac{-c_p a^2 \cos^2\phi}{2\tan\phi}\left[\frac{ln(350/\theta_s)}{F}\overline{T}_\phi + \left(\frac{\eta}{F} - 1\right)\frac{d(ln\theta_s)}{d\phi}\frac{\partial \overline{T}}{\partial \eta}\right] \tag{5.9}$$

Again, for $\theta > 350$K, θ_s is independent of ϕ and 5.9 reduces to 5.4. Since $\overline{p}$, and not $\overline{T}$, is chosen as a model variable, the derivatives of $\overline{T}$ must be replaced with derivatives of $\overline{p}$. Defining

$$A = ln\theta_s, \quad B = \left(\frac{\eta}{F} - 1\right), \quad C = BA_\phi, \quad D = \frac{p_\eta}{p}, \quad E = \frac{p_\phi}{p}, \quad G = \frac{1}{F}ln(350/\theta_s)$$

and noting from the definition of θ, that $\Delta(lnT) = \Delta(ln\theta) + \kappa\Delta(lnp)$ and from 5.7, that $\Delta(ln\theta) = -B\Delta A$ along a surface of constant η, and hence that $\overline{T}_\phi =$

$\overline{\mathrm{T}}(\kappa E - C)$ and $\overline{\mathrm{T}}_\eta = \overline{\mathrm{T}}(\kappa D + G)$, eqn.5.9 can be rewritten as

$$\overline{\tau}_\eta \overline{\tau} = \frac{-c_p a^2 \cos^2\phi}{2\tan\phi} \kappa \overline{\mathrm{T}}(GE + CD) \ , \tag{5.10}$$

which replaces the derivatives of $\overline{\mathrm{T}}$ with those of $\overline{\mathrm{p}}$.

The equations to be solved now are 5.6 (with w defined by 5.8), 5.10, 5.1 and 5.2.

5.2.2 The stream function as a means of finding $\overline{\mathrm{V}}$

From equations 5.1 and 5.2 and 2.5 it is seen that $\overline{\mathrm{p}}_t$ (with suitable boundary conditions) and $\overline{\tau}_t$ are determined uniquely by $\overline{\mathrm{V}}$, $\overline{\mathrm{W}}$ and $\mathcal{F}$. Now, since $\overline{\mathrm{W}}$ is determined from the heating rate calculations, and $\mathcal{F}$ is specified, the constraint of thermal wind balance between $\overline{\mathrm{p}}_t$ and $\overline{\tau}_t$ is sufficient to determine $\overline{\mathrm{V}}$. Differentiating 5.10 with respect to time results in the expression

$$\begin{aligned} \overline{\tau}_{zt}\overline{\tau} + \overline{\tau}_z \overline{\tau}_t \ = \ & \frac{-c_p a^2 \cos^2\phi}{2\tan\phi} \kappa \overline{\mathrm{T}} [(GE + CD) \left(\frac{\overline{\mathrm{p}}_t}{\overline{\mathrm{p}}} (\kappa - 1) - BA_t \right) \\ & + \frac{\overline{\mathrm{p}}_{\phi t}}{\overline{\mathrm{p}}} G + \frac{\overline{\mathrm{p}}_{zt}}{\overline{\mathrm{p}}} C - A_t \frac{E}{F} + DBA_{\phi t}] \end{aligned} \tag{5.11}$$

(since $\overline{\mathrm{T}}_t = \overline{\mathrm{T}}(\kappa \frac{\overline{\mathrm{p}}_t}{\overline{\mathrm{p}}} - BA_t)$, $\quad G_t = -A_t/F, \quad C_t = BA_{\phi t}$)
Expressions for $\overline{\tau}_t$ and $\overline{\mathrm{p}}_t$ in terms of $\overline{\mathrm{W}}$, $\mathcal{F}$ and $\overline{\mathrm{V}}$, must now be found and substituted into 5.11, after which $\overline{\mathrm{V}}$ can be found. However, in order to find an expression for $\overline{\mathrm{p}}_t$ it appears at first sight that it is necessary to integrate 5.1 vertically. This complication can be avoided by rewriting 5.1 as

$$(\overline{\mathrm{W}} - \frac{\cos\phi}{g} \overline{\mathrm{p}}_t)_z + \overline{\mathrm{V}}_\phi = 0$$

and defining a stream function ψ such that

$$\psi_\phi = \overline{\mathrm{W}} - \frac{\cos\phi}{g} \overline{\mathrm{p}}_t \tag{5.12}$$

and

$$\psi_z = -\overline{\mathrm{V}} \tag{5.13}$$

so that 5.1 is automatically satisfied. Thus the expression for $\overline{\mathrm{p}}_t$ is

$$\overline{\mathrm{p}}_t = \frac{g}{\cos\phi} (\overline{\mathrm{W}} - \psi_\phi). \tag{5.14}$$

Equation 5.2 can then be rewritten, using 5.14 and 5.13, as

$$\overline{\tau}_t = \frac{1}{\overline{\rho}\cos\phi}(a\cos^2\phi\,\overline{\rho}\,\mathcal{F} + (\psi_z\overline{\tau})_\phi - (\overline{W}\overline{\tau})_z - \overline{\tau}\,\overline{\rho}_t\cos\phi), \tag{5.15}$$

This expression for $\overline{\tau}_t$ conserves the total angular momentum, to second order in the time-step Δt, as will now be shown. The changes in $\overline{\tau}$ and $\overline{\rho}$ over the time step are denoted by $\Delta\overline{\tau}$ and $\Delta\overline{\rho}$, which to a first approximation are $\overline{\tau}_t\Delta t$ and $\overline{\rho}_t\Delta t$ respectively. Equation 5.15 can be re-written as

$$(\overline{\rho}\Delta\overline{\tau} + \overline{\tau}\Delta\overline{\rho})\cos\phi = [a\cos^2\phi\,\overline{\rho}\,\mathcal{F} - (\overline{V}\overline{\tau})_\phi - (\overline{W}\overline{\tau})_z]\Delta t \tag{5.16}$$

The total angular momentum in the model is given by the sum $\Sigma(\overline{\tau}\,\overline{\rho}\beta)$, where $\beta = \delta V\cos\phi$ and δV is the volume of a model box. In the absence of $\mathcal{F}$ the change in the total momentum over one time step is then

$$\begin{aligned}\Sigma(\overline{\tau}+\Delta\overline{\tau})(\overline{\rho}+\Delta\overline{\rho})\beta - \Sigma\overline{\tau}\,\overline{\rho}\beta &= \Sigma\overline{\tau}\Delta\overline{\rho}\beta + \Sigma\overline{\rho}\Delta\overline{\tau}\beta \\ &\quad +\Sigma\Delta\overline{\tau}\Delta\overline{\rho}\beta \\ &= \Delta t\Sigma(-(\overline{V}\overline{\tau})_\phi - (\overline{W}\overline{\tau})_z)\beta/\cos\phi \\ &\quad +\Sigma\Delta\overline{\tau}\Delta\overline{\rho}\beta\end{aligned}$$

where eqn.5.16 has been used to reach the second line. The first sum on the RHS is zero since it is equal to the total flux of $\overline{\tau}$ by the mean circulation across the model boundary, and therefore the total angular momentum in the model is conserved (in the absence of $\mathcal{F}$) to 2nd order in Δt. Even with a time step of one day, this 2nd order term is negligible compared with the other sources of angular momentum.

Equations 5.14 and 5.15 are then substituted into 5.11 to get a second order partial differential equation for ψ, which is normally elliptic (see section 5.2.6). This enables the equation to be solved using the method of relaxation and ψ is then used to find the new $\overline{\tau}$ and $\overline{\mathrm{p}}$ fields which satisfy the thermal wind equation 5.4.

The finite difference forms of equations 5.14, 5.15 and 5.11 (with its time derivatives removed by using equations 5.14 and 5.15) will now be given.

5.2.3 Spatial finite differencing

The grid used to discretise 5.11 has 19 boxes from pole to pole and is 29 boxes high, and is shown in fig 5.3. Chemical mixing ratios, $\mathcal{F}$ and $\overline{\tau}$ are held in the centres

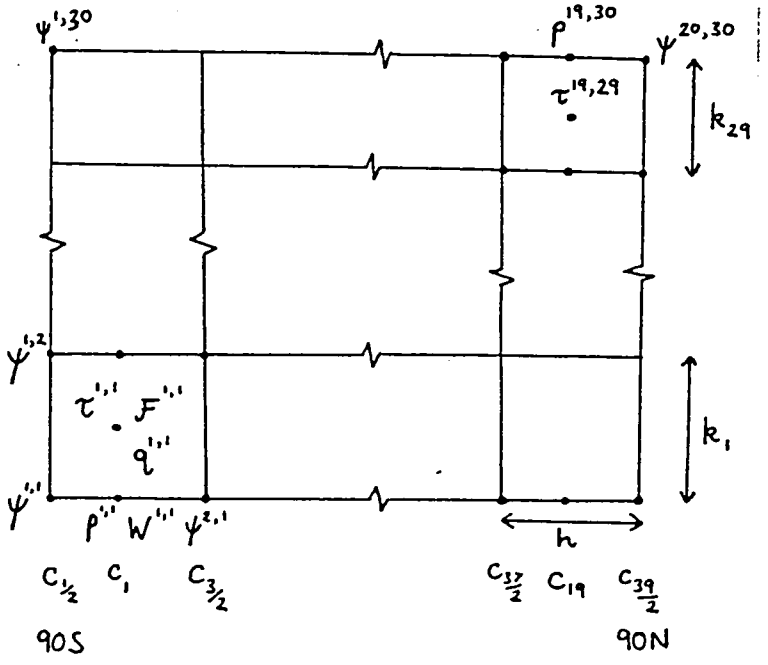


Figure 5.3: *Model Grid*

of the boxes, $\overline{p}$ and $\overline{\overline{W}}$ at the middle of the top and bottom of boxes, and ψ at the box corners. k_m is the thickness of the mth layer, h is the box width, which is a constant $\pi/19$, c_n is $\cos\phi$ at the middle, and $c_{n+1/2}$ is $\cos\phi$ at the northern edge, of the nth box from the south pole. The vertical resolution is variable, but k_m for each of the layers in the model used to produce results for this thesis is given in Table 5.1. The heights of the tops of the model boxes in September is shown in

layer (m)	1-4	5	6,7	8-14	15	16,17	18-29
box height (k_m)	0.084	0.1	0.17	0.18	0.144	0.130	0.144

Table 5.1: *Vertical Resolution*

figure 5.2 for the control run of Chapter 6. Note that the average box height is close to half a scale-height (see section 2.1), and that the boxes are reasonably well spaced in the vertical. The join between the sigma coordinates and the isentropic coordinates occurs at the top of the 4th model layer. Because the box height is variable, care must be taken when interpolating vertically. For example, $\overline{\tau}$ at the point where $p^{1,2}$ is held is approximated by the expression

$$\overline{\tau} \simeq \frac{k_2\overline{\tau}^{1,1} + k_1\overline{\tau}^{1,2}}{k_1 + k_2}$$

First order finite differences are used so that, for example, the estimate of $\overline{\tau}_z$ at the same point as above is

$$\overline{\tau}_z \simeq \frac{2(\overline{\tau}^{1,2} - \overline{\tau}^{1,1})}{k_1 + k_2}$$

The finite difference form of 5.14 is

$$\overline{p}_t^{n,m} = \frac{g}{c_n}\left[\overline{W}^{n,m} - \frac{1}{h}(\psi^{n+1,m} - \psi^{n,m})\right] \tag{5.17}$$

and the finite difference form of 5.15 is, after some rearrangement,

$$\frac{\overline{\tau}_t^{n,m}}{c_n} = M^{n,m}(\psi^{n+1,m+1} - \psi^{n+1,m}) + N^{n,m}(\psi^{n,m+1} - \psi^{n,m}) + O^{n,m} \tag{5.18}$$

where

$$M^{n,m} = \frac{1}{2\rho^{n,m}c_n^2}\left(\frac{\overline{\tau}^{n+1,m} - \overline{\tau}^{n,m}}{k_m h}\right) \tag{5.19}$$

$$N^{n,m} = \frac{1}{2\rho^{n,m}c_n^2}\left(\frac{\overline{\tau}^{n,m} - \overline{\tau}^{n-1,m}}{k_m h}\right) \tag{5.20}$$

$$O^{n,m} = a\mathcal{F}^{n,m} - \frac{1}{\rho^{n,m}c_n^2}\left[\overline{W}^{n,m+1}\left(\frac{\overline{\tau}^{n,m+1} - \overline{\tau}^{n,m}}{k_m + k_{m+1}}\right) + \overline{W}^{n,m}\left(\frac{\overline{\tau}^{n,m} - \overline{\tau}^{n,m-1}}{k_{m-1} + k_m}\right)\right] \tag{5.21}$$

and

$$\rho^{n,m} = \frac{(p^{n,m} - p^{n,m+1})}{gk_m}$$

To estimate $\overline{\tau}$ at the point where $\psi^{n,m}$ is held $\overline{\tau}^{n,m}$, $\overline{\tau}^{n-1,m}$, $\overline{\tau}^{n,m-1}$ and $\overline{\tau}^{n-1,m-1}$ must be averaged in some way. For the horizontal average of $\overline{\tau}^{n,m}$ and $\overline{\tau}^{n-1,m}$, rather than just taking $0.5(\overline{\tau}^{n,m} + \overline{\tau}^{n-1,m})$ the expression

$$\frac{c_{n-\frac{1}{2}}}{2}\left(\frac{\overline{\tau}^{n,m}}{c_n} + \frac{\overline{\tau}^{n-1,m}}{c_{n-1}} - a\Omega(c_n + c_{n-1} - 2c_{n-\frac{1}{2}})\right),$$

is used, which is equivalent to averaging $\overline{u}$. This was considered preferable because of the $\cos^2\phi$ term in the definition of τ. If this average is denoted by $\overline{\tau}^{\dot{n},m}$ the LHS of 5.11 can be written in finite-difference form as

$$\begin{aligned}\overline{\tau}_{zt}\overline{\tau} + \overline{\tau}_z\overline{\tau}_t &= \frac{2}{(k_{m-1} + k_m)^2}[(\overline{\tau}_t^{\dot{n},m} - \overline{\tau}_t^{\dot{n},m-1})(k_{m-1}\overline{\tau}^{\dot{n},m} + k_m\overline{\tau}^{\dot{n},m-1}) \\ &\quad +(\overline{\tau}^{\dot{n},m} - \overline{\tau}^{\dot{n},m-1})(k_{m-1}\overline{\tau}_t^{\dot{n},m} + k_m\overline{\tau}_t^{\dot{n},m-1})] \\ &= c_{n-\frac{1}{2}}^2\left[Q\left(\frac{\overline{\tau}_t^{n-1,m}}{c_{n-1}} + \frac{\overline{\tau}_t^{n,m}}{c_n}\right) - R\left(\frac{\overline{\tau}_t^{n-1,m-1}}{c_{n-1}} + \frac{\overline{\tau}_t^{n,m-1}}{c_n}\right)\right]\end{aligned} \tag{5.22}$$

where

$$Q = \frac{1}{(k_{m-1}+k_m)^2}[2k_{m-1}\frac{\overline{\tau}^{n,m}}{c_{n-\frac{1}{2}}} + (k_m - k_{m-1})\frac{\overline{\tau}^{n,m-1}}{c_{n-\frac{1}{2}}}] \tag{5.23}$$

and

$$R = \frac{1}{(k_{m-1}+k_m)^2}[2k_m\frac{\overline{\tau}^{n,m-1}}{c_{n-\frac{1}{2}}} + (k_{m-1} - k_m)\frac{\overline{\tau}^{n,m}}{c_{n-\frac{1}{2}}}] \tag{5.24}$$

On substituting 5.17 into the finite-difference form of the RHS of 5.11, and 5.18 into 5.22 the finite-difference form of 5.11 is finally obtained :-

$$\begin{aligned}
& \psi^{n-1,m-1} && \{RN^{n-1,m-1} - XY\} \\
+ \; & \psi^{n,m-1} && \{R(M^{n-1,m-1} + N^{n,m-1})\} \\
+ \; & \psi^{n+1,m-1} && \{RM^{n,m-1} + XY\} \\
+ \; & \psi^{n-1,m} && \{-QN^{n-1,m} - RN^{n-1,m-1} - X(\frac{2G}{h} + Z)\} \\
+ \; & \psi^{n,m} && \{-Q(M^{n-1,m} + N^{n,m}) - R(M^{n-1,m-1} + N^{n,m-1}) + \frac{4G}{h}\} \\
+ \; & \psi^{n+1,m} && \{-QM^{n,m} - RM^{n,m-1} - X(\frac{2G}{h} - Z)\} \\
+ \; & \psi^{n-1,m+1} && \{QN^{n-1,m} + XY\} \\
+ \; & \psi^{n,m+1} && \{Q(M^{n-1,m} + N^{n,m1})\} \\
+ \; & \psi^{n+1,m+1} && \{QM^{n,m} - XY\} \\
& = && -2hX\{Z\overline{\mathrm{W}} + G\overline{\mathrm{W}}_\phi + C\overline{\mathrm{W}}_z - \overline{\mathrm{p}}c_{n-\frac{1}{2}}(DBA_{\phi t} - a_t\frac{E}{F})\} \\
& + && R(O^{n-1,m-1} + O^{n,m-1}) - Q(O^{n-1,m} + O^{n,m})
\end{aligned} \tag{5.25}$$

where

$$X = \frac{c_p a^2 g\overline{\mathrm{T}}\,\kappa}{4\overline{\mathrm{p}}h\sin\phi}$$

$$Y = \frac{C}{(k_m + k_{m-1})}$$

$$Z = (\kappa - 1)(GE + CD) + G\tan\phi$$

and these terms are all evaluated at the point where $\psi^{n,m}$ is held.

This finite difference equation is solved by the method of successive over relaxation (see Haltiner and Williams(1976)), relaxation being halted when the maximum change in ψ in one iteration altered $\overline{\mathrm{V}}$ by no more than 10^{-5}m/s.

5.2.4 Time-stepping

Once the discretised ψ is found, 5.15 and 5.14 are used to calculate $\overline{\tau}_t$ and $\overline{p}_t$. These are then combined with those of the previous time step, using the Adams-Bashforth method (see Haltiner and Williams(1976)) to give $\Delta\overline{p}$ and $\Delta\overline{\tau}$, the amounts by which the model fields were incremented. The expression for $\Delta\overline{\tau}$ is

$$\Delta\overline{\tau} = \Delta t\{1.5(\overline{\tau}_t)_{current} - 0.5(\overline{\tau}_t)_{previous}\}.$$

The time-step Δt was initially taken to be 4 hours, but the difference in the model fields for a 100 day run on increasing it to 24 hours was insignificant so that a time step of 24 hours was then adopted. $\overline{\tau}$ is updated using the above $\Delta\overline{\tau}$ only on the vertical line at the equator and horizontally at level 16 (about 56km). The finite-difference form of the thermal wind equation, 5.9, is then used to fill in $\overline{\tau}$ over the rest of the grid, to ensure that thermal wind balance is maintained since due to truncation errors $\overline{\tau}$ and $\overline{p}$ might otherwise drift out of balance.

5.2.5 Boundary conditions

At the poles ψ must be constant, since there $\overline{V}$ is zero. At the earth's surface the pressure is taken to be constant in time, so from 5.14 it is seen that $\psi_\phi = \overline{W}$. Since it is also assumed that there is no vertical flux of material across the earth's surface $\overline{W}$ is constrained to be zero there. (It is, however, conceivable for there to be a non-zero $\overline{W}$ at the surface which is balanced by a horizontal eddy flux - see McIntyre(1980)). From definition of w in sigma-coordinates (Equation 5.8) it is seen that our boundary condition at the surface implies that there *is* diabatic heating there but that it is entirely accounted for by the change in surface temperature. Since the top level of the model is well above the stratosphere the upper boundary condition is not expected to have much effect on the stratosphere, $\overline{W}$ and ψ are both set equal to zero there, which therefore means that ψ must be set to zero over the whole model boundary. Above model level 25 $\overline{W}$ is linearly reduced to reach zero at the top boundary.

5.2.6 Ellipticity, and stability fluxes

As noted in section 5.2.2, in order that the 2nd order PDE for ψ has a unique solution, and in order to be able to find this solution using the method of relax-

ation, this PDE must be elliptic. The conditions for ellipticity, and the method of ensuring that they are fulfilled, will now be described.

At first the PDE will be considered for $\theta > 350K$ (in the isentropic coordinate domain). (It is then easy to show that the ellipticity condition when $\theta < 350K$ is equivalent). Substituting 5.14 and 5.15 into 5.11 and noting that when $\theta > 350K$, G$\equiv$ 1, C$\equiv$ 0 and $A_t \equiv 0$, it is found that

$$\psi_{zz}H - \psi_{\phi\phi}I + \psi_z H_z - \psi_\phi IJ + \overline{\tau}L_z + I(\overline{W}J + \overline{W}_\phi) = 0 \qquad (5.26)$$

where

$$H = \frac{\overline{\tau}_\phi \overline{\tau}}{\overline{\rho}\cos\phi}$$

$$I = \frac{gc_p a^2 \overline{T}\kappa\cos\phi}{2\tan\phi\,\overline{p}}$$

$$J = (\kappa - 1)\left(\frac{\overline{p}_\phi}{\overline{p}}\right) + \tan\phi$$

$$L = \frac{1}{\overline{\rho}}(a\cos^2\phi\,\overline{\rho}\,\mathcal{F} - \overline{W}\overline{\tau}_z)$$

Equation 5.26 will be elliptic if $HI < 0$, or equivalently if $\overline{\tau}_\phi \overline{\tau}\,\overline{p}_z\phi > 0$. The ellipticity condition for $\theta < 350$K is equivalent to this since, as can be seen from eqn.5.11, H is as given above and I differs only by a factor of G. A hydrostatic atmosphere is statically unstable if $p_z > 0$ (since this implies that $\theta_p > 0$). It is inertially unstable if $\phi\overline{\tau}_\phi < 0$ or $\overline{\tau} < 0$ (see Palmen and Newton (1969) p508-509). (The case of $\overline{\tau} < 0$ corresponds to the case of a too-strong anticyclone leading to an imaginary solution for $\overline{u}$ in gradient wind theory). Therefore, a statically and inertially stable model atmosphere will lead to an elliptic PDE for ψ. In order to maintain the model atmosphere in a statically and inertially stable state fluxes of momentum and heat are produced whenever any of the three conditions are close to being violated. These heat and momentum fluxes, described below, relax the atmosphere towards a 'safe' state with a time constant of 2 days.

1) $\phi\overline{\tau}_\phi < 0$:- $|\overline{\tau}_\phi|$ is expected to be potentially small only near the equator where planetary vorticity is small (noting that the zonal-mean absolute vorticity is equal to -$\overline{\tau}_\phi/\cos\phi$). Near the equator, $\overline{\tau}_\phi \simeq a\cos\phi\,\overline{u}_\phi$. Therefore an equatorward momentum flux is produced if $|\overline{\tau}_\phi|$ falls below $a\cos\phi\frac{1.5\text{m/s}}{\Delta\phi}$, where $\Delta\phi = \pi/19$, the width of a model box. These fluxes are confined to within 20 degrees of the equator, where they are almost always present.

2) $\overline{\tau} < 0$:- Near the equator, $\overline{\tau}$ is dominated by the planetary angular momentum so $\overline{\tau}$ is expected to become negative only near the pole. If $\overline{u} = -a\Omega\cos\phi$ then $\overline{\tau}$ will be zero, so a poleward momentum flux is created if $\overline{u}$ falls below $-a\Omega\cos\phi + 9\text{m/s}$, which is about -30m/s at the middle of the most poleward box (about 5 degrees from the pole). The 9 m/s above was chosen to be small enough to ensure stability without restricting $\overline{u}$ excessively.

3) $p_z > 0$:- Convection (section 5.4.3) keeps the model atmosphere from becoming statically unstable (static instability could not anyway be represented in our model since this would require the same isentropic level to occur at two heights in the atmosphere). There is, however, another way in which p_z could become positive, and this is if two isentropic levels cross over (eg. if the pressure on a lower level decreases until it is less than that of the one above it. To prevent this happening the lower level is heated and the upper level cooled if the pressure on the lower lever falls below 1.03 times that of the upper level. This occurs only sporadically, where the temperature increases unrealistically rapidly with height (about 10K per km).

5.3 Parametrisation of physical processes

The physical processes included in the model will now be described. There are basic representations of the major processes, and many of the important interactions, although it is hoped to develop them in future as better data or more refined techniques become available.

5.3.1 Treatment of vertical eddy fluxes

In the real atmosphere the main processes which give rise to vertical eddy fluxes of heat, matter and momentum are different for the troposphere, stratosphere and mesosphere. In the troposphere, unstable baroclinic waves, convective overturning and perhaps breaking gravity waves all lead to down-gradient, diffusive eddy fluxes. These processes are strongly non-adiabatic, so the convergence of the fluxes should be very similar in both isobaric and isentropic coordinates.

In the lower stratosphere convection is negligible and planetary scale Rossby waves are the dominant disturbances at mid- and high-latitudes. They can be considered

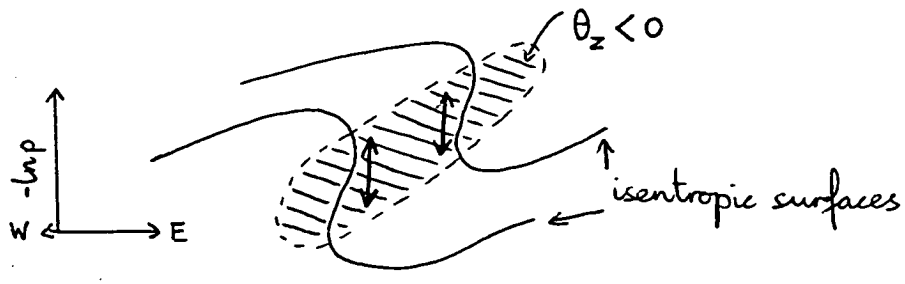


Figure 5.4: *Isentropic surfaces in a breaking gravity wave*

as almost adiabatic and therefore, in isentropic coordinates, their motion will be almost horizontal and they will produce a very small vertical eddy flux. In isobaric coordinates, as noted in Chapter 1, they will have roughly elliptical orbits in the meridional plane and can therefore create a much larger vertical eddy flux. This flux is, however, largely advective in nature and not diffusive. In the upper stratosphere the Rossby waves become less adiabatic, and gravity waves become large and perhaps start to break.

In the mesosphere breaking gravity waves are probably the main cause of a vertical eddy flux. This flux, as in the troposphere, is non-adiabatic and diffusive in nature, and will therefore be similar in both isentropic and isobaric coordinates. A brief description of the way in which breaking gravity waves are thought to produce a vertical eddy flux will now be given, so that the difficulty in parametrising them can be appreciated. As gravity waves propogate vertically their amplitudes grow, and at a certain level (the breaking level) they will produce local statically-unstable regions in the absence of turbulent diffusion (see fig.5.4). By assuming that, above this breaking level, a vertical diffusion of temperature actually prevents the formation of these statically-unstable regions, Lindzen (1981) was able to deduce a spatially-uniform vertical diffusion coefficient. This coefficient can then be used to parametrise the zonal-mean down-gradient flux of heat and constituents (eg. Holton (1983)). However, as noted by Fritts and Dunkerton (1985), this will overestimate the flux if vertical diffusion occurs only *locally* in those regions close to neutral stability (the shaded region in fig.5.4). This can be seen readily on transforming to isentropic coordinates, where the double headed arrows in the fig.5.4 indicating the direction of diffusion, become almost horizontal. The diffusion coefficient required for this local diffusion is very sensitive to the degree of supersaturation of the gravity waves (McIntyre 1987), and hence the value of the vertical diffusion coefficient needed in the mesosphere is very difficult to estimate.

The way in which the vertical eddy fluxes are parametrised in the model will now be described. The zonal mean of the continuity equation for a general quantity, eqn.2.10, is

$$\cos\phi(\overline{\rho}\,\hat{\chi})_t + (\overline{V}\hat{\chi})_y + (\overline{W}\hat{\chi})_z = \overline{S\rho}\cos\phi - (\overline{\rho v\chi^*}\cos\phi)_y - (\overline{\rho w\chi^*})_z\cos\phi \qquad (5.27)$$

Since $\rho\chi^* = \rho\chi - \rho\hat{\chi} = \rho\chi - \overline{\rho\chi} - \rho'\hat{\chi} = (\rho\chi)' - \rho'\hat{\chi}$ we can write

$$\overline{\rho w\chi^*} = \overline{w'(\rho\chi)'} - \overline{w'\rho'}\hat{\chi} \qquad (5.28)$$

Defining a vertical displacement, ξ, such that $w' = \xi_t$ and $\chi' = -\xi\,\overline{\chi}_z$, where χ can be any quantity conserved under a vertical displacement, gives

$$\overline{w'\chi'} = -\frac{1}{2}(\overline{\xi^2})_t\,\overline{\chi}_z = -K\,\overline{\chi}_z \qquad (5.29)$$

where K here is the K_{zz} of (1.16) Substituting 5.29 in 5.28 gives

$$\overline{\rho\,w\,\chi^*} = -K[(\overline{\chi\rho})_z - \hat{\chi}\,\overline{\rho}_z] = -K\overline{\rho}\,\hat{\chi}_z \qquad (5.30)$$

In eqn.5.6, $\overline{w'\rho'}$ can now be parametrised as $-K\overline{\rho}_z$. Thus diffusion alters the vertical velocity in isentropic coordinates. This corresponds to diffusion of $\ln\theta$ in pressure coordinates, since diffusion in either coordinate system is trying to produce a uniform vertical distribution of p_z.

The term involving the vertical diffusion of angular momentum, from 2.16, is $\overline{\rho}\,\overline{w^*\tau_z}$. This can be rewritten as $\overline{\rho}\,(\overline{w'\tau'_z}) - (\overline{\rho'w'})\overline{\tau}_z$, which is then parametrised as

$$\overline{\rho}\,\overline{w^*\tau_z} = -K(\overline{\rho}\,\overline{\tau}_{zz} - \overline{\rho}_z\overline{\tau}_z).$$

K has been estimated in geometric coordinates from GCM data, by Kida(1983) and Plumb and Mahlman(1987), and from satellite data by Lyjak(1987). Values of about $5m^2s^{-1}$ are typical of the troposphere with K decreasing to about $0.2m^2s^{-1}$ in the stratosphere, although Lyjak deduced values of about $1m^2s^{-1}$ during a NH major warming. Garcia and Solomon (1985) use a value of $0.35m^2s^{-1}$ in their stratospheric model. In the mesosphere, K might be typically $100m^2s^{-1}$ (see WMO(1986) p333), although, as noted above, it might be much smaller than this. To estimate K in isentropic coordinates the flux convergence will be assumed to be the same in both isentropic and isobaric coordinates, i.e.

$$\frac{1}{\overline{\sigma}}(\overline{\sigma}\,K^*\hat{\chi}_{z^*})_{z^*} = \frac{1}{\overline{\rho}}(\overline{\rho}\,K\,\hat{\chi}_z)_z$$

where σ, K* and z^* are density, diffusion coefficient and height in geometrical coordinates. Although this might be valid for the mesosphere and troposphere, the vertical flux in the stratosphere should be much smaller in isentropic than in isobaric coordinates, as noted above. In the absence of a direct estimate of K in isentropic coordinates, however, this method was used to obtain K throughout the atmosphere. Since $dz^* = z^*_z dz$ and $\sigma = z_{z^*}\rho$, it follows that

$$K = K^*(z_{z^*})^2$$

From the definition of θ the expression $z_{z^*} = (T_{z^*} + \kappa g/R)/T$ can be obtained. Taking $T_{z^*} \simeq 2Kkm^{-1}$ and $T \simeq 240K$ in the stratosphere gives $z_{z^*} \simeq 5x10^{-5}m^{-1}$. For the troposphere, $T_{z^*} \simeq -7Kkm^{-1}$ and $T \simeq 270K$ implies $z_{z^*} \simeq 1x10^{-5}m^{-1}$. Taking $K^* = 0.35m^2s^{-1}$ in the stratosphere and $K^* = 5m^2s^{-1}$ in the troposphere then results in $K = 8x10^{-10}s^{-1}$ in the stratosphere and $K = 5x10^{-10}s^{-1}$ in the troposphere. In our model, a value for K of $7x10^{-10}s^{-1}$ is used in both the troposphere and stratosphere, which is considered reasonable bearing in mind the uncertainties first in estimating K* and then in transforming to isentropic coordinates. A value of $49x10^{-10}s^{-1}$ is used in the mesosphere (above model level 16).

Although this treatment is far from perfect, especially in the stratosphere, it is an appropriate first step.

5.3.2 Planetary waves in the stratosphere - PV flux

Attempts at having a wave-mean flow interaction in a 2D model have been made by, for example, Hitchman and Brasseur (1988), Pyle and Rogers (1980) and Holton and Wehrbein (1980). Hitchman and Brasseur's model specified the wave activity at the tropopause and then predicted its propogation into a stratosphere whose refractive index depended on the mean state. Pyle and Rogers calculated stationary wave amplitudes on a background mean state, and Holton and Wehrbein's model was in effect a 3D spectral model, severely truncated in the zonal direction. Although these techniques have been fairly successful, it was felt that specifying $\mathcal{F}$ independently of the model state would be both simpler and would also allow the reliability of a $\mathcal{F}$ calculated from satellite data to be asessed. However, if our model is to be used in a predictive role effectively it will be necessary for $\mathcal{F}$ to be parametrised from the model state.

The part of $\mathcal{F}$ (the eddy forcing of $\overline{u}$) due to the planetary waves was assumed to be fairly well approximated by the PV flux calculated from satellite data in the

stratosphere. A full description of the calculated PV flux and its effect on the model will be given in Chapter 7. The estimates of PV flux made in Chapter 8, where it was assumed that the flux is due entirely to dissipation of zonal asymmetries of PV, were also applied to the model stratosphere.

5.3.3 Friction and gravity waves in the mesosphere and stratosphere

For a long time Rayleigh friction was included in the mesospheres of 2D models as the only way to obtain closure of the stratospheric jets and a stratopause. It later became accepted that gravity waves created in the troposphere would tend to break in the mesosphere, accelerating the zonal wind towards their phase speed (see e.g Lindzen 1981). If the gravity waves had a phase speed of zero then this would provide some basis for the use of mesospheric friction. With the inclusion of a spectrum of phase speeds the reversal of the zonal winds above the mesopause has been modelled (Garcia and Solomon 1985), which would not be possible using Rayleigh friction. However, since our interest in this study does not extend to those heights, and since the effect of a $\mathcal{F}$ in the mesosphere would only be expected to influence the top few scale heights of the stratosphere (McIntyre's (1987) 'downward influence' principle), there seems to be no need to use some of the more sophisticated gravity wave parametrisations. Therefore, suitable Rayleigh friction coefficients are used (see Table 5.2) which produce the required closure of the stratospheric jets. Throughout the stratosphere a very small coefficient is used.

Pressure scale height	2.75-7.75	8.25	8.75	9.25	9.75	10.25	10.75	11:25-14.25
Time constant (days)	200	15.4	13.3	7.7	4.4	3.0	1.9	1.3

Table 5.2: *Frictional relaxation time constants*

5.3.4 Radiative Heating in the stratosphere and mesosphere

Below 25km and above the troposphere, a Newtonian cooling parametrisation is used, with the radiative equilibrium temperature taken from WMO(1986) p281 and relaxation times (shown in Figure 5.5) calculated so as to reproduce Dopplick's

heating rates, using the observed temperature field for January of Barnett and Corney (1985a). The actual heating rates in this region of the atmosphere are difficult

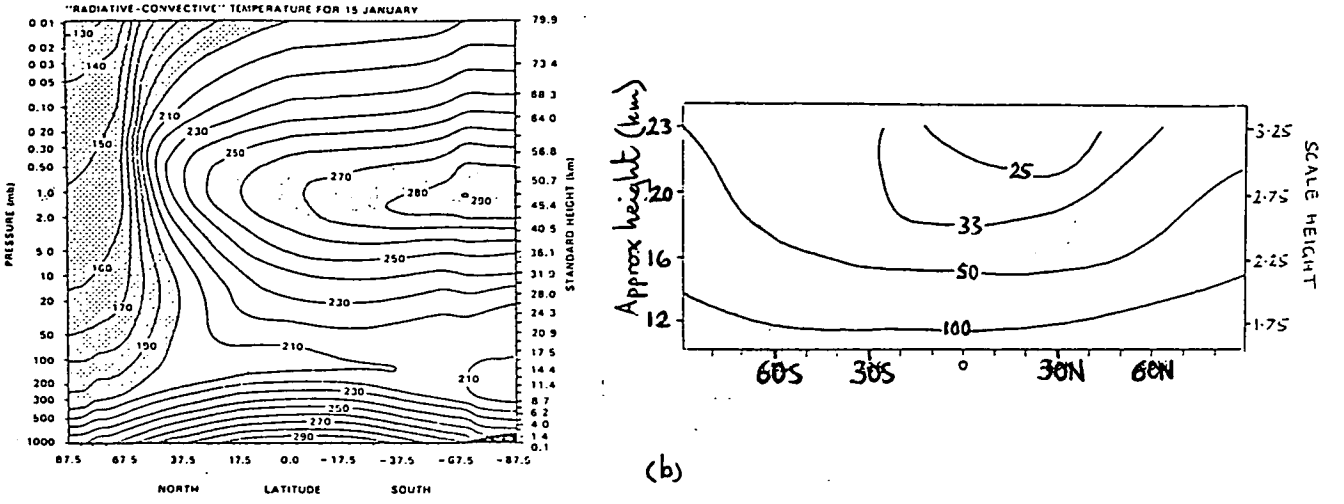


Figure 5.5: *(a) Equilibrium temperature and (b) relaxation times (days)*

to determine, but can have a large effect particularly on the ozone distribution (see Haigh 1984, Harwood and Pyle 1980).

The radiation scheme used above 25km is basically that of Haigh and Pyle (1982), with black-body cooling calculated using the Curtis matrix method for the 15μm band of carbon-dioxide, and a cooling-to-space approximation for the 9.6μm band of ozone. The absorbtion of solar radiation by molecular oxygen, ozone and nitrogen dioxide is calculated using a simple flux-depletion method. The ozone concentrations used by the heating routine were seasonally varying values obtained from McPeter et al's (1984) data set. Molecular oxygen was assumed to have a constant mixing ratio and nitrogen dioxide was specified as a function of height only.

5.3.5 Variation of daylight hours with height

The above scheme was modified to take into account the variation of the hours of daylight with height. This was not considered in the original scheme (nor does it appear to have been considered in any previous two-dimensional calculations of heating rate), but since it seemed fairly simple to carry out, it seemed worthwhile to include it and investigate its effect on the model.

Starting at a point on the surface near the edge of the region of polar night or

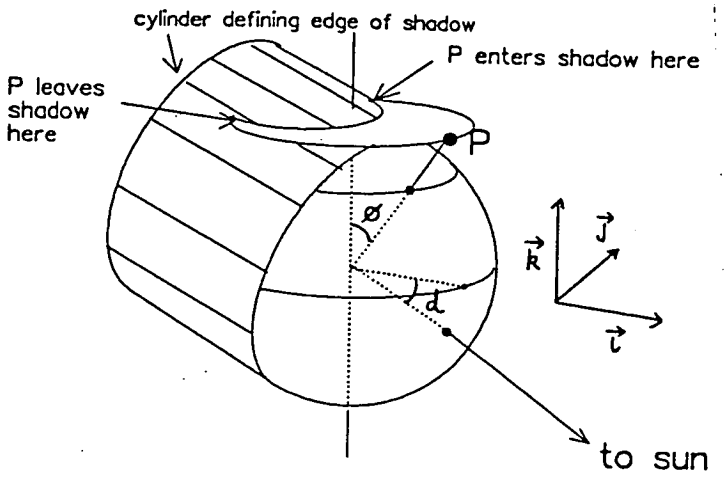


Figure 5.6: *Picture showing the plane P rotates in and where it cuts the earth's shadow*

polar day, the hours of daylight increase quickly with height, the quickest increase being at the equinox, when the atmosphere above the plane tangent to the pole at the surface is in sunlight all the time, while any point on the surface only has half a day of sunlight. The old routine assumed that the number of daylight hours at any point in a vertical column was the same as at the surface.

What is needed is an expression for the hours of daylight per day seen by a point, P, in the atmosphere, rotating with the angular velocity of the earth, which lies h km vertically (ie. radially out from the earth's centre) above a point on the earth's surface at latitude ϕ (see fig.5.6). The following definitions are made:-

- d is the sun's angle of declination (which is the latitude at which the sun is directly overhead at mid-day).
- $\vec{s}$ is the unit vector which points directly at the sun from the centre of the earth.
- $\vec{k}$ is the unit vector pointing from the south pole to the north pole.
- $\vec{\imath}$ is the vector perpendicular to $\vec{k}$ lying in the plane defined by $\vec{k}$ and $\vec{s}$, making an acute angle (which is d) with $\vec{s}$.
- $\vec{\jmath} = \vec{k} \wedge \vec{\imath}$, completing the coordinate vectors.
- $\vec{p}$ is the vector which points at P from the earth's centre.

We can now write

$$\vec{s} = \vec{\imath} cosd + \vec{k} sind \tag{5.31}$$

and

$$\vec{p} = (a+h)\{\vec{\imath} cos\phi cos\lambda + \vec{\jmath} cos\phi sin\lambda + \vec{k} sin\phi\}, \tag{5.32}$$

where a is the earth's radius and λ is the longitudinal angle, defined to be zero when $\vec{p}$ is pointing at the sun. There is a cylindrical region of space 'behind' the

earth which is in shadow (see figure 5.6). To calculate the hours of daylight per day, the two angles, λ_1 and λ_2, at which P falls into, and emerges from, this shadow must be found. If the earth's rotation is neglected so that $\lambda_2 \simeq 2\pi - \lambda_1$ there will be $24\lambda_1/\pi$ hours of daylight per day. Now, P will be on the edge of the shadow when

$$|\vec{p} \wedge \vec{s}| = a, \tag{5.33}$$

in other words when the perpendicular distance from P to the line between the sun's and the earth's centres is equal to the earth's radius (which is a characteristic of any point on the surface of the cylindrical shadow region). Substitution of 5.31 and 5.32 into the square of 5.33 gives, after some calculation,

$$(a+h)^2\{cos^2\phi\, sin^2(d-\lambda_1) + sin^2\phi\} = a^2$$

which can be solved to get

$$\lambda_1 = cos^{-1}\left\{-tan\phi\, tand - \frac{(h^2+2ah)^{1/2}}{(a+h)cos\phi\, cosd}\right\} \tag{5.34}$$

and hence the hours of daylight. Note that if $\cos\lambda_1 < -1$ then P is never in shadow. The hours of daylight are used in two ways when calculating the day- and zonal-mean heating rate. Firstly, in order to estimate the average heating rate when P is illuminated, a 5 point Gaussian integration is performed, which involves calculating the heating rates for a column of air at five different longitudes, when $\lambda = 0, \pm 0.538\lambda_1, \pm 0.906\lambda_1$. In fact, because of the symmetry about $\lambda = 0$ it is only necessary to calculate the heating rate at three longitudes. Because it is not possible to have a λ_1 which varies with height for this calculation the value of λ_1 at about 40km (which is near the ozone maximum) is used at all levels. The second way in which λ_1 is used is in finding the average daily heating rate, which is simply the Gaussian-averaged heating rate found above multiplied by the hours of daylight divided by 24. For this calculation the correct λ_1 is used at each height.

However, at latitudes where the surface is in polar night, the whole atmospheric column is treated as being in the dark, for the following reason. In the real atmosphere there will be upper levels at these latitudes which do intercept rays from the sun. The rays in this region will, however, be travelling away from the earth's surface (see figure 5.7), and cannot be dealt with using the present scheme, which follows a ray down from the top of the atmosphere to the earth's surface.

The model was run once with the old scheme and once with the new scheme and the differences (new minus old) in the heating rates, temperatures and zonal winds

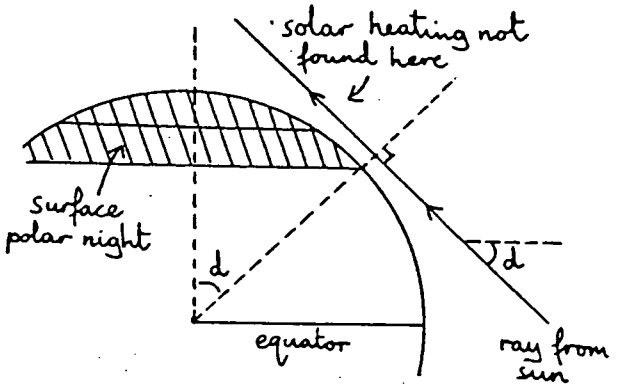


Figure 5.7: *Region above surface polar night where rays are climbing and cannot be dealt with*

for June and September are shown in Figure 5.8. In June the biggest increases in heating rates are at about 70S ($2°\text{Cday}^{-1}$) and 70N ($0.7°\text{Cday}^{-1}$), the bigger increase being near the polar night. The cooling nearer the poles is due to the general rise in temperature. In September, the greatest increases are at the poles, with $5°\text{Cday}^{-1}$ at the south pole and $1.5°\text{Cday}^{-1}$ at the north. These are significant changes, especially in September, when they are of the same order of magnitude as the total heating rate near the pole (see Figure 6.5 and the related text). As can be seen, temperature increases of 10K or more occur near the poles, along with zonal wind changes of the order of 10ms^{-1}.

5.3.6 The Troposphere

Even though the primary purpose of this thesis is to study stratospheric dynamics, it was felt it would be worthwhile to have a model troposphere which would interact with the stratosphere, to provide at least some idea of the influence stratospheric changes (eg. in ozone, PV flux) would have on the troposphere, and vice versa. More importantly, perhaps, when the model is extended to include chemistry, the troposphere will act as the source or sink of many stratospheric species. At that time the need for as realistic a cross-tropopause flow as possible will be even more important.

In the real troposphere, the total heating rate is mainly a combination of emission and absorbtion of thermal IR radiation and absorbtion of solar u.v. radiation (by ozone, oxygen, water and carbon dioxide in particular), the latent heat released by water-vapour condensing in cumulonimbus clouds, and the turbulent flux of

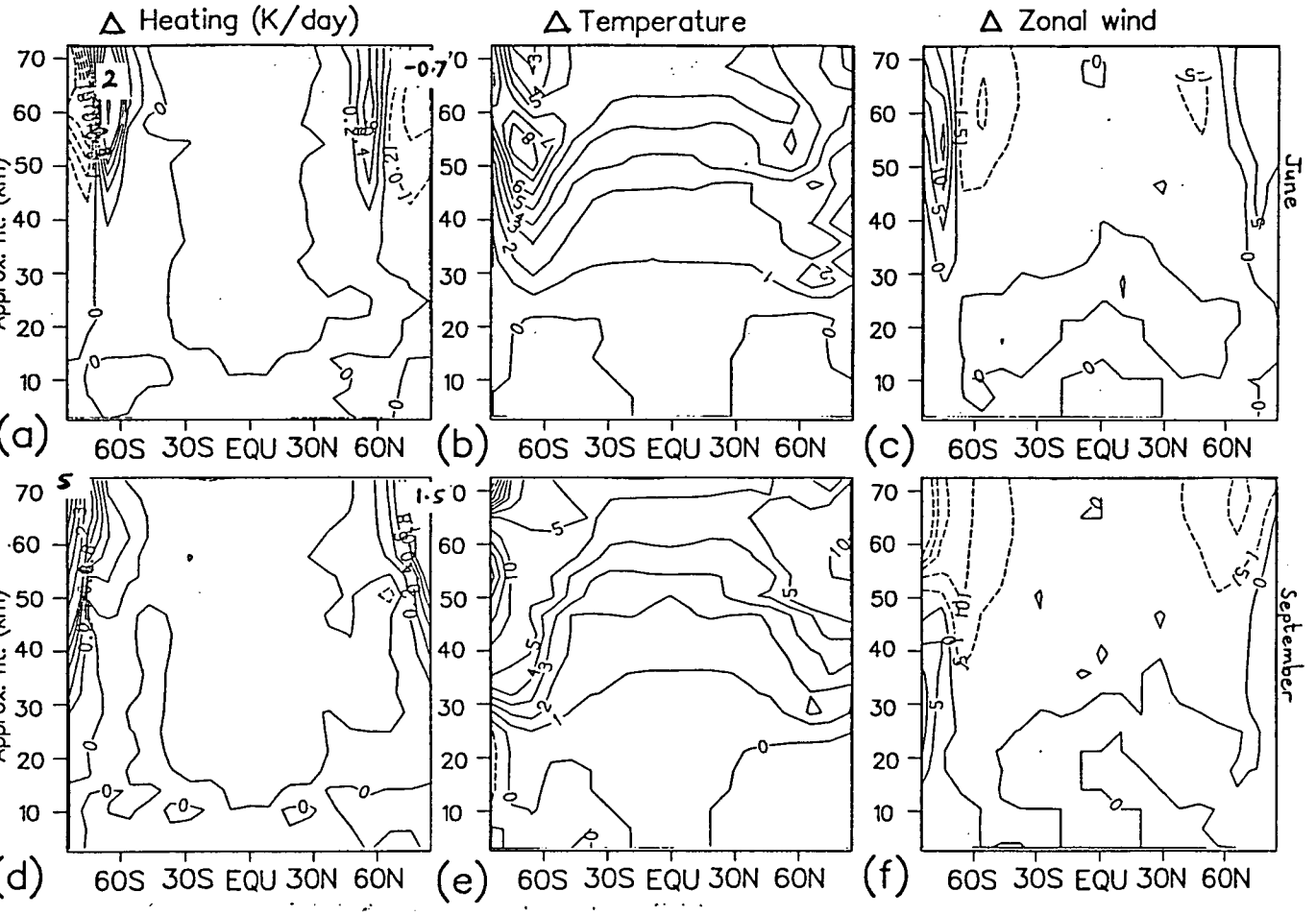


Figure 5.8: *Change in heating rate (contours at 0.2,0.4,..,1.0,2.0,..* °Cday^{-1}*), temperature (K) and zonal wind (m/s) for June (top) and September (bottom) due to the change in daylight hours with height. Dashed contours indicate negative values.*

sensible heat in the boundary layer. The sources of momentum are turbulent friction with the ground and mountain torque, which is due to the displacement of air by the mountains. Waves transport momentum from their source region to the region where they are dissipated. Their source might be on the earth's surface or in the free atmosphere (for example gravity waves may be created by storm clouds, and planetary waves by surface air temperature patterns) thus the waves might transfer momentum from the earth to the atmosphere, or vice versa, or redistribute the momentum of the atmosphere. Similarly, convective motion redistributes potential temperature. All these effects are important to the working of the real troposphere and must be parametrised in some way in our model's representation.

5.3.7 Radiative Cooling in the troposphere

Two calculations of the combined effect of the main radiative gases are shown in Figure 5.9. They include the cooling due to black-body radiation and the heating due to absorbtion of solar radiation. As can be seen, in their details they vary considerably. This is mainly due to the uncertainty in the quantity and effect of clouds. A broad summary of their common features is an average cooling

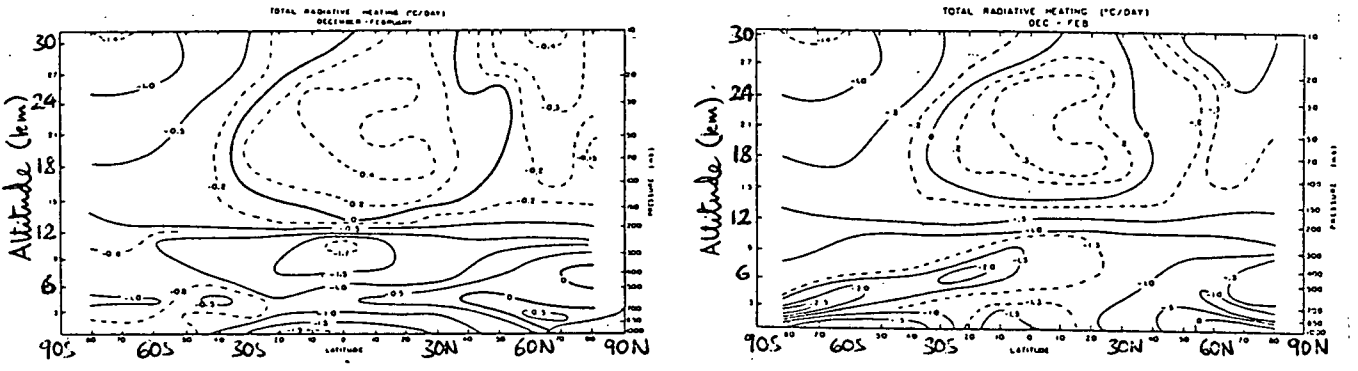


Figure 5.9: *Tropospheric radiative heating rates (Dec-Feb). Left: Newell et al. (1972), Right: Dopplick(1979)*

rate of about 1 degree per day throughout the troposphere, though increasing by about 50% near the equator, with a winter hemisphere's rate exceeding that of the summer hemisphere again by about 50%. The rates all fall to near zero at the tropopause. Because of the variation among the values it seemed pointless to use the exact values of one set in preference to another, so our values were chosen to lie somewhere amongst them.

The rates specified poleward of 20N and 20S did not vary with latitude but varied sinusoidally in time, with extremes at the solstices. These values are shown in Table 5.3. The rates equatorward of 20 degrees were the average of those poleward of 20 degrees and so did not vary with time. An improvement on having a fixed

Winter	Summer	level
-0.3	-0.15	4
-0.75	-0.45	3
-1.2	-0.6	2
-1.2	-0.6	1

Table 5.3: *Tropospheric radiative heating rates used in model (in K/day)*

cooling rate would be to have a cooling rate which depended on the tropospheric temperatures, which would provide another feedback mechanism, but which would also require more approximations and assumptions (such as a an equilibrium temperature and a radiative time constant).

5.3.8 Convection

Convection occurs in the model when the lapse rate exceeds a certain critical value. However, unlike most 2D models, which have a critical lapse rate of 6.5K/km, independent of latitude (which is close to the observed tropical lapse rate - see, for example, Houghton (1986)), the critical lapse rate was chosen here to vary with latitude, as shown in Table 5.4, since in the tropics, where the air is moist, the air becomes unstable at a lower lapse rate than in the drier extra-tropical regions. The lapse rate in the tropics is determined by convection, whereas in the extra-tropics the air is generally convectively stable, and the lapse rate cannot be maintained by convection - the air aloft must be being warmed by the heat transport of planetary waves. The critical lapse rates, chosen to give reasonable results, confine convection to the tropics (except in the model's bottom layer). If the lapse rate

latitude	0.0	9.5	18.9	28.4	37.9	47.4	56.8	66.3	75.8	85.3
c. l. r.	5.5	5.5	7.5	8.0	8.5	8.5	8.5	8.5	8.5	8.5

Table 5.4: *Critical lapse rates used in model (in K/km)*

across a certain layer exceeds the critical lapse rate at that latitude an upward heat flux is created by increasing $\overline{W}$ at the top of that layer, and decreasing it below, to relax the lapse rate towards the critical one with a time constant of 2 days. Since $\overline{W}$ is increased on the upper level by the same amount as $\overline{W}$ on the lower level is decreased, there will be a small creation or destruction of potential energy, but because of the crudeness of the representation of the heat sources and sinks in the troposphere this was not felt to be important.

5.3.9 Surface Heating - sensible and latent heat fluxes

The surface potential temperature is specified and is seasonally varying, with the extremes at mid-January and mid-July, as shown in Figure 5.10

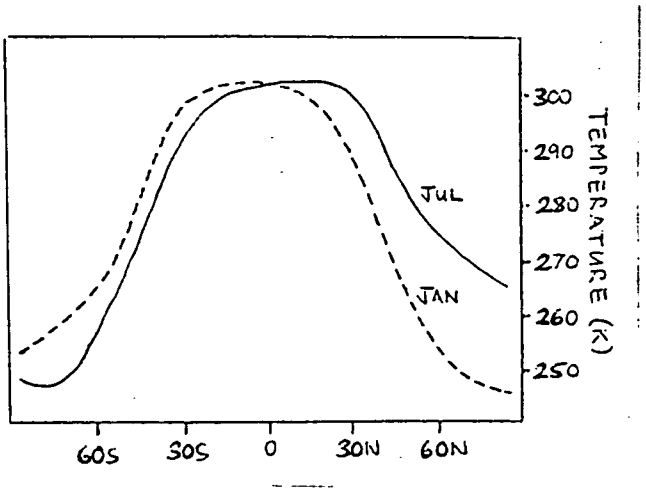


Figure 5.10: *January and July surface temperatures*

The values were derived from the SSU satellite data used elsewhere in this thesis, by extrapolating to 1000 mb using the local lapse rate. The surface temperature is not affected by the rest of the atmosphere but affects it through convection in the bottom model layer. Though in the rest of the model convection involves a heating at the top of a box and cooling at the bottom, at the surface the boundary condition of $\overline{W} = 0$ means that convection there produces only a heating at the second model level. This is supposed to crudely represent sensible and latent heat fluxes from the surface. Convection in the rest of the atmosphere then redistributes this heating.

5.3.10 $\mathcal{F}$ in the troposphere and Surface Friction

Calculations of ∇.F in the troposphere from models (Plumb and Mahlman (1987), Boville (1986)) and observations (Randel (1987) and Edmon et al. (1981)) and of $\mathcal{F}$ by Yang et al.(1990) show a negative region from about 3 to 10 km and 20° to 70°, with a minimum of about -15ms^{-1}day^{-1} in winter and a maximum of about -10ms^{-1}day^{-1} in summer, and a positive region near the surface with a typical value of about 10ms^{-1}day^{-1}. However, there is a large variation amongst the different studies, with Yang et al.'s winter minimum being about -20ms^{-1}day^{-1} and Randel's being about -10ms^{-1}day^{-1}. Since the negative values are roughly correlated in time with the jet velocities they are parametrised here as Rayleigh friction, with a time constant which varies with latitude, as shown in Table 5.5. The small, but non-zero, coefficient at level 5 is supposed to represent the breaking of gravity waves at the tropopause (see Tanaka (1986)). Since the size of this effect is uncertain from observations, the level 5 coefficient was simply chosen to give a

latitude	0.0	9.5	18.9	28.4	37.9	47.4	56.8	66.3	75.8	85.3
Levels 2 to 4	0.24	0.24	0.24	0.32	0.4	0.48	0.64	0.8	0.8	0.8
Level 5	0.09	0.09	0.09	0.12	0.15	0.18	0.24	0.3	0.3	0.3

Table 5.5: *Raleigh friction coefficients (days^{-1})*

reasonably realistic separation of the tropospheric jet from the winter stratospheric jet. The variation with latitude was chosen after experimentation to obtain reasonable zonal wind and $\mathcal{F}$ fields. In the bottom layer two contributions to $\mathcal{F}$ are included. One is a specified positive $\mathcal{F}$, whose values are shown in Table 5.6. This

latitude	0.0	9.5	18.9	28.4	37.9	47.4	56.8	66.3	75.8	85.3
$\mathcal{F}$ (ms^{-1}day^{-1})	0.0	0.0	2.6	5.2	9.1	9.1	9.1	7.8	5.2	2.6

Table 5.6: *Positive $\mathcal{F}$ specified in level 1*

structure of $\mathcal{F}$ implies that eddies are transporting angular momentum downwards, their production in the lower troposphere (whether by orography or convection) accelerating the zonal wind and their dissipation higher up decelerating it. The other contribution to $\mathcal{F}$ in the bottom layer is Rayleigh friction, with a time constant of 2 days, which is supposed to be a parametrisation of surface turbulence and mountain drag.

5.4 Summary

In this chapter a 2D model of the atmosphere from 0 to about 100km was described. Its coordinates above the $\theta = 350$K surface are isentropic while below, sigma-type coordinates are used to avoid intersection of coordinate surfaces with the ground. A stream-function is used to determine the meridional circulation which, when combined with the dynamical forcing due to $\mathcal{F}$ and the diabatic heating, ensures that thermal-wind balance is maintained between the $\overline{u}$ and pressure fields as they evolve in time. A time step of 1 day is used with an Adams-Bashforth scheme, and the model boxes are about 10 degrees wide, with the vertical resolution, which is able to be varied in $\ln(\theta)$ coordinates, chosen to give a box height of about 3.5km.

The parametrisation of some of the real physical processes in the troposphere is simple, but allows an interaction with the stratosphere. These processes are; convection, latent and sensible heat flux from the surface, vertical heat and momentum

eddy fluxes above the surface which are parametrised as diffusion, a $\mathcal{F}$, part of which is specified, part of which is parametrised as Rayleigh friction, and a specified radiative heating. In the stratosphere and mesosphere, radiative heating is calculated using a scheme which is basically that of Haigh and Pyle (1982), with solar heating found from a flux-depletion method and long-wave cooling found by using a Curtis matrix. The modification to their scheme made in this thesis is the inclusion of the effect of the variation of daylight hours with height. The changes to the $\overline{T}$ field due to this modification are largest at the equinoxes, where $\overline{T}$ increases by over 5K in large regions near the poles above 40km. The momentum deposition due to breaking gravity waves in the mesosphere is parametrised as Rayleigh friction. Vertical eddy fluxes of heat and momentum are parametrised as diffusion, with a diffusion coefficient which is small in the stratosphere but large in the mesosphere.

The behaviour of this model will be described in Chapter 6, and the effect on the model of two different estimates of the stratospheric $\mathcal{F}$ due to planetary waves will be presented in Chapters 7 and 8.

Chapter 6

Control run and effect of changes in tropospheric parametrisations

How big is it? - well, I don't know!
What if it's twice as small, or, say
not half as big? How will it show?
- and does it matter anyway?

In this chapter results are presented from a version of the model whose stratospheric $\mathcal{F}$ is due only to the very small Rayleigh friction there. A PV flux due to planetary waves has *not* been included. Comparison of model results with observations then reveals, within model limitations, the effect of planetary waves on the real atmosphere.

The other purpose of this chapter is to describe the results of variations in the tropospheric representations of radiative heating, convection, diffusion and $\mathcal{F}$. Bearing in mind that quantities calculated from observations, such as heating rates and $\mathcal{F}$, often have fairly large uncertainties associated with them, such perturbations show us the effect on the model's dynamics of variations in these quantities within the bounds allowed by uncertainty.

6.1 Model run without stratospheric $\mathcal{F}$

The model version used in this section is exactly as described in Chapter 5 but with $\mathcal{F}$ in the stratosphere being solely due to the small Rayleigh friction there, the PV flux due to planetary waves not being included.

Similar experiments have been performed, both with mesospheric drag (e.g. Holton 1983, Schneider et al. 1989, Schoeberl and Strobel 1978) and without it (Shine 1987) to deduce the state of an atmosphere in which wave effects are absent. Departures of observations from modelled winds and temperatures in these studies were then attributed to the effects of these waves. In the models which included a mesospheric friction, solstice winds were fairly well reproduced while large departures were observed at the equinoxes.

6.1.1 Comparison with observations

The model results shown in this section are the monthly-averaged $\overline{u}$ and $\overline{T}$ for July and October of the third year of the run, and have been extrapolated to pressure surfaces. The model was started on February 20th with a temperature structure dependent only on height. By the third year an annual periodicity had set in, and July and January were very nearly mirror images of each other (the $\overline{u}$ fields differing by less than 5 m/s) as were April and October, the small asymmetries being attributable to the asymmetries in the specified ozone concentrations, the earth's orbit and the surface temperature. Only the modelled July and October fields are shown, then, since they represent closely the fields for all four seasons. First, the stratospheric $\overline{u}$ and $\overline{T}$ fields will be compared with those of of Barnett and Corney (1985a), which were derived from 10 years of satellite data. Then the tropospheric fields will be briefly considered.

We shall start with the summer and autumn since these are the seasons when the observed inter-hemispheric differences (IHDs) are smallest. The approximate height scale used in the graphs, of both modelled and observed fields, is ln(1000/p) x 7km. In summer, the observed easterlies extend down to about 18km in both hemispheres (fig. 6.1(c) and (d)) and reach a peak of over -60 m/s at 65-75km and 40-50 degrees N or S. In the model, however, (fig. 6.3(a)) the easterlies do not extend much below 40km and peak near the equator at about 60km. This difference between modelled and observed $\overline{u}$ is accompanied by a discrepancy in $\overline{T}$, the model stratopause (fig. 6.3(c)) being about 10°C warmer than observed (fig. 6.2(c) and (d)) at the pole (though it occurs at about the right height) and the lower stratosphere being about 20°C too cold.

The observed autumn $\overline{u}$ (fig. 6.1(a) and (b)) shows a larger IHD than the summer $\overline{u}$, the SH westerly jet maximum of about 70 m/s being about 15 m/s stronger

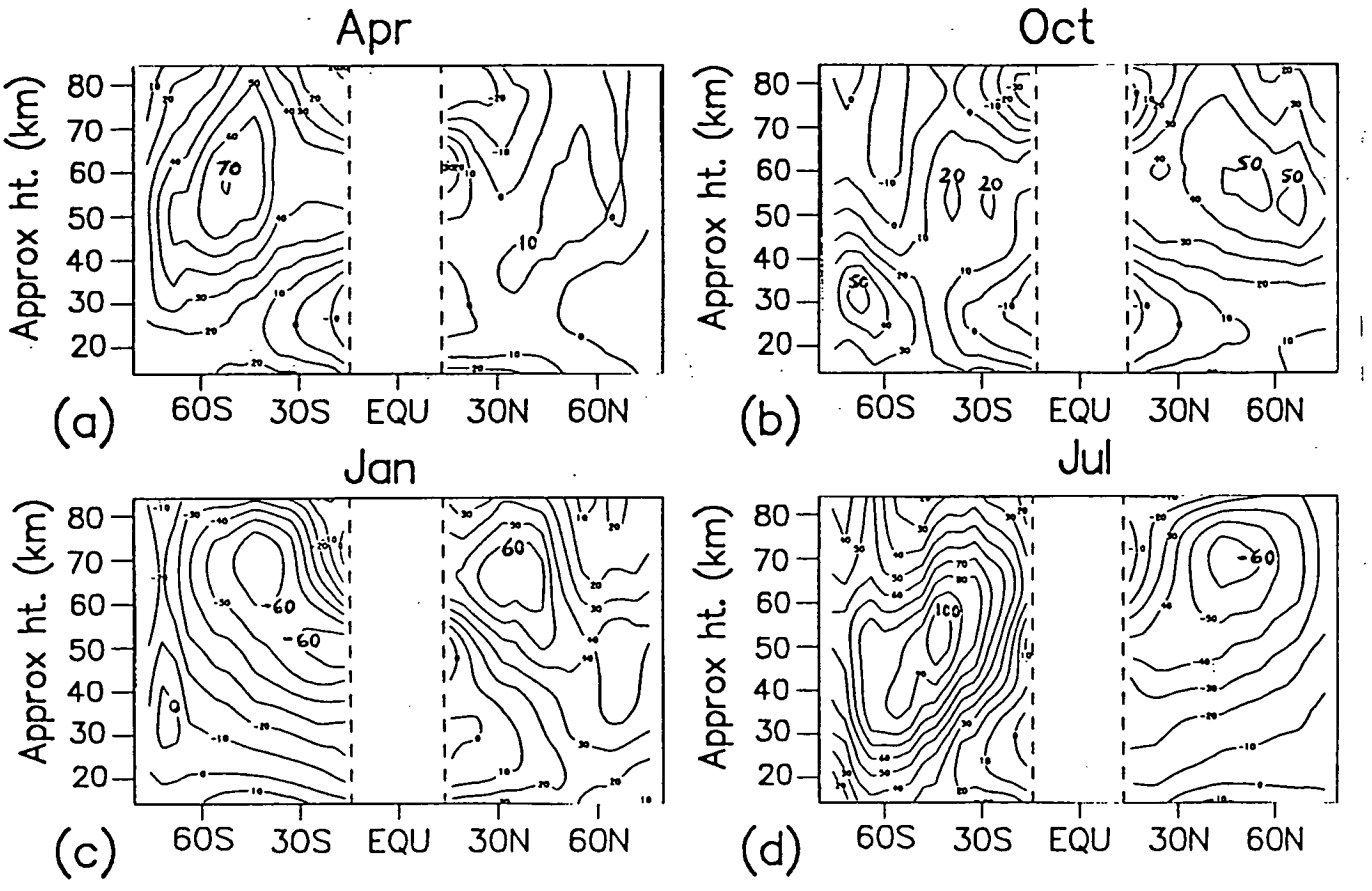


Figure 6.1: *Observed* $\bar{u}$ *for (a) April, (b) October, (c) January and (d) July. (from Barnett and Corney (1985a)*

than that of the NH. The modelled $\bar{u}$ and $\overline{T}$ fields (fig. 6.3(b) and (d)) are very similar to the observed NH fields, although the stratopause is too warm by about 10°C.

The observed winter hemispheres are very different from each other, the westerly jet in the SH being about 40 m/s stronger, about 10 degrees closer to the pole, and about 15 km lower than that of the NH. Thus the NH $\bar{u}$ has been decelerated with respect to the SH $\bar{u}$ from about 40N to 70N and from about 20 to 60km. The modelled winter westerly jet has about the same strength as that of the observed SH jet, but occurs about 5km too high and about 5 degrees closer to the pole. Corresponding to these winds the observed SH polar minimum in $\overline{T}$ is about 25°C lower than that of the NH, and the stratopause is about 15°C warmer. Again, the model is closer to the SH, with the polar minimum $\overline{T}$ falling just below 180K, as in the observed SH, though it occurs at about 40km as opposed to 30km. Above about 60km , near the pole, the model is warmer than both the observed

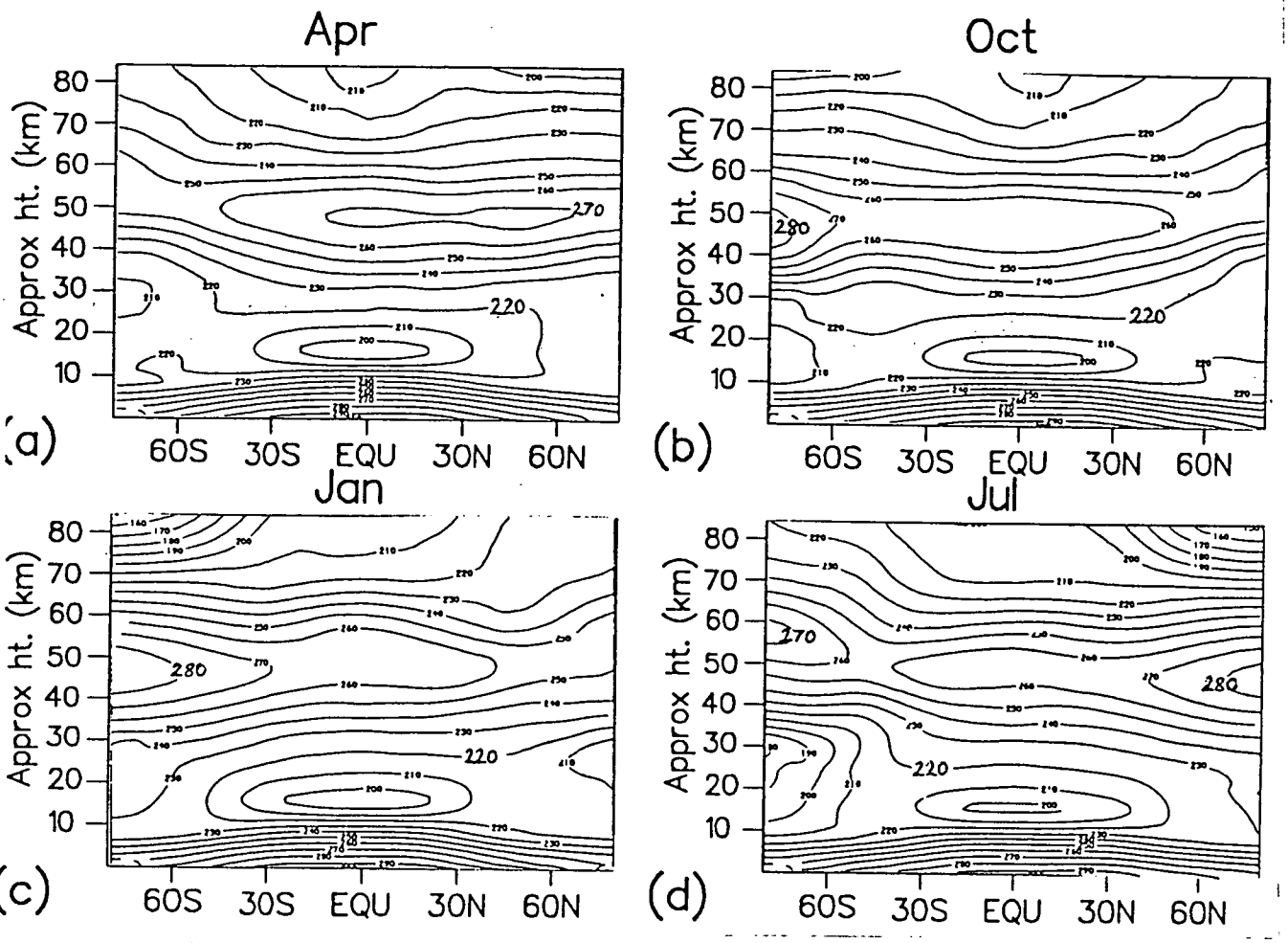


Figure 6.2: *Observed* $\overline{\mathrm{T}}$ *for (a) April, (b) October, (c) January and (d) July. (from Barnett and Corney (1985a)*

hemispheres, and colder throughout the 20km or so below 60km. This model deficiency could probably be cured by having an additional negative $\mathcal{F}$ at about 60km (see fig. 4.6(c) and the accompanying text), and so is presumably due to the over-simple parametrisation of gravity-wave drag in the model. Although not shown, the observed SH westerly jet decreases by about 20 m/s from June to July, while the model's jet increases by about 10 m/s.

The observed spring hemispheres are also very different from each other, and from the model spring. In the SH the westerly jet has been pushed down and polewards compared with the model, indicating a region of negative $\mathcal{F}$ during early spring in the real SH at mid-latitudes and above 30km. The observed NH westerly jet has virtually disappeared, but its deceleration from 60 to about 0 m/s is less than that in the SH but greater than that in the model, indicating perhaps a negative $\mathcal{F}$ in

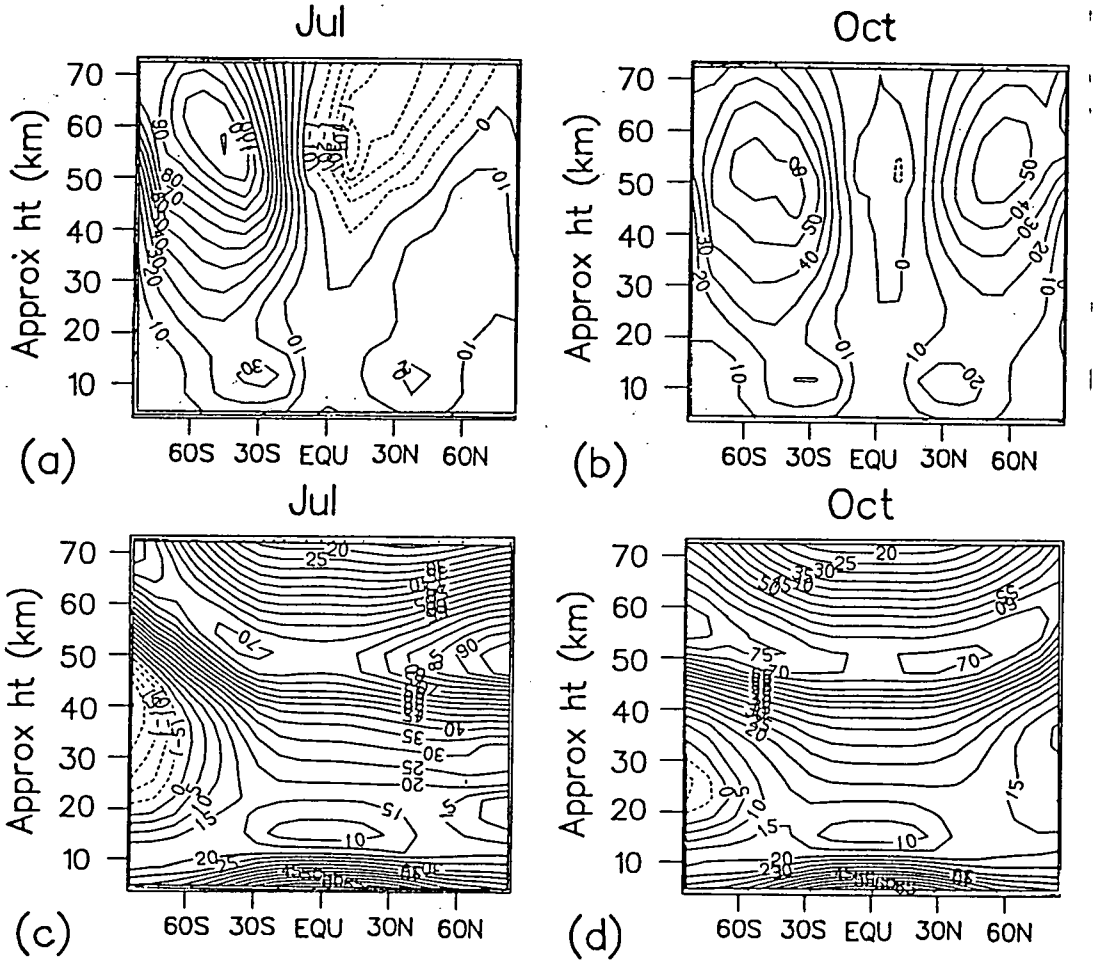


Figure 6.3: *Modelled* $\overline{u}$ *for (a) July and (b) October, and modelled* $\overline{T}$*-200K for (c) July and (d) October. Dashed contours indicate negative values.*

the NH over early spring which is smaller than that in the SH.

The modelled tropospheric temperatures may be compared with those of Barnett and Corney. The model reproduces the observed tropopause minimum over the equator but is about 5K too warm. It remains at about 205K throughout the year, while the observed minimum has an annual cycle, being warmest in August at about 202K and coolest in January at about 197K (this was remarked upon in section 4.5.2). Observational values of tropospheric winds are taken from Newell et al. (1972). They calculate the seasonal averages, which may be summarised as follows. In the NH the westerly jet decreases from a maximum of over 35 m/s in winter to a minimum of about 15 m/s in summer while moving north from 30N to 45N. In the SH the jet decreases from 30 m/s in winter to about 20 m/s in summer while moving polewards from just under 30S to about 40S. Winds below

about 15km are predominantly easterly within about 10 degrees of the equator, reaching a peak value of about -8 m/s during June-August. As noted before, the model shows little inter-hemispheric difference. The westerly jet resembles more closely the SH jet than the NH jet, decreasing from over 30 m/s in winter to about 20 m/s in summer, and also moving polewards from about 30 degrees to about 40 degrees. The winds near the equator, however, never become easterly, except within a few km of the surface.

To summarise :-

(1) The model autumn is similar to the observed autumn hemispheres, and its winter is close to the observed SH winter. These are times when the PV flux in the real stratosphere is expected to be small.

(2) The deceleration of the observed SH westerly jet from June to July is not reproduced by the model, however, so as well as indicating a possible negative PV flux in the SH in June it also implies that the similarity between model and observations in July was slightly fortuitous.

(3) The summer hemispheres are not modelled well (with westerlies remaining in the too-cold model lower stratosphere), even though planetary waves are small in the real summer hemispheres. The reason for this is perhaps insufficient heating in the model lower stratosphere over late spring/early summer due to the simple heating parametrisation used there, though it might also be due to the presence of a negative PVF in the real lower stratosphere during spring and summer. We note that other models which have realistic heating rates but do not include stratospheric $\mathcal{F}$ (Schneider et al. 1989, Holton 1983) have also had summer easterlies which do not extend down to the troposphere. Some models (eg. Hitchman and Brasseur 1988, Garcia and Solomon 1985, Holton 1980) have produced easterlies in the summer lower stratosphere without any wave driving in the stratosphere, but they employ a cooling routine which relaxes the temperature towards an observed profile producing unrealistic heating rates, as noted by Holton (1983).

(4) The NH winter and both spring hemispheres are not reproduced well by the model, from which it is deduced that a negative PVF occurs in the real NH from winter till spring, and in the SH during spring. A poor representation of gravity wave drag near the stratopause could also be a cause of the discrepancy between model and observations, especially in winter.

(5) The troposphere is generally well reproduced, except for the observed equatorial easterlies.

6.1.2 Heating rates, $\mathcal{F}$ and meridional circulation

A brief description of the remaining important model fields will now be given, for future reference. Figure 6.4 contains the heating fields for June 20 and September 20. Several calculations have been made of stratospheric heating rates (Callis et al.

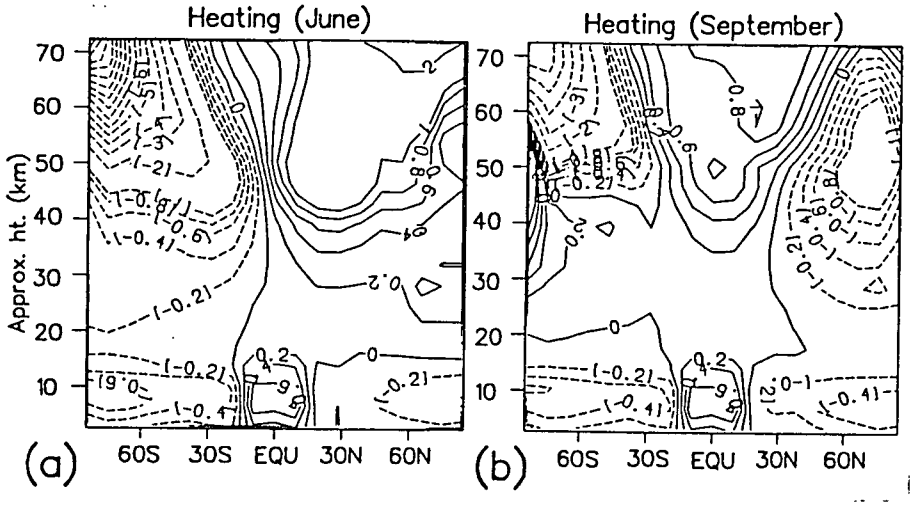


Figure 6.4: *Heating rates for (a) June and (b) September (contours at 0.2,0.4,..,1.0,2.0,..* $^{\circ}\mathrm{Cday}^{-1}$*). Dashed contours indicate cooling.*

1987, Kiehl and Solomon 1986, Rosenfield et al. 1987, Gille and Lyjak 1987) and our rates compare quite favourably with these, to within a few K/day, although there is considerable variation between the studies. There is a strong cooling in winter, of about 10K/day at about 60km over the pole, and a weaker heating near the summer pole. However, the local maximum of about 3 K/day in the summer hemisphere near the equator found in the above studies is not so pronounced in our heating rates. In spring and autumn there is also qualitative agreement, with cooling over both poles, separated by equatorial heating. In comparison with these studies our cooling is stronger over the spring north pole at about 70km, consistent with a warmer-than-observed spring and weaker over the autumn pole by a few K/day. This weaker cooling and the patch of heating over the spring pole at about 45km in fig.6.4(b) is due, at least in part, to the inclusion of the effect of variation of daylight hours with height (see section 5.3.5) which is not considered in the other studies.

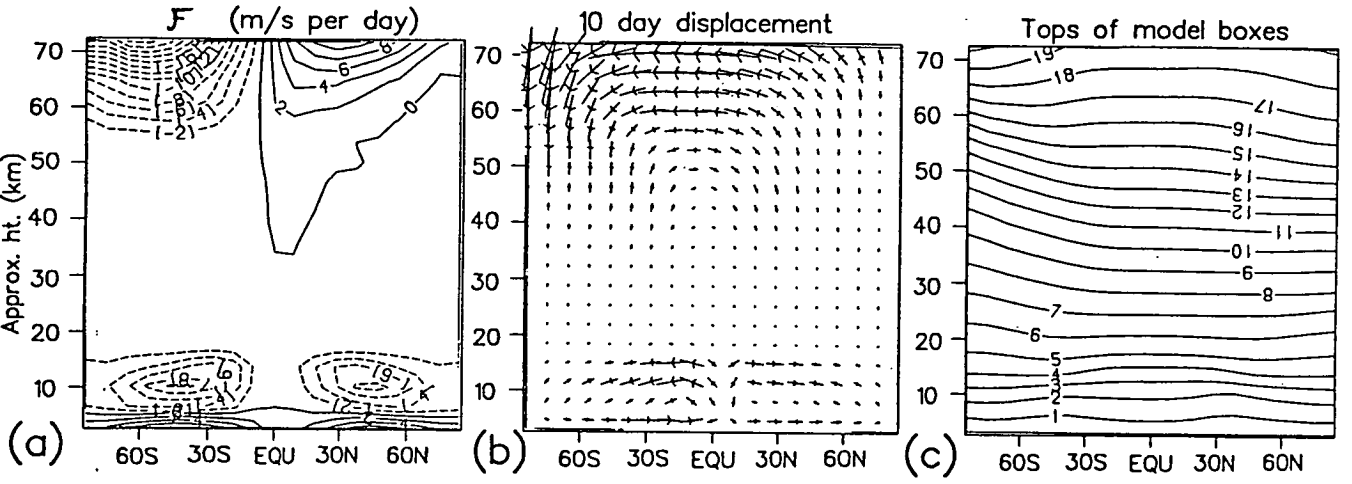


Figure 6.5: *(a)* $\mathcal{F}$ *in* $\mathrm{ms}^{-1}\mathrm{day}^{-1}$, *(dashed contours indicating negative values) (b) the meridional circulation (10 day displacement) and (c) the heights of the tops of the model levels for June*

The $\mathcal{F}$ field for June 20 shown in figure 6.5(a) shows the specified positive flux near the surface with the Rayleigh frictional negative values in the upper troposphere. They are of the same general size as those calculated in other studies although, as noted in section 5.3.10, there is a large spread of values between the studies. They have the observed seasonal dependance, being stronger in the winter than in the summer. In the mesosphere the large $\mathcal{F}$ due to the frictional parametrisation of breaking gravity waves is apparent. The meridional circulation (fig. 6.5(b)) shows the stronger Hadley cell in the winter hemisphere and strong descent in the mesosphere over the winter pole. Comparing the size of the arrows, which indicate a 10 day displacement, with the height of a model box (fig 6.5(c)) we see that a parcel will pass through a model level in the mesosphere in about 3 days. Thus the model remains stable with a 1 day time step, but if the box height were to be reduced significantly then numerical stability problems would no doubt be encountered.

6.2 Sensitivity to Tropospheric Changes

Because there are large uncertainties in most of the parameters used in our representation of the tropospheric processes (see section 5.3), it is useful to know how sensitive the model is to these uncertainties. The way in which the various processes in the model troposphere balance each other will first be described, so

that the shift in balance on changing one of these processes may be understood. In the absence of a tropospheric $\mathcal{F}$ the only specified forcing in the troposphere would be due to the radiative cooling and the surface heating, both of which are time-dependent. Together, these would act to increase the lapse rate. This increase would be balanced by the representation of convection, which transports heat upwards from the surface if the lapse rate exceeds a critical value (see section 5.3.8). The balance would not be exact however, the small residual being due to the time-dependence of the specified radiative cooling and surface heating. The mean meridional circulation would therefore be small. Let us now consider the effect of the specified positive $\mathcal{F}$ in the model's bottom layer (see section 5.3.10). To maintain thermal wind balance in the presence of this $\mathcal{F}$ the model induces an equatorward mean flow in the bottom layer (see section 4.5.1). This adiabatically cools the lower troposphere near the equator and heats it near the poles, thus stimulating convection near the equator and damping it near the poles. The meridional circulation is closed by a polewards flow in the upper troposphere (see fig. 6.5(b)). The Coriolis force on $\overline{u}$ due to this circulation is balanced by the negative $\mathcal{F}$ due to Rayleigh friction (see fig. 6.5(a)). The positive $\mathcal{F}$ near the surface produces surface westerlies which, because of the surface friction, transfer angular momentum from the atmosphere to the ground at mid-latitudes, while easterlies remain at the poles and equator and transfer momentum from the ground to the atmosphere.

The following subsections will describe the effect of changing some of the parameters in our representations of the tropospheric processes. Four changes will be considered :-

6.2.1) Doubling the tropospheric frictional time constants in layers 2 to 5 (i.e. reducing the friction).

6.2.2) Multiplying the specified tropospheric radiative cooling by a factor of 1.5.

6.2.3) Increasing the tropical critical lapse rate from 5.5 to 6.5 $^{\circ}Ckm^{-1}$.

6.2.4) Dividing K_{zz} for the troposphere and stratosphere by a factor of 7.

The conclusions drawn from these experiments will be given in section 6.3.

6.2.1 Reduced tropospheric friction

This subsection compares the control run just described with a run using a version of the model identical apart from the Rayleigh frictional time constants in the troposphere (in model levels 2 to 5), which are doubled in size.

The changes in $\mathcal{F}$, zonal wind, circulation, temperature and heating are shown in Figure 6.6. A decrease in the friction leads to a less-negative $\mathcal{F}$ in the upper tro-

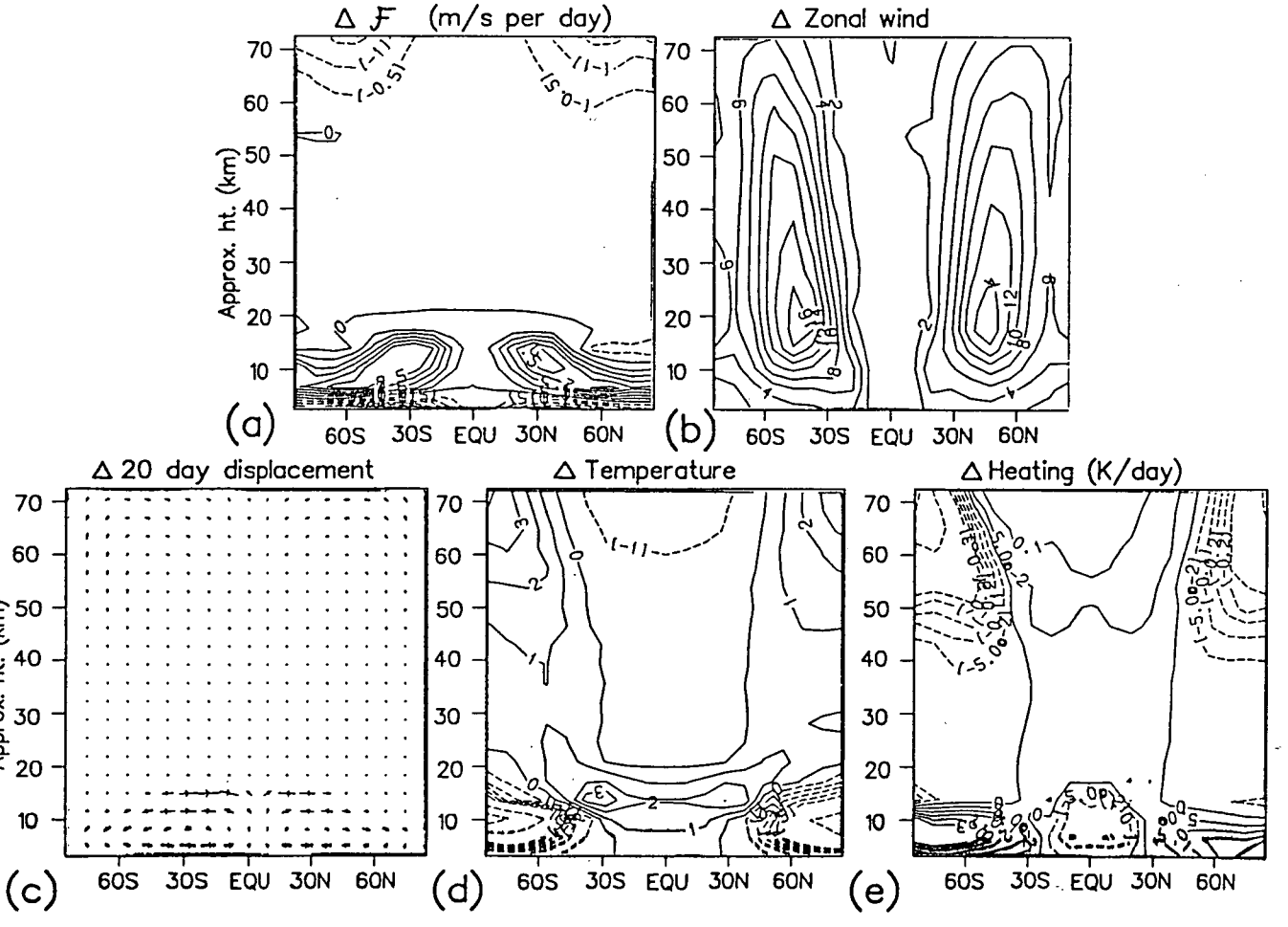


Figure 6.6: *Changes in (a) $\mathcal{F}$ ($ms^{-1}day^{-1}$), (b) zonal wind (m/s), (c) circulation (20 day displacement), (d) temperature and (e) heating (contours at 0.05,0.1,..,0.25,0.5,..°Cday^{-1}) due to a doubling in the tropospheric frictional time constants. Dashed contours indicate negative values.*

posphere (fig.6.6(a)) and an increase in $\overline{u}$, an increase which extends down to the ground (fig.6.6(b)). Therefore, since the surface frictional time constant was unchanged, $\mathcal{F}$ in the bottom layer becomes less positive (as it is a combination of the

specified positive $\mathcal{F}$ and the surface friction). Since $\mathcal{F}$ plays a large part in driving the meridional circulation (see section 6.1.1), a reduction in the size of both the positive and negative $\mathcal{F}$ leads to a reduction in the strength of the Hadley cells (see fig.6.6(c)). This change in circulation tends to increase the surface density near the poles and lower it near the equator, thus lowering the pressure (and hence temperature) of the polar isentropes and raising the equatorial temperature (fig.6.6(d)), until these temperature changes induce balancing convective changes (polar heating and equatorial cooling - see fig.6.6(e)). The temperature changes are in thermal wind balance with the changes in $\overline{u}$. Because the change in temperature is small above the troposphere, the change in $\overline{u}$ decreases only slowly with height (through thermal wind balance), so the large changes in $\overline{u}$ continue into the stratosphere. The increased frictional force on this increased $\overline{u}$ induces a poleward circulation (which is about 5% of the size of the total circulation, shown in fig.6.5(b)). This induced circulation increases with height (as the Rayleigh frictional time constants decrease) so the poleward sides of the stratospheric isentropes are pushed down (thus raising the temperature) while the equatorward sides are raised. Note the greater change in temperature at higher latitudes (see section 4.5.1). This change in temperature induces a balancing change in the heating rate. It is perhaps remarkable that there is such a strong correlation between temperature and heating rate in the stratosphere, and this fact is made use of in Chapter 7 in calculating Newtonian cooling time constants for the stratosphere.

In summary, reducing the tropospheric friction allows $\overline{u}$ to increase throughout the troposphere (by up to 16 m/s) and stratosphere, the increase being still large (about 10 m/s) at about 50km. There is a small increase in the strength of the upper stratospheric meridional circulation. There is a large decrease in the tropospheric $\mathcal{F}$ of about 30% and the Hadley cells decrease in strength by about 40%. The temperature at the polar tropopause falls by up to 15°C.

6.2.2 Increased Cooling

This subsection compares the control run with a run using a version of the model identical except that the specified radiative cooling in the troposphere (see section 5.3.7) was doubled.

Due to the increased cooling, the polar tropopause falls by about 20K, while the tropical tropopause falls by less than 10K, because of the stronger tropical con-

vection. (See Figure 6.7(d)). This change in temperature gradient gives rise to

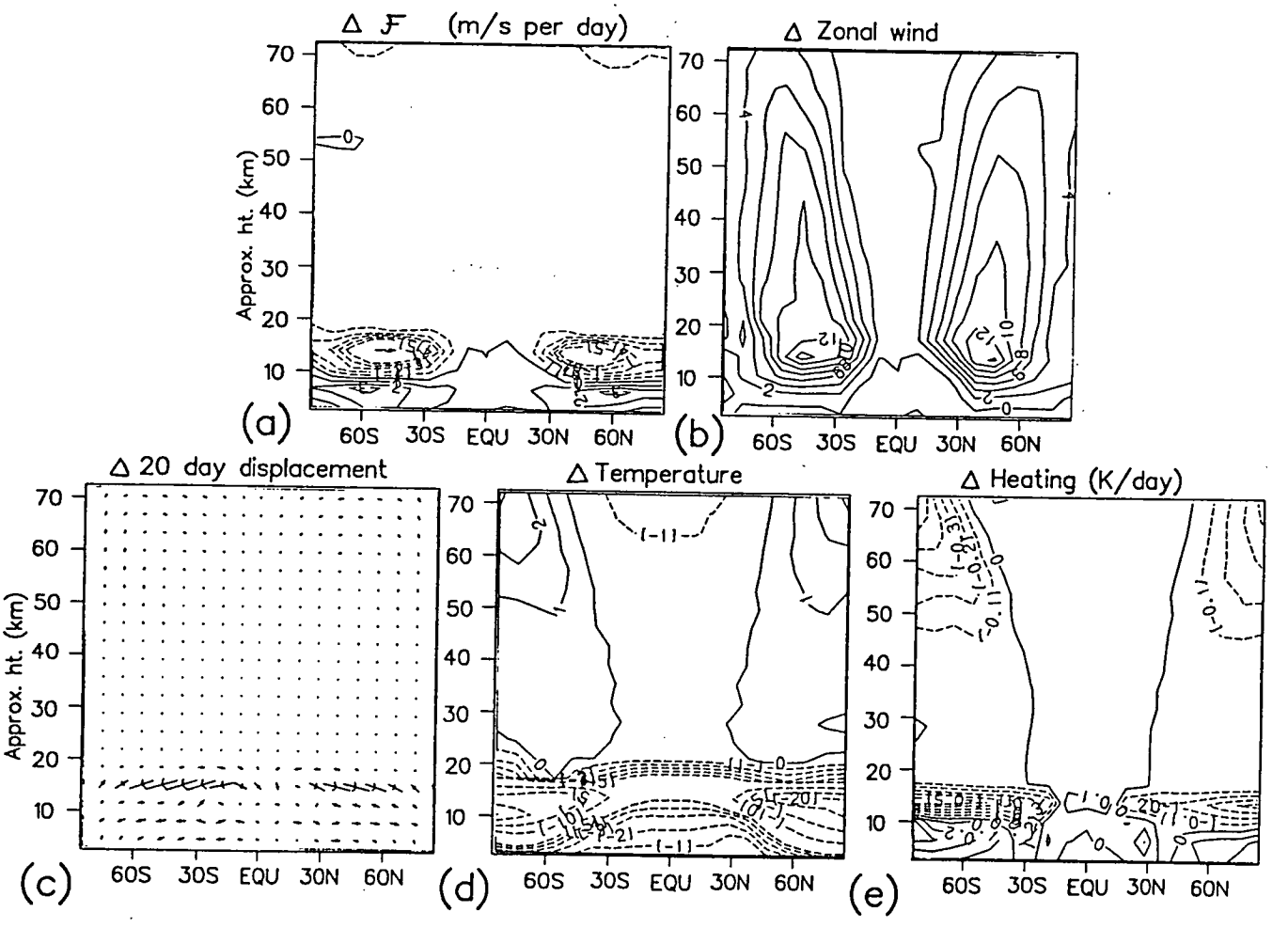


Figure 6.7: *Changes in (a) $\mathcal{F}$ ($\mathrm{ms^{-1}day^{-1}}$), (b) zonal wind (m/s), (c) circulation (20 day displacement), (d) temperature (contours at 1,2..5,10,15..K) and (e) heating ($^{\circ}\mathrm{Cday^{-1}}$) due to an increase in the tropospheric radiative cooling. Dashed contours indicate negative values.*

a change in $\overline{u}$ (fig.6.7(b)) which, as in the previous experiment, extends into the stratosphere, since the main temperature changes are in the troposphere. The increase in $\overline{u}$ in the upper troposphere, and the decrease near the surface, generally strengthen the Hadley cell (though at about 7km it appears to weaken it) because of the frictional parametrisation of $\mathcal{F}$. As before, the increased stratospheric and mesospheric $\overline{u}$ induces a change in $\mathcal{F}$, circulation, temperature and heating rates.

The changes in these fields above about 20km are very similar in size and shape to those in fig.6.6, which suggests that the effect on the stratospheric $\overline{u}$ and $\overline{T}$ of doubling the tropospheric cooling rates would be almost balanced by *halving* the tropospheric frictional time constants. The tropospheric Hadley cells would be

strengthened, by perhaps 50% or more. The equatorial tropopause would fall by about 10K, increasing the heating rate there by about 0.1 K/day. Temperatures in the polar upper troposphere would fall slightly.

6.2.3 Increased Critical Lapse Rates

This subsection compares the control run with a run using a version of the model identical except that the critical lapse rates (see section 5.3.8) within 10 degrees of the equator were increased from 5.5 to 6.5 K/km. Only the rates near the equator are increased since convection in the mid-troposphere is assumed to be mainly confined to equatorial regions.

Increasing the tropical critical lapse rate allows the temperature of the tropical tropopause to fall by about 5K. (See Figure 6.8(d)). The temperature falls not just at the equator, but throughout the troposphere and most strongly near the tropopause, falling by a maximum of over 6K at about 40N and S. Because the critical lapse rates polewards of 10N and S have not been changed, this temperature change induces an increase in convectional heating at about 20° N and S. Equatorial convection is reduced, leading to a decrease in the heating there. Thermal wind balance results in an increase in $\overline{u}$ with height equatorwards, and decrease polewards, of 40° N and S. Again, changes in $\overline{u}$, $\overline{T}$ and circulation occur in the stratosphere. Compared with the changes induced by the two previous experiments, however, the changes caused in both the troposphere and stratosphere by the increased critical lapse rates are small.

6.2.4 Decreased Vertical Diffusion

This subsection compares the control run with a run using a version of the model identical except that the vertical diffusion coefficient K_{zz} was reduced in the stratosphere and troposphere from $7x10^{-10}$ to $1x10^{-10}$ (see section 5.3.1).

Since, in the troposphere, or in the stratosphere, pressure decreases with height, density ($\equiv \frac{-1}{g}p_z$) will also decrease with height. Since diffusion tries to smooth out vertical variations in density, it will pump matter upwards across the isentropic surfaces (ie: it will heat the air). The biggest heating will be at the tropopause where the vertical gradient of density is the greatest. Therefore reducing the

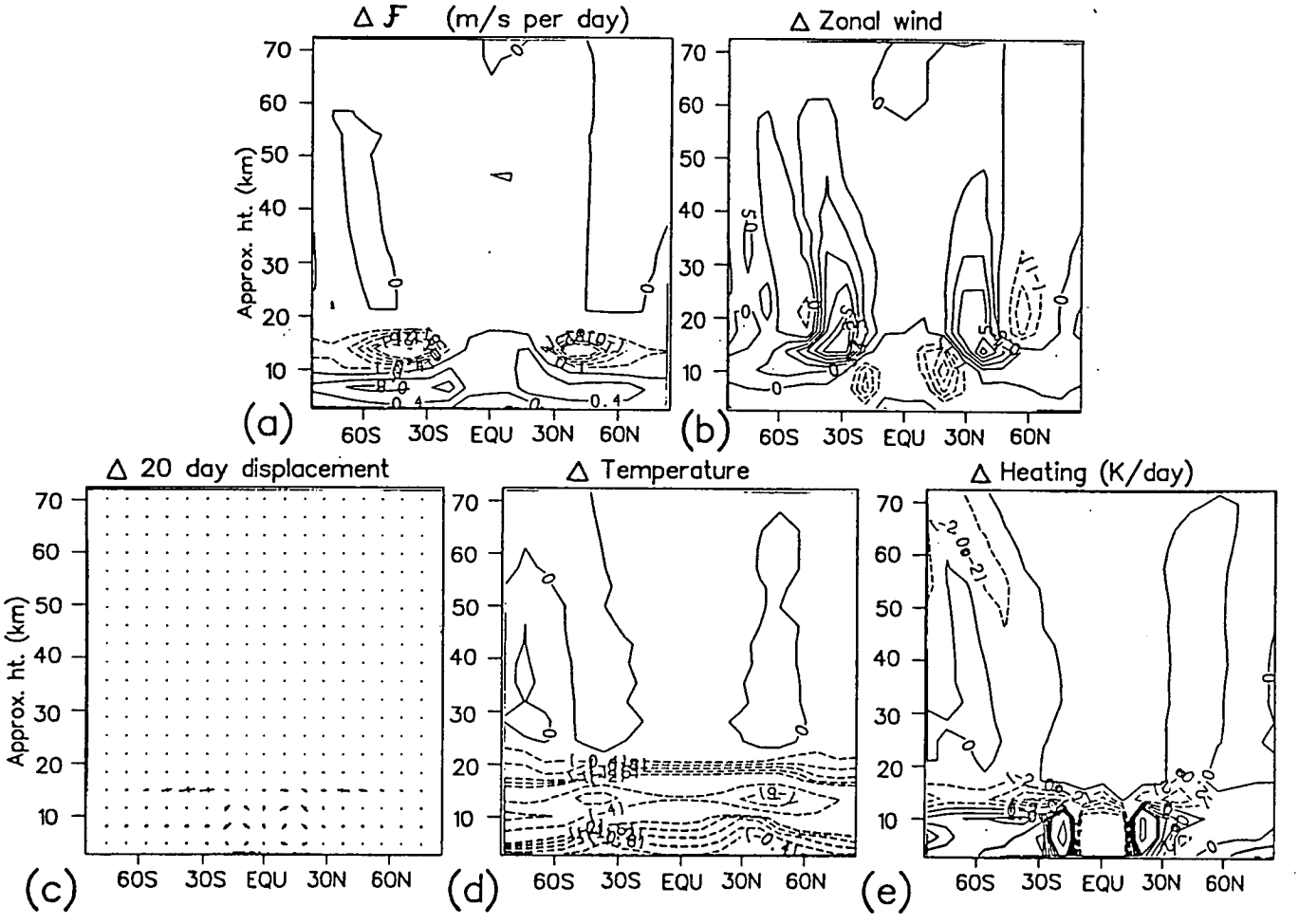


Figure 6.8: *Changes in (a) $\mathcal{F}$ ($\mathrm{ms^{-1}day^{-1}}$), (b) zonal wind (contours at 0.5 m/s intervals), (c) circulation (20 day displacement), (d) temperature (contours at 0.4,0.8,..2.0,4.0,.. K) and (e) heating (contours at 0.02,0.04,..,0.1,0.2,.. $^{\circ}\mathrm{Cday^{-1}}$) due to an increase in the tropical lapse rate. Dashed contours indicate negative values.*

diffusion will reduce the heating rate, with the greatest reduction at the tropopause (see Figure 6.9(e)). With the fall in the tropospheric temperature (fig.6.9(d)) associated with the reduced heating, convection increases and hence we see the increase in heating rate near the surface, at least polewards of 40° N and S. Nearer the equator, convection occurs right up to the tropopause, and almost cancels out the change in heating rates due to lack of diffusion. It also prevents the temperature from falling as far as it does at the poles, and the resulting change in horizontal temperature gradient produces the change in $\overline{u}$, which again extends into the stratosphere. The changes in all the dynamical fields are very similar to those produced by increasing the tropospheric cooling (comparing figs.6.7 and 6.9). In the model, both heat and momentum are diffused vertically, but it appears

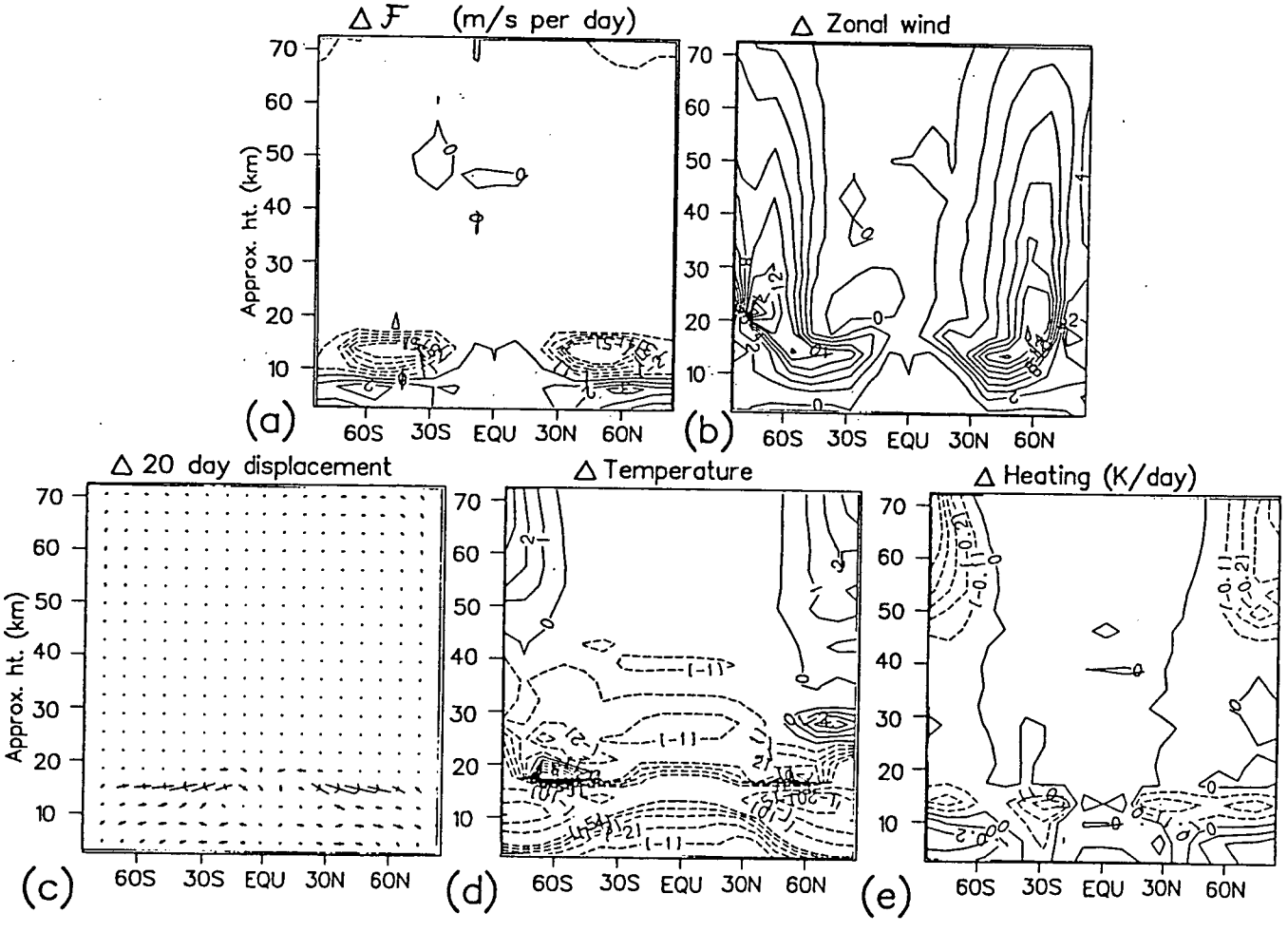


Figure 6.9: *Changes in (a) $\mathcal{F}$ (contours at 1ms^{-1}day^{-1} intervals), (b) zonal wind (contours at 2 m/s intervals), (c) circulation (20 day displacement), (d) temperature (contours at 1,2,..5,10,.. K) and (e) heating (contours at 0.1 °Cday^{-1} intervals) due to a decrease in vertical diffusion. Dashed contours indicate negative values.*

that diffusion of heat is the more dynamically important process, at least in the troposphere. This is deduced from the observation that the change in $\mathcal{F}$ appears to be mainly due to tropospheric friction acting on a change in $\overline{u}$, (and not due to vertical diffusion of momentum) since comparing figs.6.9 and 6.7 in which the increase in the tropospheric $\overline{u}$ is about the same, we see that the change in $\mathcal{F}$ is also very similar.

6.3 Summary and Conclusions

A version of the model in which the PVF due to stratospheric planetary waves is not included appears to simulate fairly well the real atmosphere in the autumn and in the SH winter, when stratospheric planetary waves are small. The observed deceleration of the SH $\overline{u}$ from June to July is not reproduced by the model, implying a negative PVF in the real SH stratosphere in June. In summer, although planetary waves are also expected to be small, at least above the lower stratosphere, the model easterlies do not extend down to the tropopause as in the real atmosphere. It is suggested therefore that the observed summer easterlies are due to the lingering effects of the winter and spring $\mathcal{F}$ combined with a negative $\mathcal{F}$ in the lower stratosphere over spring and summer, and perhaps also to a lack of heating in the model lower stratosphere over spring and summer. In spring and in the NH winter, when planetary waves are present in the real stratosphere, the model departs significantly from observations.

It was noted that the model troposphere is fairly realistic, with $\overline{u}$ generally lying within 5 m/s of the observed values. The polewards movement of the jet from winter to summer is also reproduced, and the equatorial tropopause minimum is only about 5K too warm. Either halving the tropospheric friction or increasing the cooling by 50% leads to unrealistically strong (by about 10 m/s) tropospheric westerly jets, although such changes do not push the values of $\mathcal{F}$ and heating in the troposphere outside the broad limits set by observation of the real atmosphere. Significant increases in $\overline{u}$ extend upwards to over 50km, with small changes in temperature and circulation occuring in the mesosphere and upper stratosphere.

However, a suitable combination of *increased* friction and increased cooling would have only a small effect on $\overline{u}$ (in both the stratosphere and troposphere), the main effect being to increase the strength of the tropospheric Hadley cells. In other woeds, for values of $\overline{u}$, $\overline{T}$, $\mathcal{F}$ and heating lying within observational limits, a wide range of Hadley cell strengths is possible, the strongest being associated with the largest values of tropospheric friction and cooling. This will be useful when chemistry is incorporated in the model (as an extension of the work of this thesis), since it will allow the strength of the Hadley cell, which will have a large effect on gases with tropospheric sources and sinks, to be altered without producing unrealistic tropospheric winds and temperatures.

The 80% reduction in diffusion of $\overline{\rho}$ has a very similar effect on the dynamics as a

doubling of the radiative cooling, while diffusion of $\overline{\tau}$ appears to give only a small contribution to $\mathcal{F}$. Therefore diffusion appears to have a significant effect on the dynamics, and should therefore be treated with more care in future versions of the model. The effect of this vertical diffusion of density could be mainly absorbed by a rescaling of the heating rate. Down-gradient diffusion of chemicals would, however, act in opposition to this additional mean circulation due to diffusion of $\overline{\rho}$.

A change in equatorial critical lapse rates has only a small effect (compared with the previous changes) and its effect on temperature is greatest at the *mid-latitude* tropopause.

All of the above changes made to the tropospheric parameters have an effect on the model fields, throughout the troposphere and stratosphere, which is very nearly independent of the time of year.

This completes the description of a run of the model without a stratospheric $\mathcal{F}$ due to planetary waves. In the following chapter an estimate of the PV flux due to stratospheric planetary waves will be applied to the model, and the subsequent differences from the model of this chapter will be discussed.

Chapter 7

PV flux estimates and their effect on the model

Do the waves slow the wind, or it accelerate?
From satellite data it's tricky to state.
To this question an answer I'd like to be knowing
so tell me, please, which way is PV a-flowing?

A number of studies (eg. Geller et al.(1984), Mechoso et al.(1985)) have calculated climatologies of $\nabla.F$ using geostrophic winds derived from satellite data. In these calculations of $\nabla.F$, positive regions appeared. It was shown by Robinson (1986), using a 3D model, that the use of geostrophic winds could lead to spurious positive values, which therefore cast doubt on the reality of a positive $\nabla.F$ occuring in the atmosphere. A possible test of the climatological $\nabla.F$s is to feed them into a Transformed Eulerian Mean (TEM) model, comparing its response to that of the real atmosphere. Although such a test has not been carried out, with a similar purpose Crane et al.(1980) drove a 2D isobaric model with observed momentum fluxes while parametrising the heat fluxes. They thus specified only part of the wave driving. Driving an isentropic model with PV fluxes calculated from winds derived from satellite data is similar to driving a TEM model with $\nabla.F$, and will provide a test of our ability to estimate in this way the total dynamical forcing in the stratosphere due to planetary waves.

In this chapter we will present our estimate of the PVF for the year from July 1980 to June 1981, calculated from the approximate winds described in chapter 2. This PVF will be applied to the 2D model of Chapter 5, assuming that PVF is a good approximation to $\mathcal{F}$. The effect on the model fields and meridional circulation will be described and discussed. We will then compare the resulting $\overline{u}$ and $\overline{T}$ fields with observations, highlighting points which indicate deficiencies in

the calculated PVF, or the model. Also, features will be identified which occur in the real atmosphere, and in the model when it is driven by this PVF - in particular the effect of a strong PVF in one hemisphere on the $\bar{u}$ of the other, and the cooling of the tropical tropopause by a stratospheric PVF.

7.1 Sources of uncertainty in the calculated PV flux

There are a number of factors contributing large uncertainties to the PVF calculated from estimated winds. Because PVF and $\nabla.F$ are similar quantities, we expect that many of the conclusions reached about $\nabla.F$ in other studies should also apply to PVF, in particular those relating to the sources of error in $\nabla.F$.

A main feature of $\nabla.F$ calculated from satellite data, which has been noted by several authors (eg. Geller et al.(1984), Randel(1987) and Robinson(1986)), is the appearance of regions where it is positive. Robinson suggests that they are due to the over-estimation of momentum fluxes through using geostrophic winds, and both Randel and Robinson have noted that using balance winds reduces the magnitude and extent of these regions. However, it is still very difficult to achieve accurate estimates of winds from the temperature data supplied by satellites (see Marks(1989) and Li (1991) for recent attempts) so this remains a primary source of uncertainty in $\nabla.F$.

Other sources of uncertainty lie in the retrieval of geopotential heights from the satellite data. Satellite instruments measure the radiance of certain slabs of the atmosphere. The spatial resolution of the data depends on the instrument used to obtain the radiances. The SSU nadir-viewing device, from which the data used in this thesis came, looks almost vertically and so has good horizontal resolution (wavenumber 12) but poorer vertical resolution (10-12km). From the radiances the temperatures of these slabs can be estimated. To obtain the temperature for a larger number of slabs, statistical regression can be used to provide an estimate based on climatology, introducing a certain amount of uncertainty (see eg. Jackson et al. (1990)). From the temperature, geopotential heights of pressure levels can be built up, using the hydrostatic approximation, providing the height of the bottom pressure level is known. Because the calculation of PV requires the differentiation of pressure twice with respect to potential temperature, poor vertical resolution can result in the loss of features of PV plotted on isentropic (horizontal) surfaces.

Clough et al(1984) and Grose and O'Neill(1988) have calculated $\nabla.F$ and PV for the same day using data from different instruments and different geopotential base-height fields and find substantial discrepancies. Although it is difficult to estimate the uncertainty in calculations of PV flux and $\nabla.F$, consideration of the above points should certainly make us more cautious about the reliability of these calculations. Nevertheless, McIntyre (1987) commented that despite all the reasons for doubting them, PV maps derived from satellite data seem to be remarkably accurate.

Further uncertainties arise when we take the monthly-mean of daily values of PV flux because daily values can exceed the monthly average, sometimes by a factor of about ten, due to vacillations of the vortex about the pole which are accompanied by large PV fluxes of alternating sign. These vacillating fluxes should probably cancel to a large extent in the monthly-mean, but due to missing days of data and the errors in daily PV flux calculations already mentioned, the uncertainty in monthly-mean fluxes could be large.

7.2 Calculation of PV flux from estimated winds

7.2.1 Features of PV flux

In this section we will describe the main features of the PVF calculated directly from the winds derived from satellite data (described in chapter 2). The fluxes were only calculated up to about 45km, and above that they were linearly reduced to zero at about 70km before applying them to the model. This applies to the PVFs shown in this chapter and in chapter 8. This treatment of the fluxes was roughly guided by the NH wavenumber-one geopotential height amplitudes of Barnett and Corney (1985b) which, when they were large, reached a peak at about 55km. However, they were small but still not negligible at 85km, the upper limit of the data. In the model mesosphere, though, $\mathcal{F}$ is dominated by the frictional representation of gravity wave breaking, so the treatment of PV fluxes there will not be very important. The PV fluxes are shown in Fig(7.1) for the months from July 1980 to June 1981, and their features are given in note form.

In the NH :-

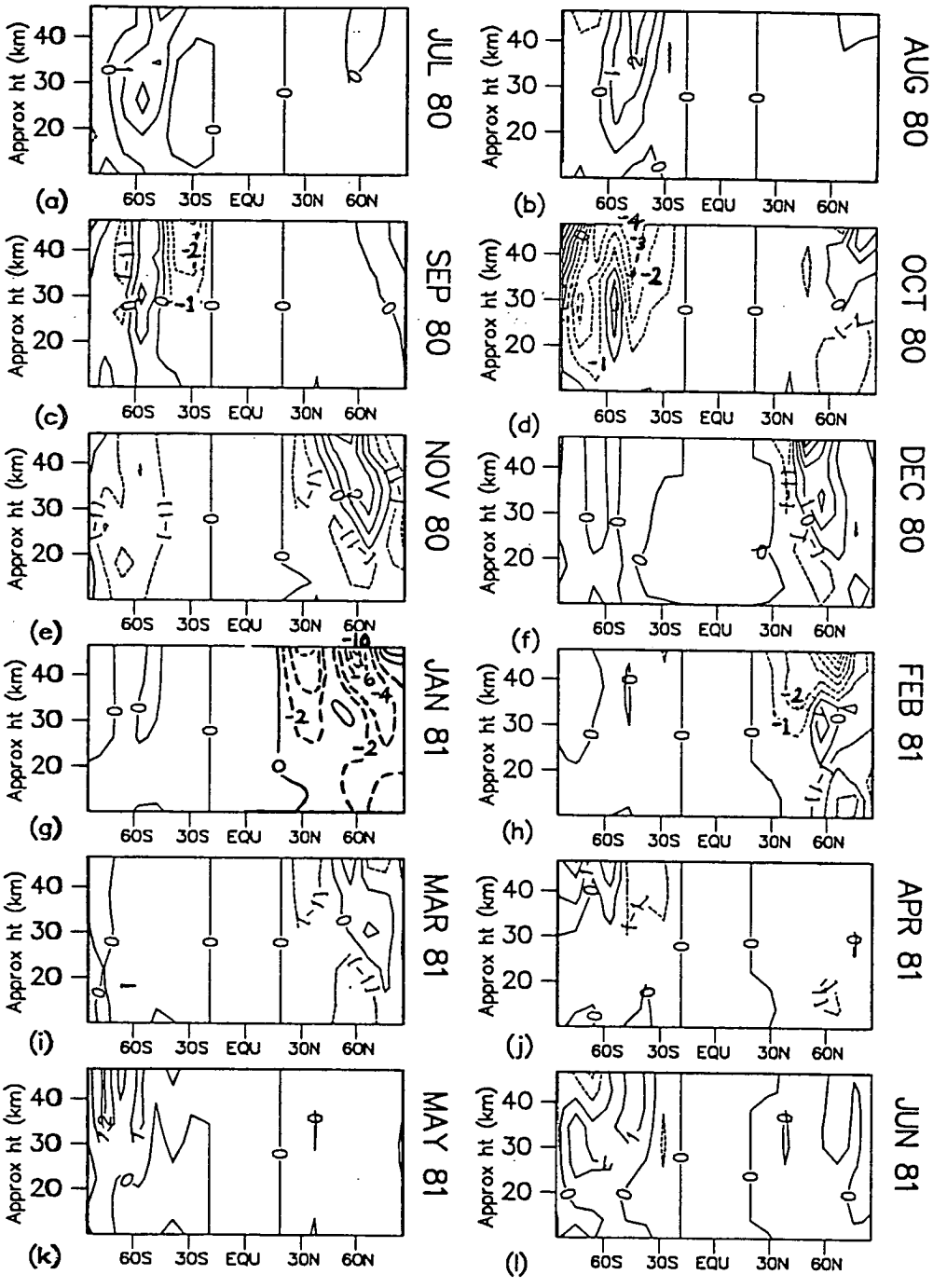


Figure 7.1: *Direct estimate of PV fluxes, in* $\mathrm{ms^{-1}day^{-1}}$. *(a)July 80 ... (l)June 81. Dashed contours indicate negative values.*

(1) The PVF is small (less than 1 $\mathrm{ms^{-1}day^{-1}}$ in magnitude) for the six months from April to September. From the calculations of $\bar{u}$ made in Chapter 2 (but not shown in this thesis), and from Barnett and Corney's (1985a) data, these

are the only months, apart from February, when the monthly-mean $\bar{u}$ is small (less than 20 m/s everywhere north of 20N) or easterly, in the NH. In February, as noted in section 3.5.1, $\tilde{u}$ is not small although $\bar{u}$ is small, due to the distortion of the vortex. We can therefore say that the PVF is small when $\tilde{u}$ is small or easterly.

(2) The main negative PVF occurs during January and February, reaching -10 $ms^{-1}day^{-1}$in January, but negative fluxes exceeding -1$ms^{-1}day^{-1}$ are present from October till April.

(3) Significant negative PVFs occur in the lower stratosphere (below 25km) from October till April.

(4) Positive PVFs appear from October till March, being largest in December (about 4 $ms^{-1}day^{-1}$) and lying almost exactly over the westerly jet maximum in $\bar{u}$. In January this positive region is small but there is an obvious double peak in the negative flux, with the dip between the peaks occurring at the position of the westerly jet maximum, at about 50N. In February, when $\bar{u}$ is small, the double negative peak does not occur as it does in the other months, although there is a positive region below about 35km.

(5) Other positive regions occur near the pole in October and January.

In the SH :-

(6) The PVF is negligible for the four months from December to March. Again, these are months when $\bar{u}$ is small or easterly. In November $\bar{u}$ is also small, while the PVF is not small. However, $\tilde{u}$ is about 10 m/s stronger than $\bar{u}$. As in note (1) above, the PVF is small when $\tilde{u}$ (not $\bar{u}$) is small or easterly.

(7) The main negative PVF occurs from September till November, reaching -4 $ms^{-1}day^{-1}$in October.

(8) Significant negative PVFs occur in the SH in October and November.

(9) Positive fluxes occur from April till October, their positions again correlated well with the westerly jet maximum. However, in November, when $\bar{u}$ is small, there is neither a positive region nor a dipole structure.

(10) Other positive regions occur near the pole in April, May and October.

In summary, the PVF is small when $\tilde{u}$ is small (less than 20 m/s) or easterly. The same cannot be said of $\bar{u}$ since there are times (Feb in the NH, Nov in the SH) when

$\bar{u}$ is small but the PVF is large. From November till March in the NH, and from April till November in the SH, the PVF is not small, but can be either positive or negative. There appears to be a systematic positive error in PVF (a claim which will be further supported later in this chapter) in the regions where $\bar{u}$ is large and westerly. Sometimes this dominates any negative PVF and at other times just results in a double peak in the negative PVF. The correlation between the positive error and $\bar{u}$ is observed at times when both the error and $\bar{u}$ are large (Dec in the NH and Oct in the SH) and when both are small and yet there is a significant negative PVF (Feb in the NH and Nov in the SH). When this postulated systematic error is taken into account it is possible that the actual PVF is significantly negative in the NH from November till March, and in the SH from April till November.

7.2.2 Comparison with Chapter 4

Comparing the $\mathcal{F}$ deduced from $\mathcal{F}^D$ in Chapter 4 with the PVF calculated in this chapter, assuming that the real PVF is a good approximation to the real $\mathcal{F}$ in the stratosphere, we note the following :-

(1) There is no evidence in chapter 4 either for the double peak in the negative PVF, or for the regions of positive PVF, obtained in this chapter.

(2) The peak PVFs of -10ms^{-1}day^{-1} in January and -4ms^{-1}day^{-1} in October agree well with the values obtained in chapter 4, of -9ms^{-1}day^{-1} for Jan/Feb and -5ms^{-1}day^{-1} for Sep/Oct, when the faster diabatic time constants are used.

(3) The large negative $\mathcal{F}$ in the NH in November and December, found in chapter 4, are not found here. Instead, positive PVFs of about 3ms^{-1}day^{-1} are calculated. Similarly, in the SH in August and September positive PVFs occur at the latitudes where, in chapter 4, a large negative $\mathcal{F}$ was deduced.

(4) The regions of positive PVF obscure any polewards movement of PVF which might, as inferred in chapter 4, occur in the real atmosphere.

7.2.3 Effect on model

When the fluxes of Fig(7.1) were added to the stratospheric $\mathcal{F}$ of the 'control' model of chapter 6, the resulting $\bar{u}$ field had an unrealistically large latitudinal

gradient (for example fig(7.2) shows $\overline{u}$ for January). In chapter 4 it was explained

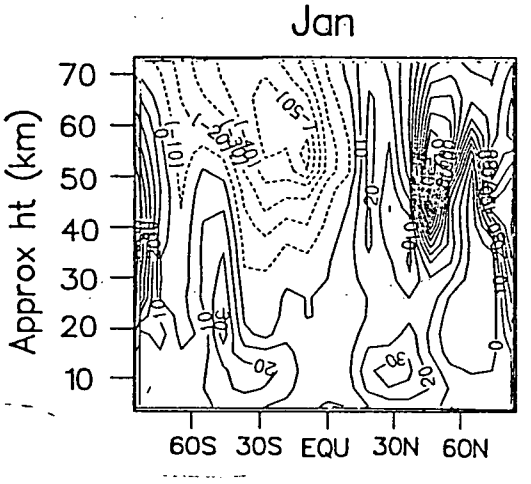


Figure 7.2: *The $\overline{u}$ field for January for the model driven by the unsmoothed PVF*

how a region of negative $\mathcal{F}$ induces a poleward $\overline{V}$ which opposes the deceleration of $\overline{u}$ inside the region, but continues outside the region, where it accelerates $\overline{u}$. Thus in January, in between the two negative peaks in PVF, the acceleration due to the poleward $\overline{V}$ is likely to dominate the small negative PVF, resulting in an acceleration of $\overline{u}$ where $\overline{u}$ is strong already (it was noted that the positive PVF, or the dip between the two negative peaks, occured at the jet maximum). As already noted, the use of geostrophic winds when calculating ∇.F can result in regions where ∇.F is positive. Moreover, any evidence for regions of positive PVFs, or double peaks, was not observed in the $\mathcal{F}^D$ of chapter 4.

Thus, it seems reasonable to conclude that the regions of positive PVF and the dipole structure are not real and are a result of a systematic error correlated with a strong westerly $\overline{u}$ which arises from the use of approximate winds.

In an attempt to make use of our estimates of PVF in the model we removed the regions of positive PVF, by setting them to zero, and smoothed the resulting field. As can be seen (fig.7.3), the shape of the fields has changed considerably, the double peaks having merged into one at about 60 degrees north and south. The peak values of the negative fluxes, however, have not changed by much, the biggest change occuring in January, the peak falling from -10 to -8ms^{-1}day^{-1} . They are similar in size and shape to the $\mathcal{F}$ deduced in chapter 4, apart from the lack of a significant negative PVF in the NH in November and December, and in the SH in August.

The value of using such highly processed fluxes in the model might appear doubtful.

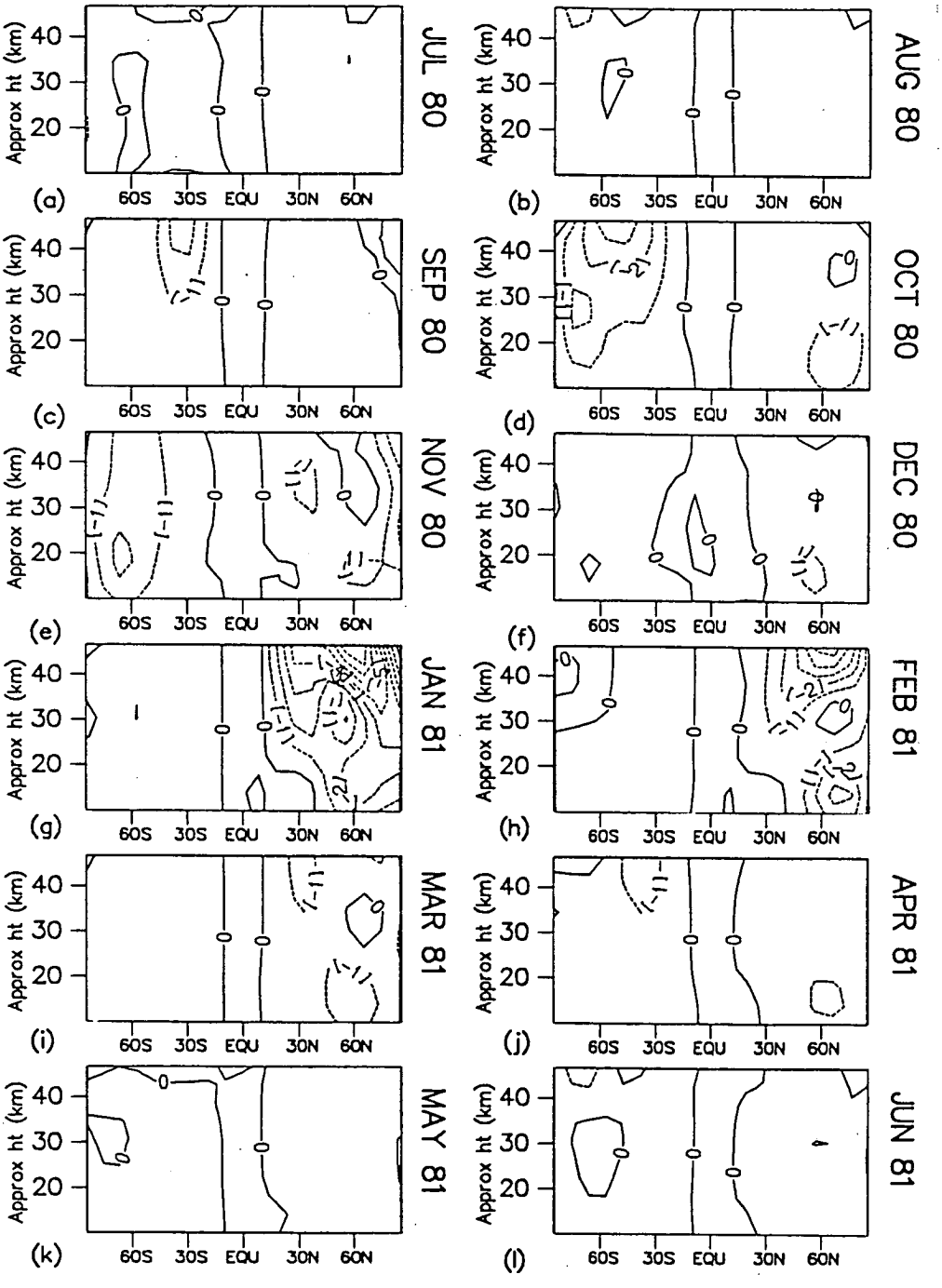


Figure 7.3: *Smoothed PVFs for (a) July 1980 ... (l) June 1981 in* $\mathrm{ms}^{-1}\mathrm{day}^{-1}$.

However, because of their similarity to the $\mathcal{F}$ of chapter 4, comparison of the model output with observations will provide some idea of the accuracy of the estimates of chapter 4. It will also reveal how the model responds to such a stratospheric $\mathcal{F}$.

The smoothed fluxes were therefore added to the small frictional $\mathcal{F}$ in the strato-

sphere of the 'control' model used in chapter 6. The model was run for three years until an annual periodicity set in and then compared with with the third year of a run of the control model. Some of the differences from the control run, on the 15th of January, induced by the addition of the PV fluxes are shown in fig.(7.4). The meridional circulation induced by the PVF (fig.(7.4)(a)) shows the strong poleward flow where $\mathcal{F}$ is large and the two-cell structure described in chapter 4 (see figures 4.6(a) and 4.8). Note also that the northward flow crosses the equator from the SH, a point which will be taken up in section 7.4.

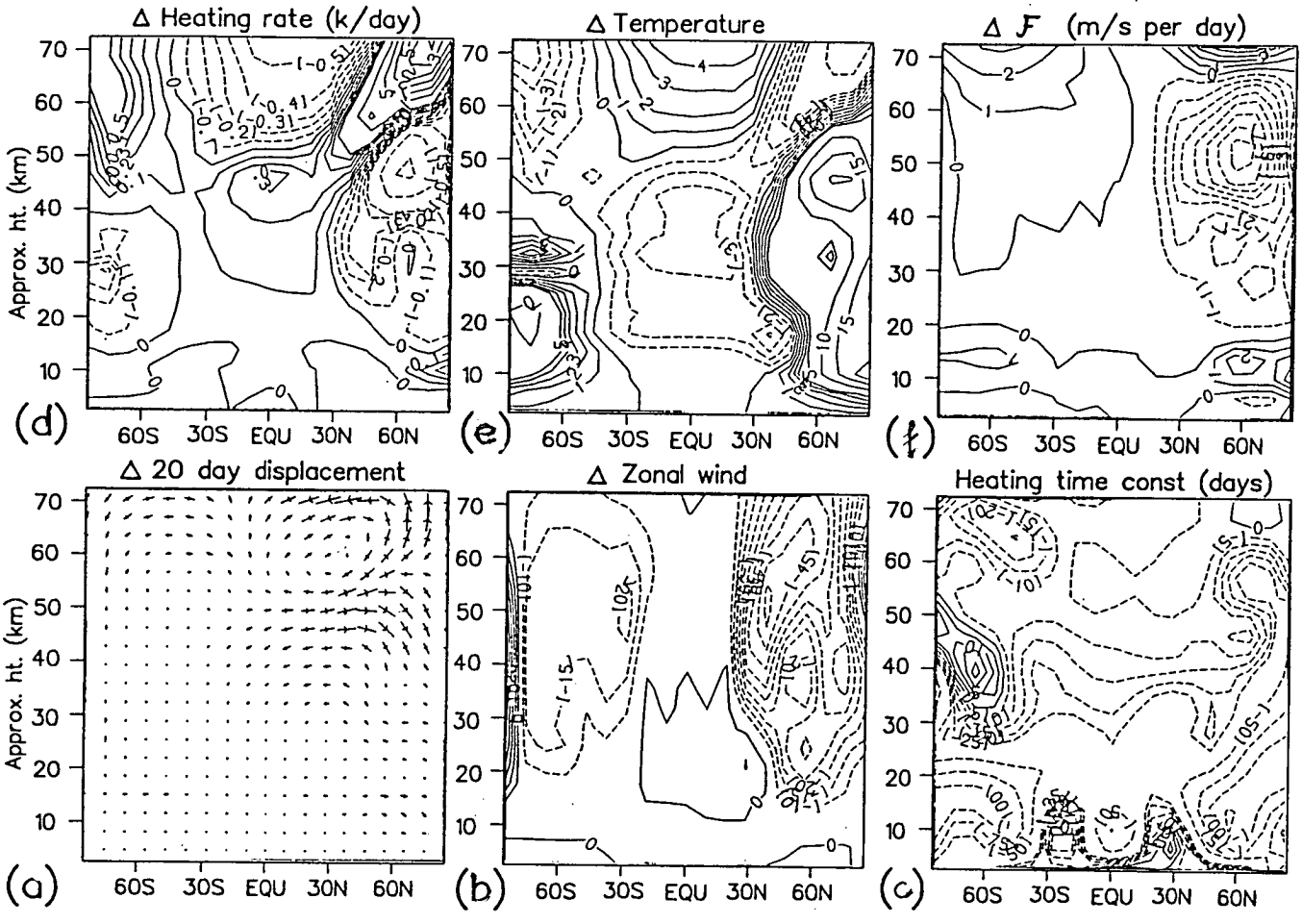


Figure 7.4: *Changes from the control run induced by the smoothed PVF on 15 Jan (a) meridional circulation (20 day displacement) (b) zonal wind (c) temperature (d) heating rate (degrees per day) (e) $\mathcal{F}$ ($\mathrm{ms^{-1}day^{-1}}$). (f) is the negative of the smoothed radiative time constant in days. Dashed contours indicate negative values.*

The zonal wind (Fig 7.4(b)) has been reduced by up to 50 m/s in the NH. The shape of the reduction resembles the shape of $\mathcal{F}$ for January fairly closely because

$\mathcal{F}$ for January is much larger than the $\mathcal{F}$ over the previous months, the difference being due to the frictional part of $\mathcal{F}$. The reduction in $\overline{u}$ (fig 7.4(b)) extends down below about 14km (which is the lower limit of the PV flux applied to the model) due to the lower, equatorwards, return flow. This polewards flow is commensurate with the positive change in the frictional $\mathcal{F}$ below 14km due to the reduction in $\overline{u}$. In the SH, $\overline{u}$ has also been significantly reduced, presumably because of the negative PVF in the SH during October and November.

The change in heating (fig 7.4(d)) which accompanies the change in temperature (fig.(c)) enables us to estimate the radiative time constant ($1/\alpha \equiv -\Delta T/\Delta Q$ of Chapter 4) (see fig 7.4(f)) as ranging from about 100 days in the lower stratosphere to about 5 days in the mesosphere. Notice that the good spatial correlation between the changes in temperature and heating seems to justify our use of the Newtonian cooling approximation in Chapter 4. Fig 7.4(f) has been smoothed, but even so, because of the erratic spatial variations we approximate it at all latitudes by the values in table 7.1, which is a rough estimate of $1/\alpha$ at mid-latitudes.

Approx ht. (km)	56	49	42	35	28	21
$1/\alpha$ (days)	10	15	15	25	50	100

Table 7.1: *Values of* $1/\alpha$ *used in Chapter 4.*

Looking at the change in temperature in the NH (Fig 7.4(c)) we see also the quadrupole structure of figure 4.6(c) with its centre at about 45N, 55km. In the mesosphere, the parametrised friction has produced a positive change in $\mathcal{F}$ (fig 7.4(e)) as a result of the reduction in $\overline{u}$ caused by the equatorwards branch of the upper cell of the induced meridional circulation. Thus there is an interaction between the mesospheric $\mathcal{F}$ and the stratospheric $\mathcal{F}$ through the meridional circulation. The change in $\overline{u}$ is negative almost everywhere, (and this applies also to the other months), and in the NH it exceeds -10 m/s throughout the year, while it is smaller in the SH.

7.3 Comparison of model $\overline{u}$ and $\overline{T}$ with observations

Now the modelled monthly-mean $\overline{u}$ and $\overline{T}$ for the four mid-seasonal months will be compared with those of the model of Chapter 6, which did not have a specified

stratospheric PVF, and with those of Barnett and Corney's (1985a) climatology. The conclusions we reach in this section would be unchanged if we had considered the very similar observed $\overline{u}$ and $\overline{T}$ fields of Marks (1989), which are derived from 2 to 3 years of data. The model $\overline{u}$ fields are shown in fig 7.5, while those of Barnett and Corney's are shown in Chapter 6 in figure 6.1.

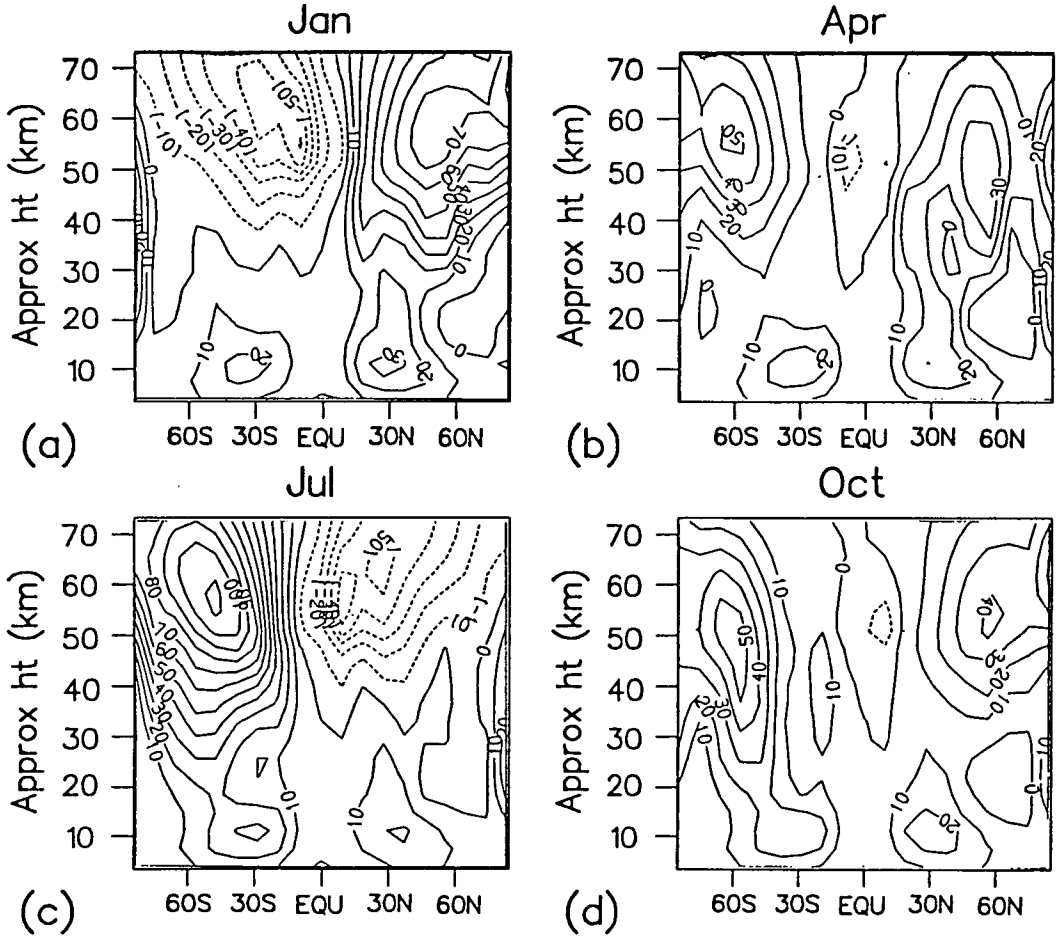


Figure 7.5: *Model $\overline{u}$ fields for (a) Jan (b) Apr (c) Jul (d) Oct. Dashed contours indicate easterly winds.*

Compared with the climatology, we notice the following features of the model $\overline{u}$.

(a) The SH westerly jet in July has about the right strength. However, the modelled jet in July is about 10 m/s stronger than in June (which is the same as the model of Chapter 6), while the observed July jet is about 20 m/s weaker than June's. (June values not shown).

(b) The NH jet is too strong (by about 10 m/s) in Jan, (and also in Nov and Dec

although not shown) and still exists in spring while the observed jet has almost disappeared. Compared with the model of Chapter 6, though, the spring jet has at least been reduced by about 30 m/s.

(c) In October, the SH jet is higher than the observed jet by about 20km, although it has about the right strength.

(d) The easterlies in both hemispheres do not extend as far downwards as observed in summer, although they are more extensive than those of the model of Chapter 6.

(e) Both the modelled and observed autumn (Apr SH/Oct NH) westerly jets are stronger in the SH than in the NH (by about 10 m/s in the model and about 15 m/s in observations).

(f) Both the modelled and observed summer (Jan SH/Jul NH) easterly winds near the equator are stronger in the SH than in the NH (by about -5 m/s in the model and about -20 m/s in observations).

Discussion of these points will be delayed till we have compared the $\overline{T}$s. The model $\overline{T}$ fields are shown in fig 7.6, while those of Barnett and Corney's are shown in Chapter 6 in figure 6.2. Note that 200K has been subtracted from the model temperatures.

We note the following;

(g) The tropical tropopause is colder in Jan and April than it is in July and October, in both the model (by about 3K) and the observations (by about 4K). The model tropopause is also a few degrees colder than that of the model of Chapter 6 (see also fig 7.4(c)).

(h) Near the modelled north pole in January, and the south pole in July, $\overline{T}$ is too cold (by up to about 50K) from about 30 to 60km, and too hot (by up to about 30K) from about 60km to 70km. This feature was also noted in the model of Chapter 6.

(i) At about 30km in winter, the model north pole is about 10 degrees warmer than the south pole, as compared with an observed difference of over 20 degrees.

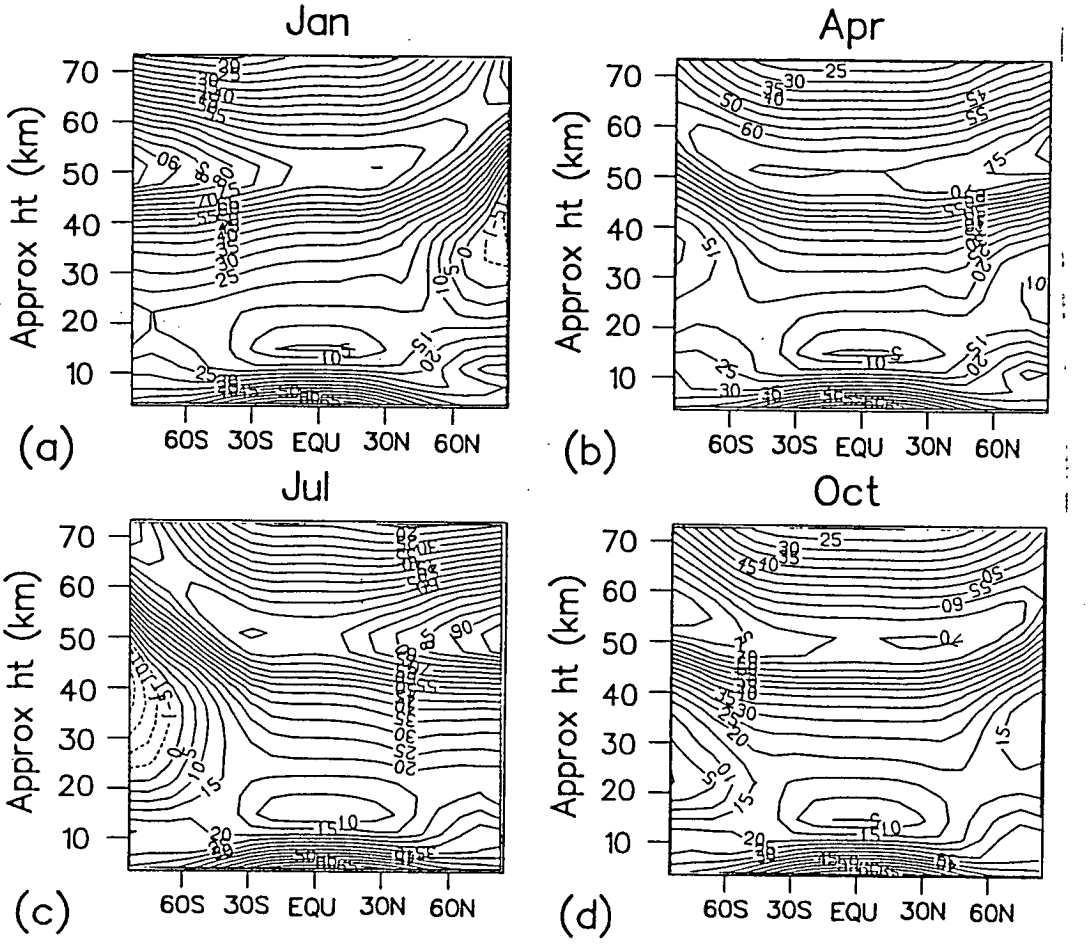


Figure 7.6: *Model* $\overline{\mathrm{T}}$*-200K for (a) Jan (b) Apr (c) Jul (d) Oct. Dashed contours indicate temperatures below 200K.*

(j) There is good agreement between the model and observed temperatures at all heights in April and October. In particular, the spring and autumn polar temperatures at about 25km and 30km respectively are less by about 5K in the SH than in the NH in both the model and observed fields.

We will now try to deduce, from the preceding points, similarities and differences between the real $\mathcal{F}$ and the smoothed PV flux of this chapter. From (a) we infer that there is a negative $\mathcal{F}$ in the real atmosphere in July at about 50km (the height of the westerly jet) which does not appear in the smoothed PVF. This would be in accord with the results of Chapter 4, which indicated a negative $\mathcal{F}$, stronger in the SH than in the NH, above 45km, in July/August. This is above the upper limit of our calculations of PVF. Points (b) and (i) imply a negative $\mathcal{F}$ missing from our calculations in the NH before January, and likewise point (c) implies a lack of $\mathcal{F}$ in the SH upper stratosphere before October. However, point (j) indicates that the

cumulative effect of the calculated $\mathcal{F}$ on $\overline{T}$ by spring is about right, implying a good estimation of $\mathcal{F}$ for the late winter/early spring months, with any mis-calculation of $\mathcal{F}$ in early winter having a small effect on the spring fields.

The addition of planetary wave PV fluxes has not cured the defects noted in points (d) and (h). These defects were also found in the model of Chapter 6. It was noted there that the lack of summer easterlies in the lower stratosphere (point (d) above) is probably due in part to a lack of heating in the lower stratosphere (where the heating is only crudely parametrised). Also, the presence of a negative $\mathcal{F}$ in the lower stratosphere during summer would help to create the easterlies. That such a $\mathcal{F}$ is plausible is suggested by the 3D 'SKIHI' model of Mahlman and Umscheid (1987), where ∇.F exceeds -1 $ms^{-1}day^{-1}$ for much of the NH lower stratosphere (from 5 to 20km) in mid-January. Point (h) is probably due to the over-simple representation of gravity wave drag.

The inferences made above all depend on the assumption that the model represents the physical processes in the atmosphere well enough so that departures of the modelled $\overline{u}$ and $\overline{T}$ are in large part due to misrepresentation of the $\mathcal{F}$ applied to the model. Bearing this in mind, then, it seems that the smoothed $\mathcal{F}$ applied to the model is fairly close to the actual $\mathcal{F}$ in the real atmosphere except in November and December in the NH and July and August in the SH, where regions of significant negative $\mathcal{F}$ are missing. This could be because they occur above the limit of the satellite observations or because they are masked by the systematic error connected with strong $\overline{u}$ discussed above, since in these months the westerly jet maximum is greater than 60 m/s. The negative $\mathcal{F}$ in the summer lower stratosphere might also be under-estimated or obscured.

7.4 Effect of $\mathcal{F}$ on the tropopause, and the cross-equator effect

The simulation of the observed annual cycle of the tropical tropopause temperature (see section 4.5.2) by the model (point (g) of the previous section) with a similar amplitude (about 3K) and phase (minimum in January) indicates that it is due in large part to the effects of stratospheric planetary waves in the real atmosphere. From the discussion of section 4.5.1 we see that a negative $\mathcal{F}$ in the stratosphere of either hemisphere will lower the temperature of the tropical tropopause. Therefore, from September till February, when $\mathcal{F}$ is large and negative in at least one of the

hemispheres, the temperature of the tropical tropopause will be lower than it is during the rest of the year. The size of this effect in our model depends on our parametrisation of the troposphere, but it seems reasonable to infer that planetary wave activity in the stratosphere can have a noticeable effect on the troposphere.

Point (e) might, at first glance, appear to be due to a stronger $\mathcal{F}$ in the NH in Sept/Oct than in the SH during Mar/Apr, and indeed, this may be a factor in the real atmosphere. However, the smoothed PVF applied to the model was small in the autumn (less than -1ms^{-1}day^{-1}) and yet the same behaviour was seen in the model, suggesting that perhaps another process is at work here. To eliminate the effect of any IHD in $\mathcal{F}$ on the modelled autumn fields, the model was run again with $\mathcal{F}$ as before, except that it was set to zero in the NH from June till October, and in the SH from December till April ie: during summer and early autumn. Therefore, unless there is another process at work, we would expect the westerly jets in mid-autumn to be similar in both hemispheres, since the summer hemispheres are very similar. The resulting $\overline{u}$ fields in April and October are shown in fig.7.7(a) and (b), which show that the SH westerly jet in April is still stronger than the NH October jet, although now only by about 5 m/s compared with the 10 m/s noted in point (e).

The extension of $\overline{V}$ across the equator, shown in figure 7.4(a), advects air of lower angular momentum equatorwards and hence decelerates $\overline{u}$ in the hemisphere opposite to that in which a strong negative $\mathcal{F}$ is present. This is further supported by fig.7.7(c), which is the IHD in $\overline{u}_t$ for Aug/Feb, and fig 7.7(d), which is the same for Oct/Apr. Figure 7.7(c) shows that the SH $\overline{u}$ in late summer (Feb,Mar), from the equator to about 40S and between 35km and 63km, is being decelerated with respect to the NH $\overline{u}$. This is the time when $\mathcal{F}$ is large and negative in the NH. Similarly, fig 7.7(d) shows that the NH $\overline{u}$ in late autumn (Oct,Nov) is being decelerated with respect to the SH $\overline{u}$, when $\mathcal{F}$ is large and negative in the SH. These figures may be compared with similar ones for the observed atmosphere in Chapter 4, though a detailed comparison must not be made since neither the PV fluxes nor the observed $\overline{u}$ are calculated within 20 degrees of the equator (see Chapter 2). Comparing fig 7.7(d) with fig 4.9(a), which is the observed $\overline{u}_t$ for Sep/Mar, a similar negative region is seen in the autumn hemisphere, though the observed value exceeds -0.4 ms^{-1}day^{-1} while the modelled one is about -0.2 ms^{-1}day^{-1}. Also, comparison of fig 7.7(c) with fig 4.9(b) (after reflection about the equator and reversal of the signs, so that summer appears on the right side of the figure) reveals the deceleration of the SH summer $\overline{u}$ with respect to the NH summer $\overline{u}$, al-

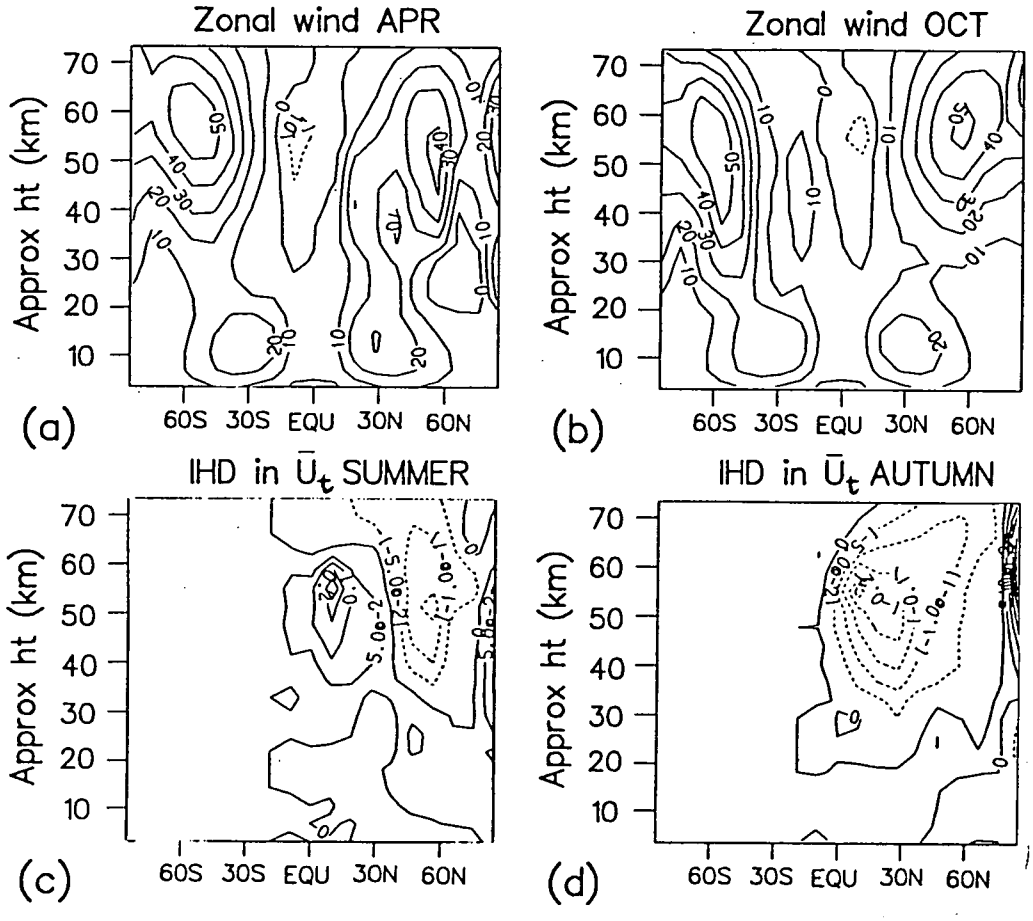


Figure 7.7: *Model* $\bar{u}$ *with no PVF in summer and autumn for (a) April (b) October, and IHDs in* $\bar{u}_t$ *fields (with the winter/spring hemispheres polewards of 20 degrees blanked out) for (c) late summer/early autumn (Aug/Feb) (d) late autumn (Oct/Apr). (Contour interval is 0.05* $\mathrm{ms^{-1}day^{-1}}$ *). Dashed contours indicate negative values.*

though the observed deceleration is again about twice the size of the modelled one. This weakness of the modelled induced cross-equatorial flow is probably due to the lack of a calculated PV flux at low latitudes, where the PV flux due to breaking planetary waves is expected to be large (see next chapter). A similar problem was encountered by Gray and Pyle (1987), who had to specify an additional easterly forcing near the equator to simulate the observed semi-annual oscillation (SAO). An increased PV flux near the equator in the winter (or spring) hemisphere would induce a stronger equatorward flow in the summer (or autumn) hemisphere, and so increase the overall deceleration of $\bar{u}$ near the equator, and broaden the latitudinal extent of the SAO, as noted by Holton and Wehrbein (1980). However, there should also be a significant annual cycle in the equatorial $\bar{u}$, as in the equatorial

tropopause temperature, with $\overline{u}$ being more easterly in January than July, due to the PV flux from September till February being stronger than from March till August.

We therefore claim that the cross-equatorial flow accounts for 50% of the modelled difference between the NH and SH autumnal westerly jets. In the real atmosphere the cross-equatorial flow could be more important than in the model if, as suggested by chapter 4, there is a significant negative $\mathcal{F}$ in the NH in November and December, since the $\mathcal{F}$ applied to the model was small in these months.

7.5 Conclusions and Summary

The calculation of $\nabla.F$, or PVF, from winds estimated from geopotential height data is known to be fraught with difficulties. The work of this chapter indicates that there is a systematic positive error in our calculations of PVF which is correlated with regions of large westerly $\overline{u}$. This is supported by the results of a model run with the calculated unsmoothed PV fluxes, the expectation that planetary waves should produce a negative PVF in the long-term average, and the results of Chapter 4. This positive error sometimes results in a positive PVF and at other times just reduces the negative PVF at mid-latitudes leaving a double peak in the negative PVF. It also appears that when $\tilde{u}$ (and not $\overline{u}$) is small or easterly, the PVF is small.

The PVFs were applied to the model of Chapter 5, after the positive regions had been removed and the result smoothed, and a comparison of the model $\overline{u}$ and $\overline{T}$ fields with observations was used to test the authenticity of these PVFs. It seemed that a real negative PVF was absent from our calculations in the NH in November and December, in the SH in July, August and September, and in the summer lower stratosphere of both hemispheres, due in part to the obscuring effect of the systematic error. There was fairly good agreement between model and observations in spring, though, indicating a good estimate of the PVF for late winter/early spring.

However, features of $\mathcal{F}$ deduced from the behaviour of our model should be treated with a similar degree of reserve as the $\mathcal{F}$ deduced by Shine (1989) and Yang et al. (1990) (which produced the observed rate of change of $\overline{u}$ and $\overline{T}$ when acting in conjunction with the calculated diabatic heating rate) since both will be sensitive

to errors in the heating rate calculation.

Model results indicated a probable deficiency in the heating calculations in the lower stratosphere over autumn and summer, and a poor parametrisation of gravity wave drag near the stratopause in winter.

The model reproduced the observed behaviour of the tropical tropopause and the effects of a cross-equator flow induced by a strong negative PVF.

Chapter 8

PV flux due to dissipation

The fight for life, the trials, the storms.
Is Nature true, or is she coy?
The waves which with this hand she forms
with other hands she will destroy.

8.1 Introduction

In Chapter 2 the two main ways believed to be the cause of a time-averaged horizontal eddy flux of PV in the stratosphere were described - namely parcel dispersion through planetary wave breaking, and dissipation of zonal asymmetries of PV.

Parcel dispersion by adiabatic planetary waves would produce a flux of any quantity down its mean meridional gradient in isentropic coordinates, but for any *tracer* the ratio of the flux to the gradient would be the same (being the negative of the tracer diffusion coefficient). Dispersion has been estimated from model and atmospheric data (see Plumb and Mahlman 1987, Kohno 1984, Lyjak 1987) and found to be large near the $\overline{u}$=0 line, which is generally at low latitudes, and smaller at high latitudes.

On the other hand, a steady, non-breaking wave will produce a meridional flux of any quantity whose zonal asymmetries are being dissipated. In this case, the flux-gradient ratio will depend on the time scale of the dissipation, so the 'diffusion coefficient' will be different for each quantity in general. The flux will be largest where the wave amplitude is largest, which is at about 60-70 degrees N or S (see

Barnett and Corney 1985b). The results of chapters 4 and 7 indicated a peak PV flux at mid-to high-latitudes, so it would appear that a large part of the PV flux at high latitudes is due to dissipation of PV. In fact, the estimate of the part of the PV flux due to dissipation of PV obtained in this chapter is generally *larger* than the total PV flux estimated in Chapter 7.

Furthermore, Kida (1983) and Tung (1984) have estimated that the K_{yy} due to a meridional dispersion of air parcels should be about $3x10^5 m^2 s^{-1}$, and Ko et al. (1984) produced realistic distributions of tracers in their model using this small value for K_{yy}. Newman et al. (1988) estimated the value needed to parametrise the eddy flux of quasi-geostrophic potential vorticity, using the flux-gradient method, to be over $40x10^5 m^2 s^{-1}$ in the NH winter upper stratosphere (though this is an over-estimate due to their use of almost-geostrophic winds). Also, Smith et al. (1988) and Newman et al. (1988) estimated the part of K_{yy} due to dissipation of PV to exceed $15x10^5 m^2 s^{-1}$ in the NH winter. This strengthens the claim that, at middle latitudes, dissipation of PV is more important than a meridional dispersion of air parcels in producing a PV flux.

In this chapter the part of the eddy flux of PV due to thermal dissipation of PV is estimated, from estimates of the horizontal displacements of air parcels and calculations of Π. It is found to be up to 50% larger than the estimates of PV flux obtained in Chapters 4 and 7, though this is an over-estimate. In section 8.2 the question of whether PV anomalies are in fact dissipated in the stratosphere is discussed. In section 8.3 the method of estimating the PV flux is described. Section 8.4 reviews previous estimates of the dissipative PV flux. Sections 8.5 and 8.6 discuss possible causes of error in our estimation of the dissipative flux, and present the estimate of the flux. Section 8.7 compares the flux with the results of chapters 4 and 7, and describes the effect of this flux on the model. The chapter is concluded and summarised by section 8.8.

8.2 Dissipation of PV

The approximation, by Newtonian cooling, of radiative dissipation of stratospheric temperature anomalies on isentropic surfaces appears to be well founded (see Pawson et al. 1991). However, dissipation of PV anomalies, although it has been assumed by several authors, (McIntyre and Norton 1990, Smith et al. 1988, Hitchman and Brasseur 1988) has not yet been shown to occur in the real atmosphere.

Therefore, the question posed in this section is whether PV anomalies are in fact dissipated. This will be answered by asking another question namely, whether thermal dissipation of stratospheric planetary waves can by itself lead to a dissipation of Π', the eddy PV.

Because of the lack of evidence of gravity wave breaking in the mid- to high-latitude stratosphere, and because dispersion due to planetary wave breaking seems to occur at low latitudes (see Plumb and Mahlman 1987, Kohno 1984, Lyjak 1987) near the $\overline{u} = 0$ line, it seems reasonable to neglect frictional body forces acting on the stratospheric winds at mid-latitudes. Therefore we ask whether thermal processes alone can lead to a dissipation of eddy PV.

Thermal dissipation of T' will lead to a dissipation of ρ', since ρ' depends directly on T'. However, this need not automatically lead to a dissipation of Π'. To first order in eddy quantities,

$$\overline{\rho}\,\Pi' \simeq (\Pi\rho)' - \overline{\Pi}\,\rho' = \zeta' - \overline{\Pi}\,\rho', \tag{8.1}$$

where ζ' is the eddy relative vorticity. Hence if there is no correlation between ρ' and ζ', dissipation of ρ' need not lead to dissipation of Π'. However, if ρ' is shown to be anti-correlated with ζ', then it seems reasonable to deduce that there is a process working to maintain this anti-correlation, and hence that as ρ' is thermally dissipated, ζ', and therefore Π', will be reduced towards zero.

Plots of ρ', ζ', Π' and Π, derived from the approximate winds described in Chapter 2, on the 850K isentropic surface (at about 30km) in the NH on 25th Jan 1981 are shown in fig.8.1. The anti-correlations between ρ' and ζ', and ρ' and Π' can be clearly seen. Correlation coefficients C(ρ',ζ') and C(ρ,Π') are shown in fig.8.2, where $C(X', Y') \equiv \overline{(X'Y')}/(\overline{X'^2}\,\overline{Y'^2})^{1/2}$. Thus in the NH winter stratosphere north of about 50N, where planetary waves are large, ρ' is well anti-correlated (C less than -0.8) with ζ'. The anti-correlation between ρ' and Π' extends over a greater area, due to the explicit contribution of ρ' to Π' in eqn.8.1. This is typical of the correlation coefficients throughout the NH winter. In the SH, which is in summer, there is no significant correlation because of the absence of large planetary waves.

It is worth emphasising that although Π' would be expected to have some correlation with ρ' because of the explicit dependence of Π' on ρ', the large correlation between ρ' and ζ', the eddy relative vorticity, requires some explanation. The explanation offered here is simply that thermal wind balance links the wind and temperatures. Consider an atmosphere which is zonally symmetric apart from a

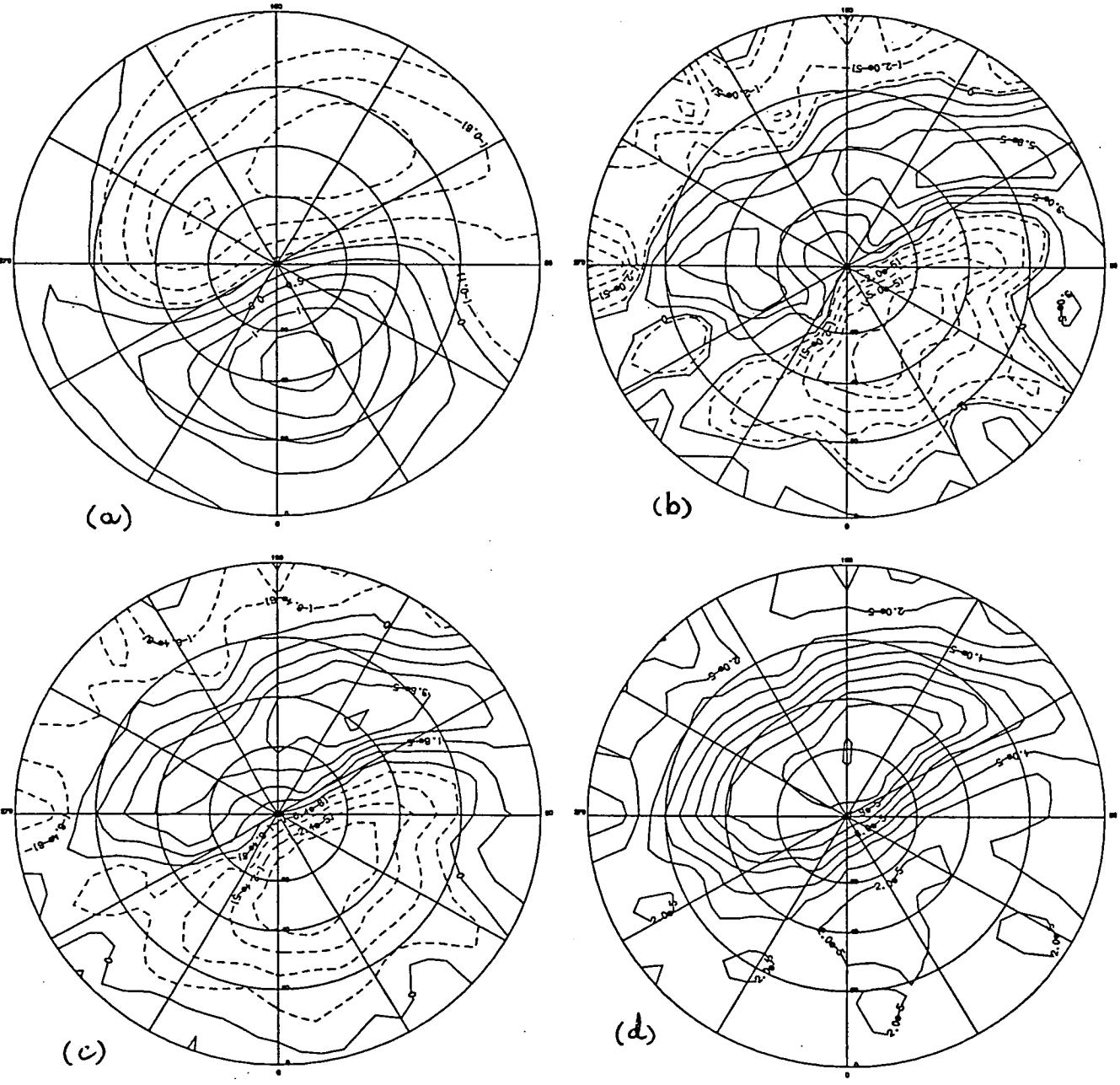


Figure 8.1: *(a) ρ', (b) ζ', (c) Π' and (d) Π (all in arbitrary units, with dashed contours indicating negative values) on the 850K isentropic surface, from 30N to 90N, on 25th Jan 1981.*

small cylindrical region where $\rho' < 0$ (see fig.8.3). It might be thought that there need be no PV anomaly, if the density anomaly is balanced by a negative relative vorticity anomaly. Thermal wind balance, however, requires there to be a *positive* relative vorticity anomaly, and thus a positive PV anomaly. The gradients of

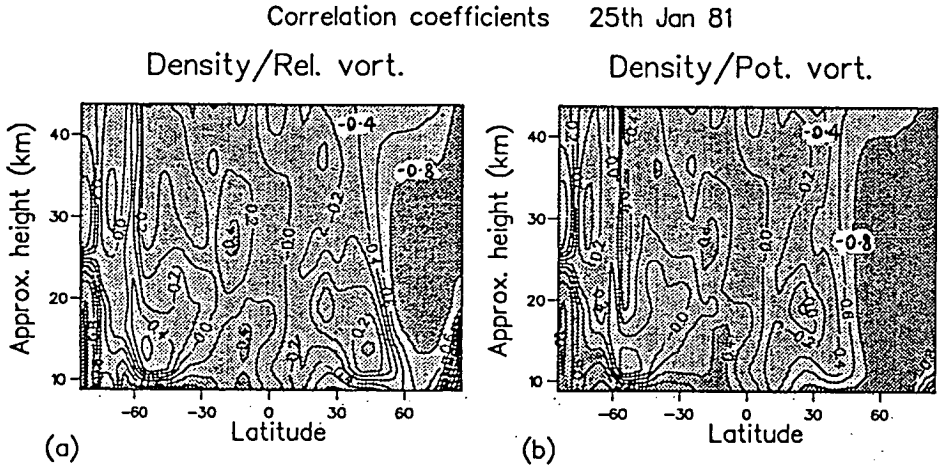


Figure 8.2: *Correlation coefficients (a) $C(\rho',\zeta')$ and (b) $C(\rho',\Pi')$ on 25th Jan 1981.*

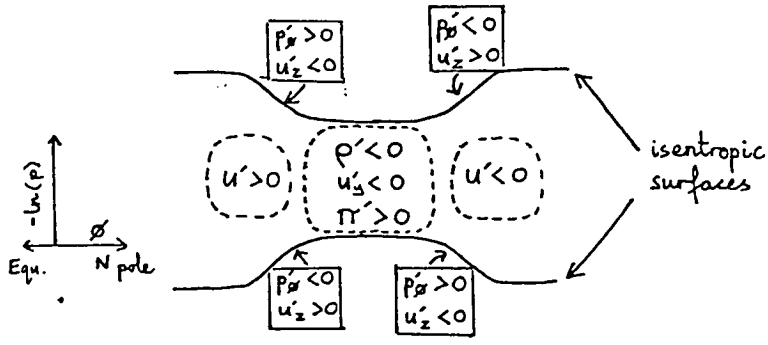


Figure 8.3: *Meridional section through positive PV anomaly 'due' to a negative density anomaly*

pressure along the isentropic surfaces are indicated in the square boxes in fig.8.3. Thermal wind balance then determines the vertical gradients of u′, and thus the values of u′ in the dashed boxes. These determine u'_y in the centre of the anomaly, and hence Π′ (see eqn. 8.1). The real atmosphere does not have only one small density anomaly, however, and since thermal wind balance links only the *vertical gradient* of the winds to the temperature, the correlation between eddy density and eddy relative vorticity depends on the state of the whole atmosphere. The correlation *is* observed, however, and thermal wind balance seems to be a plausible mechanism for its production.

Therefore it is assumed here that PV anomalies are dissipated. It is also assumed that the time constant for PV dissipation equals the radiative time constant, which we shall call $1/\alpha$. Many studies have been made of α (eg. Gille and Lyjak(1986), Dickinson (1973), Pawson et al. (1991)), with attempts to include

the scale-dependence. Gille and Lyjak obtained the time constants of figure 8.4(a) by perturbing the observed temperature by a value constant over the whole atmosphere, calculating the resulting change in heating rate and then using the relation $\alpha \equiv -\Delta Q/\Delta T$. Note the similarity to figure 7.4(f), at least near the equator, de-

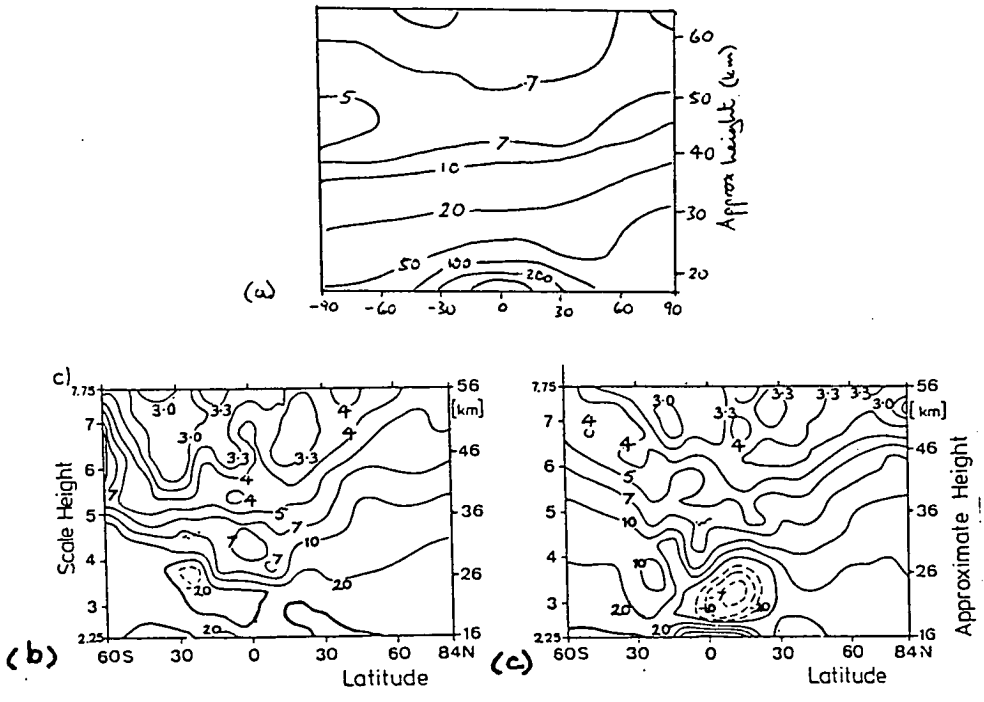


Figure 8.4: *Radiative time constants, in days, of (a) Gille and Lyjak, and of Pawson et al. for (b) January 1979 (c) April 1979*

rived from our model. They obtained much shorter time constants, however, by perturbing the temperature of a single layer of the atmosphere, finding $1/\alpha \simeq 3$ days near the stratopause and 20 days in the tropical lower stratosphere. Pawson et al. calculated α from the relation

$$\alpha = -\frac{\overline{Q'T'}}{\overline{T'^2}}$$

which does not require the temperature to be artificially perturbed. They estimated Q using satellite measurements of temperature, ozone and water vapour, and a radiation scheme which included cooling due to CO_2, ozone and water vapour, and the solar heating due to ozone. Their results, which were used in this study after being smoothed, are shown in figures 8.4(b) and (c).

8.3 Estimation and parametrisation of dissipative flux

In this section an expression will be derived for the eddy flux of PV due to dissipation in terms of a dissipative time constant, a displacement and Π. The parametrisations of this flux by Smith et al.(1988), Matsuno(1980), Garcia and Solomon(1983) and Pyle and Rogers(1980) will also be described. The linearised equation for the evolution of Π' is

$$\Pi'_t + \overline{u}\,\Pi'_x + v'\overline{\Pi}_y = -\alpha\,\Pi' \tag{8.2}$$

where the vertical advection term is neglected since the waves are assumed to be adiabatic and isentropic coordinates and are being used. Defining the operator $D_t = \frac{\partial}{\partial t} + \overline{u}\,\frac{\partial}{\partial x}$ and a meridional displacement η such that $v' = D_t\,\eta$, and assuming that $\overline{\Pi}_t << \Pi'_t$, gives

$$D_t(\Pi' + \eta\,\overline{\Pi}_y) = -\alpha\,\Pi'. \tag{8.3}$$

Multiplying by η and averaging (and noting that $\overline{(\beta\,D_t\gamma)} = (\overline{\beta\,\gamma})_t - \overline{(\gamma\,D_t\beta)}$) gives

$$-\alpha\,\overline{\eta\,\Pi'} = (\overline{\eta\,\Pi'})_t - \overline{v'\Pi'} + \frac{1}{2}(\overline{\eta^2})_t\,\overline{\Pi}_y.$$

If the waves are steady (ie: if the time-derivative of the zonal-mean of the product of two eddy quantities is zero) the first and third terms on the right-hand side vanish, leaving

$$\overline{v'\Pi'} = \alpha\,\overline{\eta\,\Pi'}. \tag{8.4}$$

From Chapter 2, however, the flux needed is $\overline{v^*\rho\Pi^*}$. Writing

$$\Pi'\overline{\rho} = \Pi^*\rho + \text{2nd order terms},$$

and since $\overline{v'\rho\Pi^*} \equiv \overline{v^*\rho\Pi^*}$ and $\overline{\eta\rho\Pi^*} \equiv \overline{\eta^*\rho\Pi^*}$ we have, from 8.4,

$$\overline{v^*\rho\Pi^*} = \alpha\overline{\eta^*\rho\Pi^*} \tag{8.5}$$

which is the equation which will be used for estimating the PV flux due to dissipation.

Now it will be shown how this flux can be parametrised in terms of the zonal-mean gradient. From eqn.8.3, if α is small, $\Pi' \simeq -\eta\,\overline{\Pi}_y$ and so

$$\overline{v'\Pi'} \simeq -\alpha\,\overline{\eta^2}\,\overline{\Pi}_y \tag{8.6}$$

This allows a parametrisation of $\overline{v'\Pi'}$ in terms of $\overline{\Pi}$, provided η can be estimated.

In considering the case where α is not small it is fruitful to look for wave solutions $\eta = A\sin\theta$ and $\Pi' = B\sin(\theta+\delta)$ where $\theta = kx - \omega t$ and B and δ are as yet undetermined. Substituting in (8.2), expanding the cos and sin terms and equating coefficients of $\cos\theta$ and $\sin\theta$ we obtain the two equations

$$(\overline{u}k - \omega)(B\cos\delta + A\overline{\Pi}_y) = -\alpha B\sin\delta$$

$$(\overline{u}k - \omega)\sin\delta = \alpha\cos\delta$$

From these it is possible to determine δ and B. However, it is only necessary to deduce from them that

$$B\sin\delta = -\frac{\alpha A\overline{\Pi}_y(\overline{u}k - \omega)}{\alpha^2 + (\overline{u}k - \omega)^2}$$

since then,

$$\begin{aligned}\overline{v'\Pi'} &= (\overline{u}k - \omega)\,A\,B\,\overline{\cos\theta\sin(\theta+\delta)} \\ &= \frac{1}{2}(\overline{u}k - \omega)\,A\,B\sin\delta \\ &= -\frac{A^2}{2}\frac{\alpha(\overline{u}k - \omega)^2}{\alpha^2 + (\overline{u}k - \omega)^2}\overline{\Pi}_y \qquad (8.7)\end{aligned}$$

If $\alpha \ll \overline{u}k - \omega$ we recover (8.6), since $\overline{\eta^2} = A^2/2$. Equation (8.7) is similar to expressions derived by Matsuno (1980), Smith et al.(1988), Garcia and Solomon (1983), Pyle and Rogers (1980), though because they were not using isentropic coordinates they considered vertical displacements also and so had a 2x2 transport matrix.

8.4 Previous estimates of the dissipative potential vorticity flux

Smith et al. (1988) estimated the dissipative flux of quasi-geostrophic potential vorticity (QPV) using an equation similar to 8.7. The displacement of a parcel in a ring of parcels all having the same potential temperature and geopotential was its displacement from the mean latitude and height of that ring. They used two dissipative time constants, the first being Gille and Lyjak's values shown in figure 8.4(a) while the second was one used by Hitchman and Brasseur(1988). This was much faster than Gille and Lyjak's, and even Pawson et al.'s(1991), being 9

days at 15km and 1 day at 60km and seemed to be chosen rather arbitrarily with little physical justification. Furthermore, it was intended to be the time constant for dissipation of wave-activity, which is proportional to $\overline{\Pi'^2}$, and thus is half as long as the time constant for dissipation of Π'. Using the first set of time constants, Smith et al. obtained a QPV flux in January with a peak at 50N, 45km of about -5 $ms^{-1}day^{-1}$, with values in November and March of about -1 $ms^{-1}day^{-1}$, which is similar to the values estimated in Chapters 4 and 7. However, comparing these fluxes to their estimate of the total flux calculated using geostrophic winds, which exceeded -20 $ms^{-1}day^{-1}$ in January, they concluded that they had underestimated the dissipation. With the faster dissipation of Hitchman and Brasseur they obtained values much closer to their geostrophic estimate. However, as noted by Randel(1987) and Robinson(1986), geostrophic winds severely overestimate $\nabla.F$. It seems, then, that their use of Hitchman and Brasseur's dissipation time constants leads to an overestimation of the dissipative QPV flux.

Newman et al.(1988) estimated the part of K_{yy} due to dissipation of QPV for the NH winter, using the dissipation rate of Dickinson (1973), the variance of the meridional velocity, and $\overline{u}$. They found it to be about 30% of their estimate of the total K_{yy} which they derived using the flux-gradient method, from the PV flux calculated from geostrophic winds. However, as noted above, geostrophic winds severely overestimate $\nabla.F$. Therefore Newman et al. probably underestimated the importance of dissipation in producing a QPV flux.

8.5 Some sources of error in estimating the PV Flux

Before examining our results, some ways in which our estimate of the dissipative PV flux might be in error will be considered.

(1) Errors in the calculation of PV, especially when the meridional gradient of PV is weak. In this case a small error in PV would result in a large error in the estimation of the displacement η, since η is defined as the displacement between the parcel's current latitude and the latitude at which the parcel's PV appears in the symmetric vortex (see section 8.6). This should result in an over-estimation of $\overline{\eta'\Pi'}$ since it is expected that Π' and η' will be anti-correlated (a positive Π' probably indicating a parcel which has been displaced from the north, and so having a negative η').

(2) An error in thermal dissipation time constant, through insufficient data coverage (the LIMS data used by Pawson et al. (1991) only extends over nine months and northwards of 60S), errors in the heating rate calculation and the neglect of scale-dependence. As noted in section 8.2, Gille and Lyjak (1986) found that scale-dependence could reduce their time constants by about 50% near the stratopause.

(3) Incorrect assumption that zonal asymmetries of PV are dissipated.

(4) The neglect of geostrophic balance in estimating the effect of diabatic heating on PV. This would also lead to an overestimation of the negative PVF. Consideration of this last point requires a section to itself.

8.5.1 Lengthening of the thermal dissipation time-constant due to local geostrophic effects

In writing eq.(8.2), and in deriving eq.(8.5), Π' has been treated as if it were a dynamically passive quantity. It is not, however, and changes in PV affect the winds and the temperature.

Consider a cylindrical PV anomaly, the cross section of which, seen in pressure coordinates, is shown in figure 8.5. The top of the anomaly is warmer than neighbouring points on the same isentropic level. Likewise, the bottom is colder than

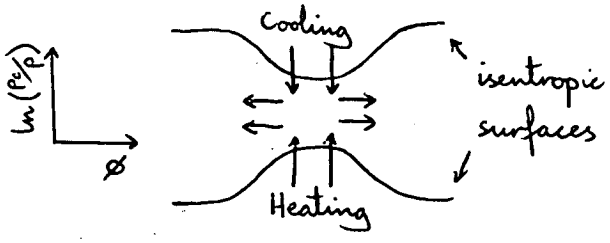


Figure 8.5: *PV anomaly and meridional circulation induced to maintain geostrophic balance*

its neighbouring points. If diabatic heating now acts to cool the top and warm the bottom, there will be a convergence of the vertical wind at the centre of the anomaly. Now, the change in temperature of the top and bottom levels will require a change in the component of horizontal velocity perpendicular to the radius of the anomaly (and hence in the relative vorticity) to maintain geostrophic balance. Neglecting vertical advection, this change in vorticity can only be achieved

by a horizontal divergence (see eqn. 2.14). This divergence will reduce the rate of change of the density of the anomaly, and hence reduce the effect of the diabatic heating.

For a more quantitative treatment consider the box in isentropic coordinates shown in fig.8.6, which could be taken as the top right-hand quarter of fig.8.5. The

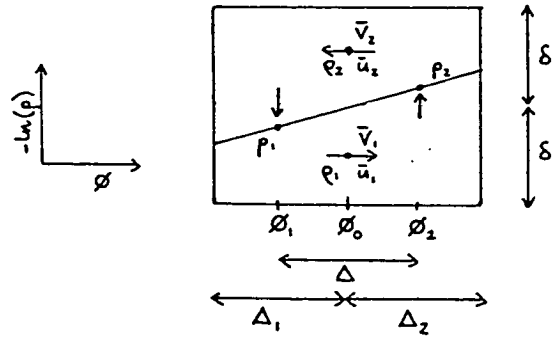


Figure 8.6: *Top right-hand quarter of previous figure.*

following restrictions will be imposed.

(1) The pressures along the top and bottom isentropic surfaces are uniform and do not change in time, and there is no mass flux across these surfaces.

(2) There is no mass flux across the sides of the box.

(3) The third isentropic surface shown divides the box in two, the total mass of air in the top half of the box equalling that in the bottom half throughout the experiment.

These three assumptions imply

$$\overline{p}_{1t}\cos\phi_1 \simeq -\overline{p}_{2t}\cos\phi_2 \tag{8.8}$$

and

$$\overline{V}_1 \simeq -\overline{V}_2 \tag{8.9}$$

These relations are approximate since the pressures and winds are values at a point but are taken to represent average values along horizontal and vertical lines. Δ, Δ_1 and Δ_2 are defined so that

$$\Delta_1\cos\phi_1 \equiv \Delta_2\cos\phi_2 \equiv \Delta\cos\phi_0, \tag{8.10}$$

which ensures that the masses of air inside the box north and south of ϕ_0 are approximately equal to each other and to that between ϕ_1 and ϕ_2. A Newtonian cooling of $\overline{p}_1$ and $\overline{p}_2$ towards a pressure p_0 (which is restricted by eqn. 8.8), with a thermal time constant α, is assumed. Therefore, use of the finite difference form of the mass continuity equation gives

$$-\frac{1}{g\delta}(\overline{p}_{1t} + \alpha(\overline{p}_1 - p_0))\cos\phi_1 = -\frac{\overline{V}_1}{\Delta_1}. \tag{8.11}$$

The meridional velocities $\overline{V}_1$ and $\overline{V}_2$ will accelerate $\overline{u}_1$ and $\overline{u}_2$. Neglecting $\mathcal{F}$ and the non-linear terms in the momentum equation gives

$$\overline{u}_{1t} \simeq af\frac{\overline{V}_1}{\overline{\rho}_1\cos\phi_1} \tag{8.12}$$

Assuming that $\overline{\rho}_1 \simeq \overline{\rho}_2$, and using 8.9, it follows that

$$\overline{u}_{2t} \simeq -\overline{u}_{1t} \tag{8.13}$$

The time-derivative of the thermal wind equation (5.4) in finite-difference form is

$$\frac{\overline{u}_{2t} - \overline{u}_{1t}}{\delta}\left(\frac{u_0 + a\Omega\cos\phi_0}{a\Omega\cos\phi_0}\right) = -\frac{c_p\kappa T}{afp}\left(\frac{\overline{p}_{2t} - \overline{p}_{1t}}{\Delta}\right),$$

where $u_0 = (\overline{u}_1 + \overline{u}_2)/2$ and so, from 8.13, $u_{0t} = 0$. Also, eqn.8.8 implies that $\overline{p}_t \simeq 0$ and $\overline{T}_t \simeq 0$ at the centre of the box. Using 8.8, 8.13 and 8.10 gives

$$\overline{u}_{1t} \simeq -\frac{\delta c_p\kappa T\overline{p}_{1t}}{2afp\Delta}\left(1 + \frac{\cos\phi_1}{\cos\phi_2}\right)\left(\frac{a\Omega\cos\phi_0}{u_0 + a\Omega\cos\phi_0}\right) \tag{8.14}$$

Rewriting 8.11, using 8.12 and 8.14 and assuming that $\overline{\rho}_1 \simeq p/(g\kappa)$ (the isothermal atmosphere approximation) leads to

$$\overline{p}_{1t}\left(1 + \left(\frac{\delta}{\Delta af}\right)^2\frac{c_pT}{2}\left(\frac{\cos\phi_0}{\cos\phi_1}\right)^2\left(1 + \frac{\cos\phi_1}{\cos\phi_2}\right)\left(\frac{a\Omega\cos\phi_0}{u_0 + a\Omega\cos\phi_0}\right)\right) \simeq -\alpha(\overline{p}_1 - p_0) \tag{8.15}$$

Except near the poles, $\cos\phi_0 \simeq \cos\phi_1 \simeq \cos\phi_2$ and $a\Omega\cos\phi_0 \gg u_0$ so that, for mid-latitudes, the time constant $1/\alpha$ is effectively multiplied by the factor

$$\left(1 + \left(\frac{\delta}{\Delta af}\right)^2 c_pT\right) \simeq \left(1 + \frac{0.29}{\sin^2\phi}\left(\frac{\delta}{\Delta}\right)^2\right) \tag{8.16}$$

Our theory should be invalid within perhaps 20 degrees of the equator, due to the breakdown of equations 8.14 and 8.12. However, excluding this equatorial region

we note the remarkable fact that, because of the $\frac{1}{\sin^2\phi}$ factor in 8.16, the lengthening of the time constant by the geostrophically-induced circulation is *greatest* at *low* latitudes. One $\sin\phi$ factor comes from 8.14 which reveals that, for a given $\overline{p}_\phi$, $\overline{u}_z$ is greatest near the equator. The other $\sin\phi$ comes from 8.12 which shows that in order to bring about the same change in $\overline{u}$ at low latitudes as at high latitudes a greater $\overline{V}$ is required. The combined effect is that at low latitudes the induced $\overline{V}$ is greater than at high latitudes for the same latitudinal gradient of heating and hence, in 8.11, the opposition to the diabatic heating is stronger and $\overline{p}_t$ is smaller.

The scale of a real anomaly is difficult to estimate, but for a Δ of 10 degrees and a δ of about 0.17 (corresponding to about 4km) the ratio (δ/Δ) is close to unity. However, this ratio appears squared in 8.16 and hence the lengthening of the time-constant is very sensitive to the shape of the anomaly, being greater for tall, narrow PV anomalies. Added to this is the fact that the thermal time-constant itself is scale-dependent (see section 8.2).

Hence the amount by which the dissipative PV flux is over-estimated is difficult to determine, but it is over-estimated more at low-latitudes than at high latitudes. For $(\delta/\Delta) = 1$, the flux is overestimated by a factor of 2.2 at 30 degrees while at 60 degrees the factor is 1.39.

8.6 The estimated dissipative flux

In order to estimate the PV flux due to dissipation using equation 8.5 the displacement η must first be estimated. By the nature of η, which is supposed to be a small displacement, its estimation must be somewhat arbitrary, since the distortion of the vortex is far from small. In our method, the undisplaced state is taken to be the symmetric vortex described in chapter 3. The displacement is then assumed to be isentropic and is defined as the difference between the parcel's current latitude and the latitude at which the parcel's PV occurs in the symmetric vortex. This definition of the displacement is likely to under-estimate the displacement, since if the PV of the parcel is actually decreasing as it moves south (through the effects of dissipation) then it will appear to have come from further south than it actually did. A similar argument applies for a northward displacement. However, as noted in section 8.5, errors in PV, especially when the meridional gradient of PV is small, will lead to an over-estimation of the PV flux.

The flux estimated in this way from July 1980 to June 1981 is shown in figure 8.7. It has been smoothed in the same way as the fluxes of Chapter 7. Due to the approximations used in deriving $\mathcal{F}$ (displacements, dissipation rates and steady waves), it did not seem unreasonable to smooth the fluxes. The unsmoothed fluxes did not have any positive regions, but there was often a strong negative region at about 30N and 30S (see figure 8.8). This is presumably due to η being badly estimated nearer the equator, where gradients of PV are weak. Equatorwards of 30 degrees the estimated flux was set to zero before the whole field was smoothed. The unsmoothed fields were not particularly spikey (figure 8.8 is typical), but the smoothed fluxes did produce more realistic model winds (ie: with less horizontal shear).

We note the following features of the smoothed PV flux of fig. 8.7.

In the NH :-

(1) The peak value never falls below -1ms^{-1}day^{-1} . It is small (not exceeding -3ms^{-1}day^{-1}) from May till September, and in April it does not exceed -4ms^{-1}day^{-1} .

(2) It is large from October till March (exceeding -5ms^{-1}day^{-1} except in November) and reaches its peak of -12ms^{-1}day^{-1} in January. Over this period the values north of 50N are smallest in December. From the discussion in section 8.5.1 the lower-latitude values are expected to be over-estimated more than the high-latitude values, and hence the actual values in December might be the smallest over this period.

(3) The peak moves polewards from about 40N in December to about 65N in February, retreating to about 50N in March.

In the SH :-

(4) The peak value never falls below -2ms^{-1}day^{-1} . However, from November until August (but excluding April) the large (exceeding -4ms^{-1}day^{-1}) values occur equatorwards of 50S. Again, from the discussion in section 8.5.1 we might expect the real PVF for these months to be small.

(5) The PVF exceeds -4ms^{-1}day^{-1} polewards of 50S in September and October, its peak value of -8ms^{-1}day^{-1} occuring in September.

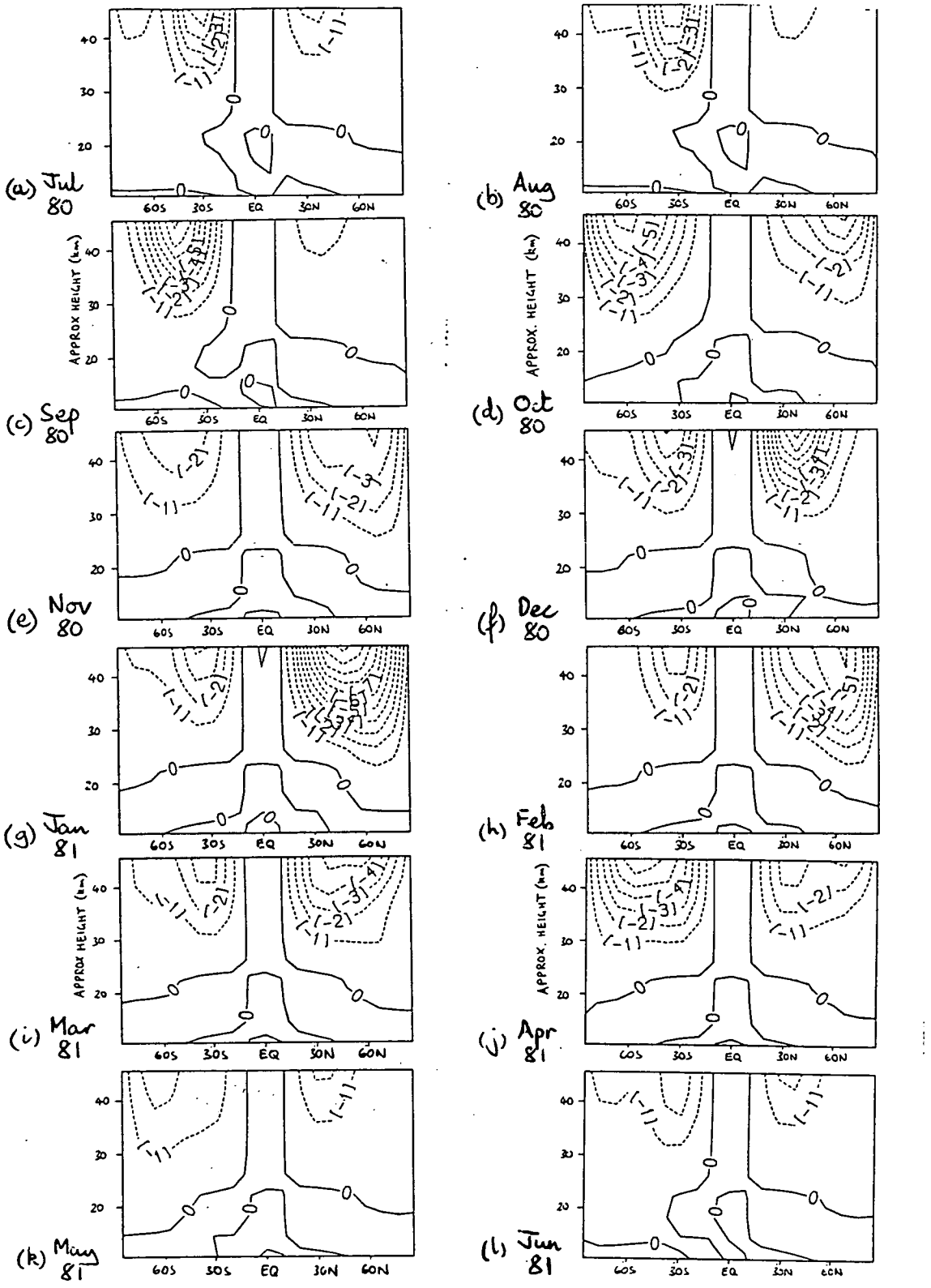


Figure 8.7: *PV flux due to dissipation, in* $\mathrm{ms^{-1}day^{-1}}$, *(a) July 1980 ... (l) June 1981*

(6) There is a large (-7$\mathrm{ms^{-1}day^{-1}}$) PVF in April.

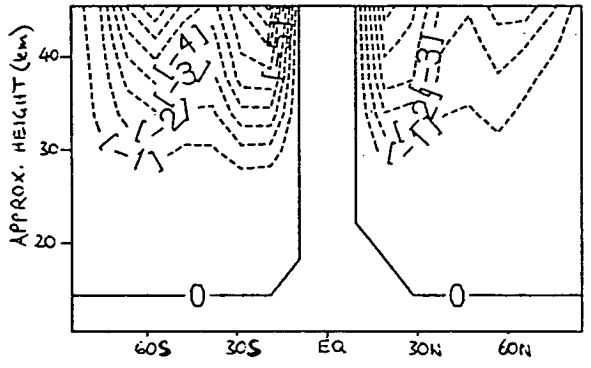


Figure 8.8: *Unsmoothed PV flux due to dissipation for April 1981*

(7) In the SH, the peak moves polewards from 30S in July to 60S in October.

8.6.1 Comparison with Chapters 4 and 7.

These values will now be compared with those of Chapters 4 and 7. A fuller discussion of the deduced features of the real PV flux will be presented in the final chapter, along with an assesment of each of the three methods' strengths and weaknesses.

(1) The 'quiet' values, over the summer and autumn, are larger than those in chapter 7 (and are larger in the SH than in the NH).

(2) Chapter 4 revealed a small $\mathcal{F}^{D}$ in mid-spring (Apr/Oct). The large values of PVF in April in the SH and October in the NH found here are not inconsistent with this finding. These large values were not, however, found in chapter 7.

(3) The occurrence of the largest PVFs in January and February in the NH, and in September and October in the SH, is in agreement with chapters 4 and 7, though the peak values are up to -5ms^{-1}day^{-1} larger.

(4) The large PVFs in the NH from October till December, and in March, are consistent with chapter 4 but are not found in chapter 7.

(5) The poleward movement from December to January in the NH agrees with chapter 4, as does the poleward movement in the SH from September to October.

The good agreement with the results of chapter 4 becomes more apparent on inspecting the IHD in the PVF fields (see fig.(8.9)) The values are generally very

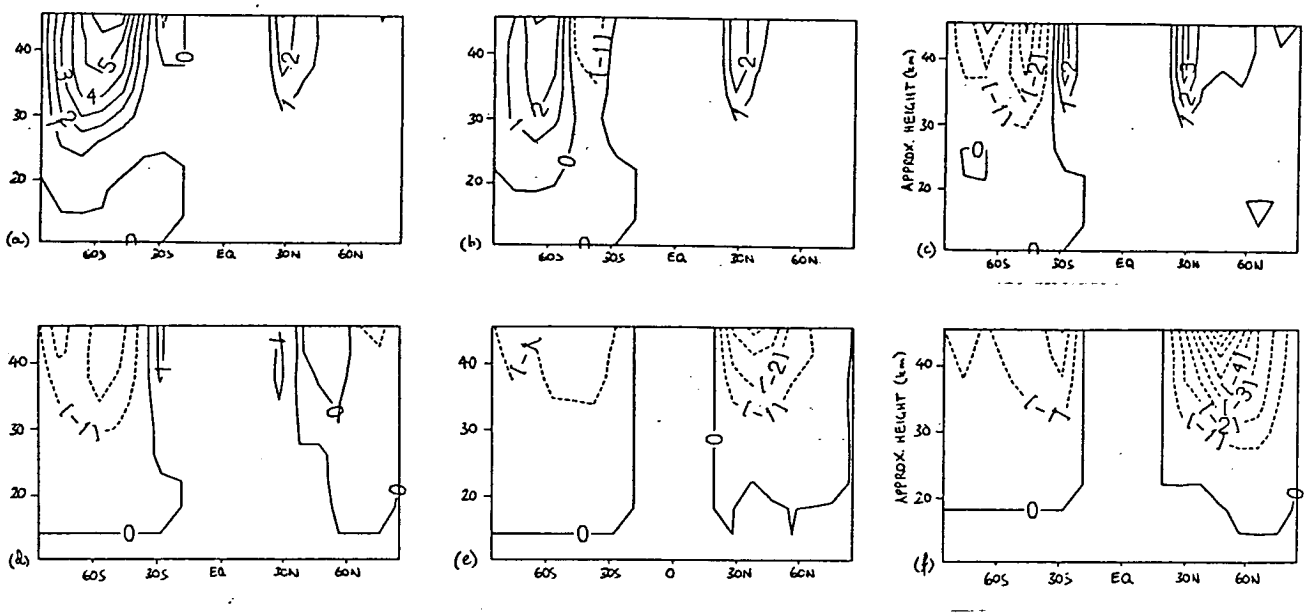


Figure 8.9: *IHD in the dissipative PVF (a) Jul/Jan (b) Aug/Feb (c) Sep/Mar (d) Oct/Apr (e) Nov/May (f) Dec/Jun in* $\mathrm{ms}^{-1}\mathrm{day}^{-1}$.

similar to those of fig.(4.5) (to within $1\mathrm{ms}^{-1}\mathrm{day}^{-1}$) all through the year. Exceptions occur in early winter (fig.8.9(f)) when the negative IHD is too large by about $3\mathrm{ms}^{-1}\mathrm{day}^{-1}$, and in late winter (fig8.9(b)) where the negative values are not as large or as dominant in the upper stratosphere. Also, over the summer and autumn the larger values of the dissipative PVF in the SH compared with the NH (point (1) above) result in positive values near 30 degrees on the right of figs.8.9(a), (b) and (c), and negative values on the left of fig.8.9(f).

8.7 Effect on the model

When these fluxes were applied to the model of Chapter 5, it produced the mid-seasonal zonal winds and temperatures shown in figures 8.10 and 8.11. Since the PV fluxes are larger than those of Chapter 7, the westerly winds should be weaker than those of that chapter. Indeed, because the peak in PV flux generally lies between 40 and 60 degrees, a minimum in the westerly $\overline{u}$ is seen at about 45N in January and at 50S in October. Because of such a strong $\mathcal{F}$, and no doubt also due to the shape of $\mathcal{F}$, the induced poleward $\overline{V}$ accelerates $\overline{u}$ strongly at the poles, resulting in the unrealistically large $\overline{u}$ at both poles for most of the year. Corresponding to the large vertical gradients of $\overline{u}$ at the poles are large horizontal gradients of temperature, consistent with thermal wind balance. Thus it seems reasonable to conclude from this that our estimate of PVF in this chapter is too

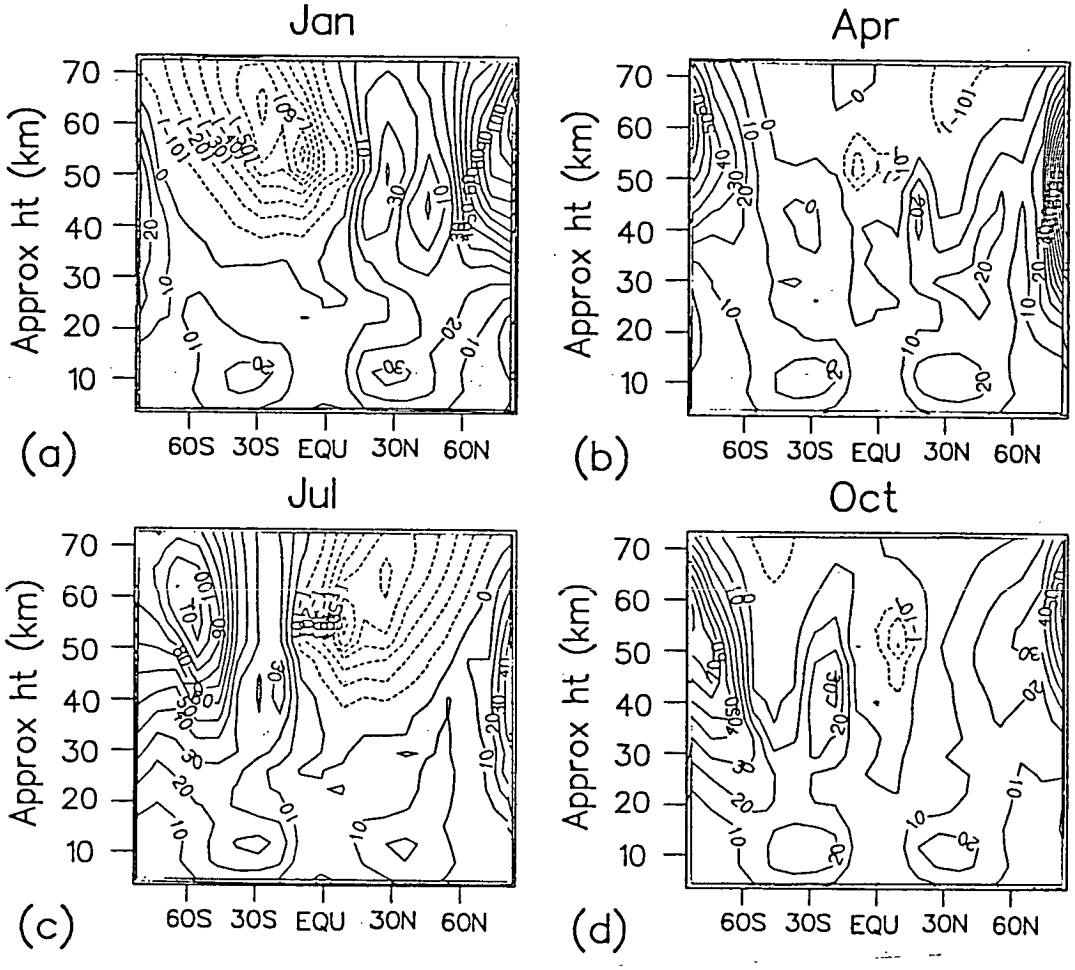


Figure 8.10: *Modelled mid-seasonal zonal winds using the dissipative flux (a) Jan (b) Apr (c) Jul (d) Oct. Dashed contours indicate easterly winds.*

large, and so the estimates of $\nabla.F$ in Smith et al's (1988) and Geller et al.'s (1984) studies of over -20 $ms^{-1}day^{-1}$ in January, calculated using geostrophic winds, must certainly be too large. This is supported by Randel's (1987) and Robinson's (1986) work in which it was demonstrated that geostrophic winds significantly over-estimate $\nabla.F$.

Because of the similarity between the IHD in the dissipative PVF and the $\mathcal{F}^D$ of chapter 4 it is informative to look at the modelled IHD in $\overline{T}$ (see fig.8.12). Comparing this with fig.(8.13), the observed IHDs, the following points are noted :-

(1) The magnitudes are fairly similar.

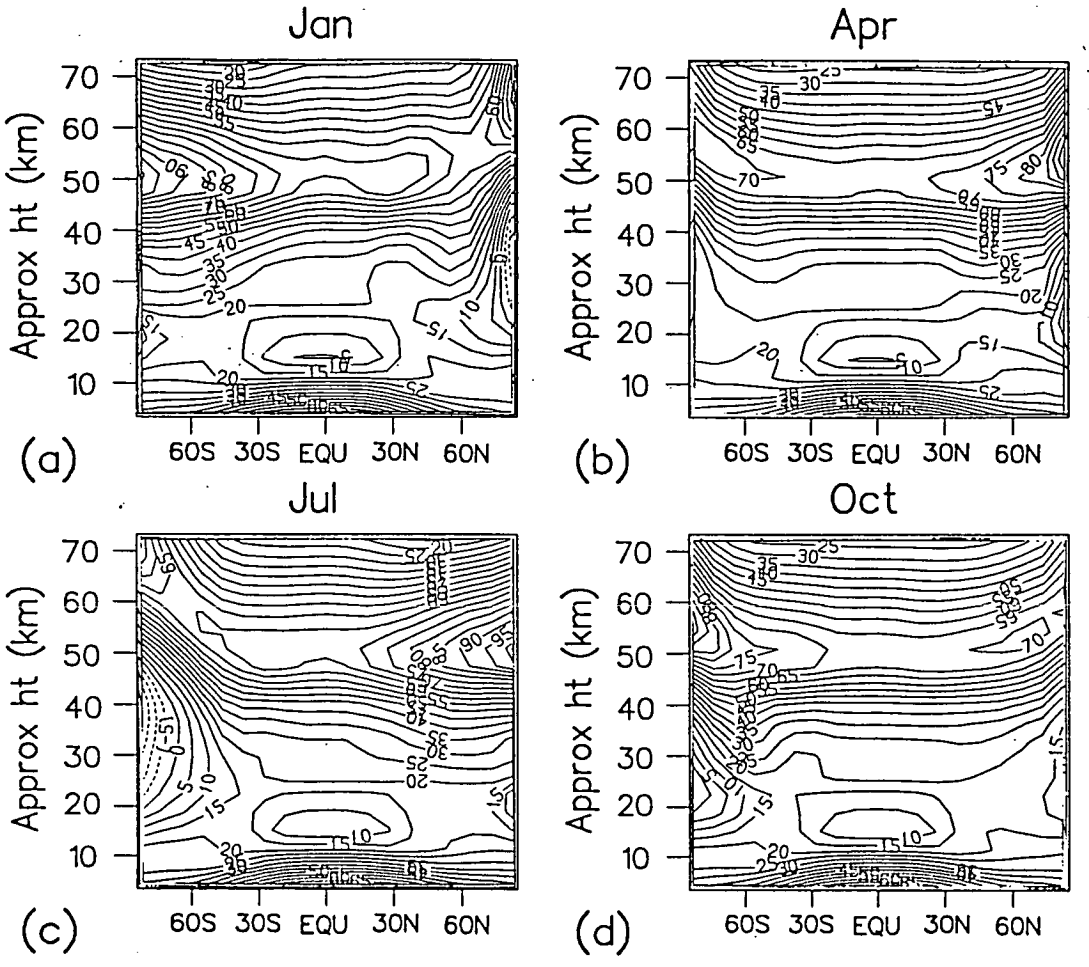


Figure 8.11: *Modelled mid-seasonal temperatures (in excess of 200K) using the dissipative PV flux (a) Jan (b) Apr (c) Jul (d) Oct. Dashed contours indicate temperatures below 200K.*

(2) The vertical movement of the patterns is not so pronounced as in the observed fields (especially in figures (a), (b) and (c)).

(3) The negative peaks in figs.8.12 (b), (c) and (d) occur at about 80 degrees, while those in fig.8.13 occur at the pole. This is connected with the behaviour at the poles mentioned in the previous paragraph.

Thus although the size of the flux is overestimated, the IHD in the flux seems to be reasonably accurate. The negative amount by which this estimate of the PV flux exceeds the actual flux, then, is about the same in both hemispheres at the same seasonal time. However, the fact that the IHD in the estimated PVF is fairly realistic implies that the dissipative flux could still account for most of the actual PVF.

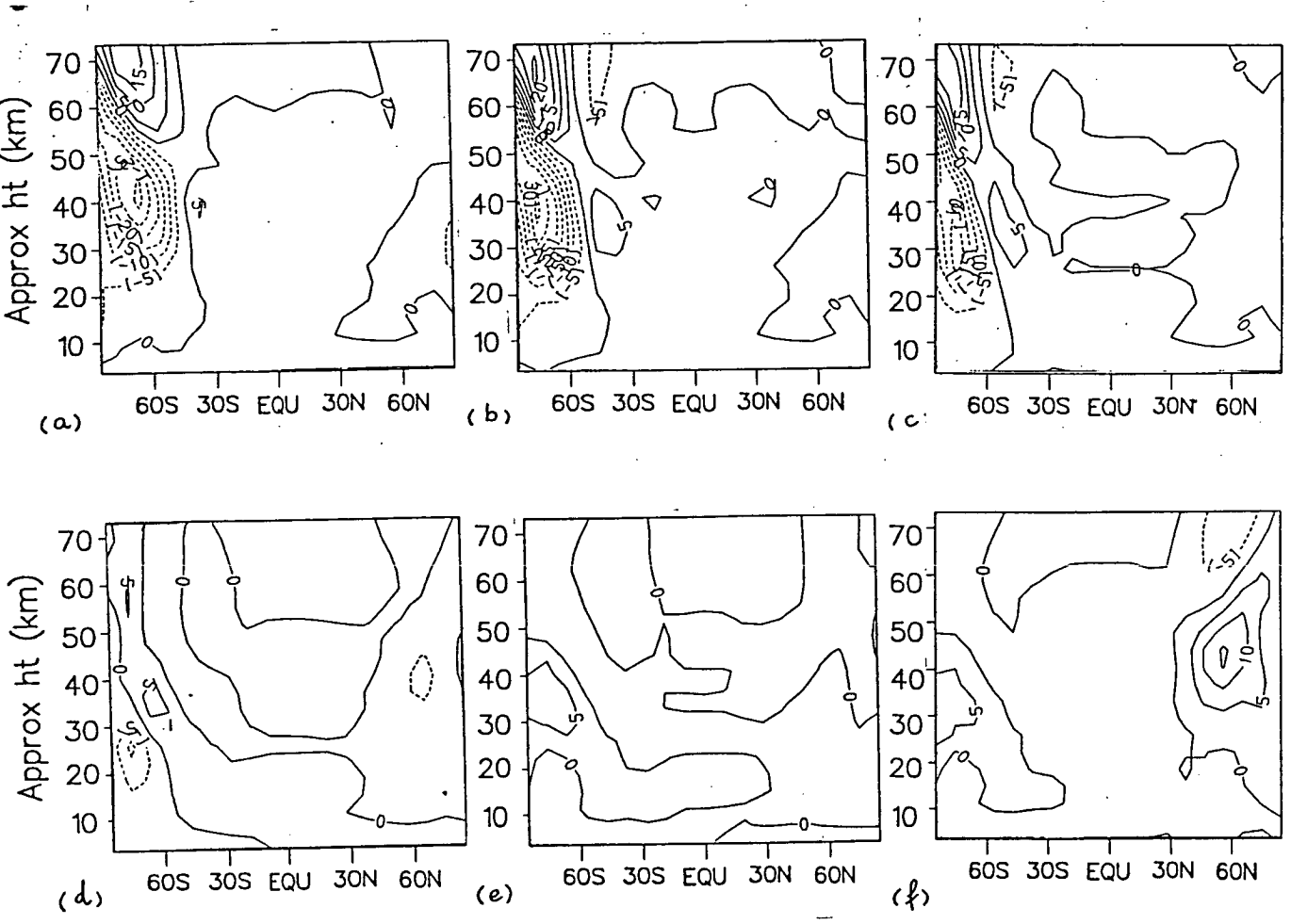


Figure 8.12: *IHD in modelled* $\overline{T}$ *for (a) Jul/Jan (b) Aug/Feb (c) Sep/Mar (d) Oct/Apr (e) Nov/May (f) Dec/Jun. Dashed contours indicate negative values.*

8.8 Summary and conclusions

In this chapter the assumption that thermal dissipation of temperature anomalies leads to a dissipation of PV anomalies, through the maintenance of thermal wind balance, was shown to be reasonably well justified. The time-constant for the dissipation of PV anomalies was therefore taken to be the same as that for the dissipation of temperature anomalies. It was shown, however, how local circulations, induced to maintain thermal-wind balance, can lengthen the time-constant for the dissipation of PV above that for the dissipation of temperature by a factor which increases towards the equator and which is greater for tall, thin PV anomalies. Our PV fluxes are therefore expected to be overestimated by this factor, particularly near the equator. Neglect of the scale-dependence of the radiative time constant could also result in significant errors. There is also expected to be

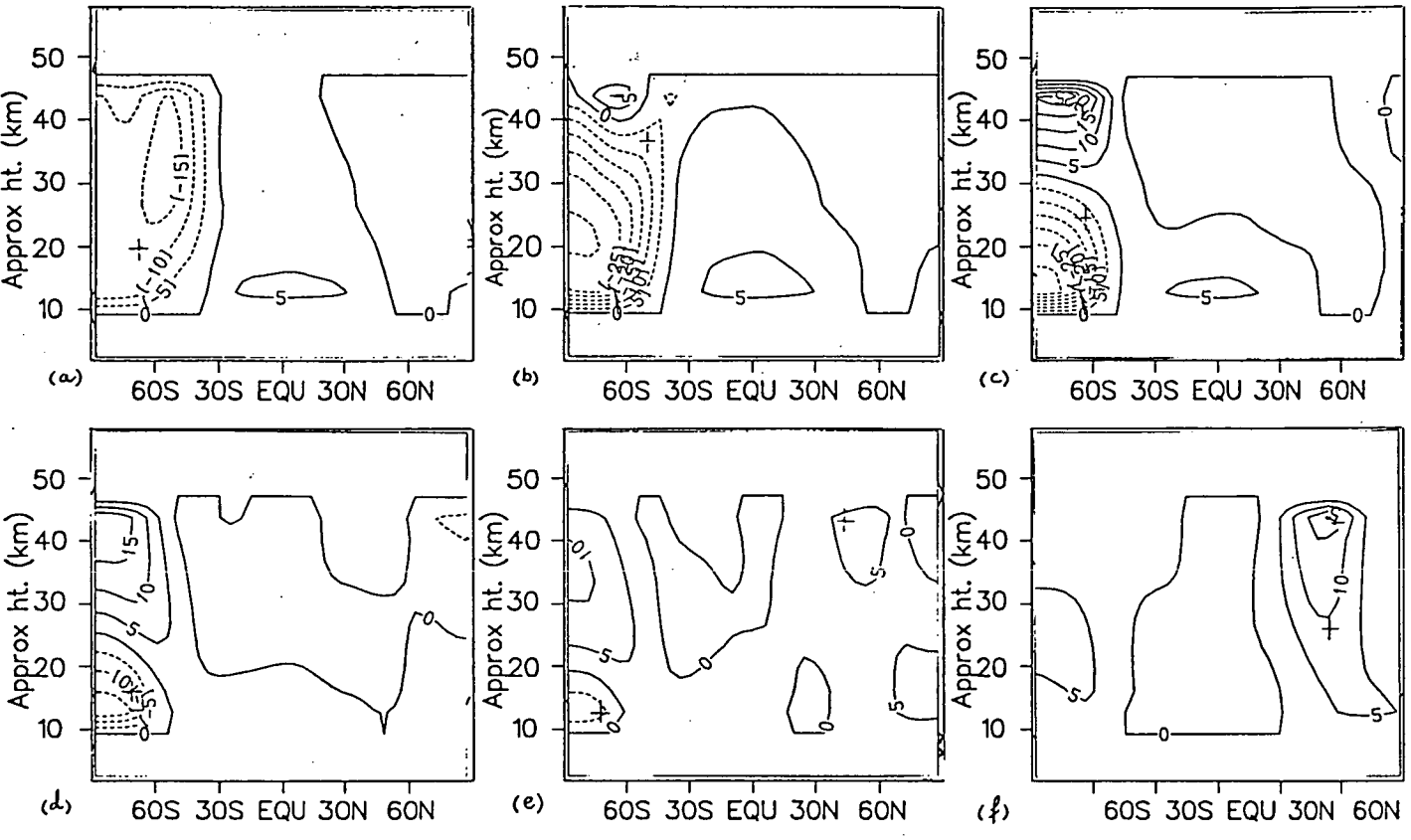


Figure 8.13: *IHD in observed* $\overline{T}$ *for (a) Jul/Jan (b) Aug/Feb (c) Sep/Mar (d) Oct/Apr (e) Nov/May (f) Dec/Jun. Dashed contours indicate negative values.*

a systematic negative error in the PV flux due to a correlation between errors in the displacement and in PV.

The PV flux due to dissipation estimated in this chapter is large in the NH from October till March, with a peak value of -12ms^{-1}day^{-1} in January.

In the SH, the PVF is largest in September, October and April (being between -6 and -8ms^{-1}day^{-1}). However, large values occur throughout the year but are closer to the equator and so are expected to be largely due to the over-estimation just mentioned.

Agreement with chapter 4 is very good, indicating that the IHD in the dissipative flux is accurately estimated. The results of a model run also suggest that the IHD is largely correct, but that the total PVF is too large. The PVF of this chapter is also larger than that of chapter 7 (by about -4ms^{-1}day^{-1} in the NH in January and the SH in September). The neglect of local geostrophically-induced

circulations mentioned above, and errors in the estimation of the meridional parcel displacement both lead to an overestimation of the PV flux.

In conclusion, dissipation appears to account for most of the stratospheric $\mathcal{F}$, but the method used here in estimating the dissipative PV flux results in a systematic over-estimation which has about the same size in both hemispheres at the same seasonal time.

Chapter 9

Summary and Conclusions

So then I could tell them
Where the wind goes ...
But where the wind comes from
Nobody knows

'Now we are six'. A.A. Milne

In this chapter the main results from the other chapters will be summarised, some conclusions will be drawn from them, and directions in which the research could be extended will be outlined.

9.1 The isentropic model

The work described in this thesis has centred on the construction of a 2 dimensional isentropic model of the atmosphere below about 100 km. Its use in this study is confined to investigating the dynamical effect of stratospheric planetary waves, though it is intended to provide the dynamical framework for a model which will predict the climate's evolution in the face of chemical perturbations. The use of isentropic coordinates leads to conceptual advantages, and will lead to practical advantages in parametrising the horizontal eddy fluxes of chemicals in this future version of the model. Its main features are described in the following paragraphs.

The prognostic equations for $\overline{T}$ and $\overline{u}$, constrained by thermal wind balance, are solved, their evolution being forced by an interactively-calculated diabatic heating

and a PV (Ertel's potential vorticity) flux specified from observations (see next section). The only other model which solves these equations in isentropic coordinates is that of Tung and Yang (1988), but in their model $\overline{T}$ and $\overline{u}$ are specified and used to deduce the PV flux. Our model is therefore more suitable for predicting the future climate in the presence of changes in the diabatic heating, if the PV flux is assumed not to change significantly or, preferably, if a suitable parametrisation of the PV flux in terms of $\overline{u}$ can be developed. The coordinate above the 350K isentropic level (corresponding roughly to the tropopause) is $\ln(\theta)$ while below, sigma-type coordinates are used to avoid the intersection of coordinate levels with the ground. The vertical resolution is variable.

A fairly accurate radiative heating scheme is used above about 25km (see Haigh and Pyle (1982)), but is modified by including the effect of the variation in the hours of daylight with height. This was found to have a significant (over 5K) effect on the temperature, especially at the equinoxes.

A simple, interactive, parametrisation of the important tropospheric processes was included. This enables an interaction between the stratosphere and troposphere. As a result, an annual cycle of the temperature of the equatorial tropopause, similar in amplitude and phase to that observed, was produced in the model on the inclusion of a stratospheric planetary-wave PV flux.

The model simulates fairly well the real stratosphere in the autumn, and in the SH in winter, without any parametrisation of a planetary wave PV flux.

In summer, the easterlies do not extend as low as observed in the real atmosphere, and the lower stratosphere is too cold, indicating either a lack of heating in the lower stratosphere or the lack of a negative PV flux there, or both.

In winter there appears to be a lack of a negative PV flux at about 60km, due to the over-simple parametrisation of gravity waves in terms of Rayleigh friction. The SH westerly jet accelerates from June to July, while the observed jet decelerates, due probably to the action of a negative PV flux in June (see below). The gravity wave drag should therefore perhaps be 'tuned' to produce a realistic westerly jet in the SH in June rather than July so that a realistic parametrisation of the PV flux in June will lead to realistic modelled June and July winds. A more sophisticated parametrisation of gravity wave drag (eg. Holton 1983) might improve the performance, although this would require a 'tuning' of gravity wave amplitudes and phase speeds.

The tropospheric winds and temperatures are simulated quite well. The strength of the Hadley cells can be easily altered, with only a small effect on the winds and temperatures, (in both the troposphere and stratosphere) by suitably adjusting both the tropospheric cooling and the tropospheric friction. This is likely to prove useful when chemicals are included in the model to which the exchange of air between the troposphere and stratosphere is important.

The model has the potential to be more interactive, if the scheme of Hitchman and Brasseur (1988), adapted to isentropic coordinates, is used to parametrise $\mathcal{F}$ using the modelled $\overline{u}$ field (although the wave activity would still have to be specified at the tropopause). The inclusion of chemicals as independent variables would then introduce the interaction between chemicals and radiative heating rate, as well as the interactions amongst the chemicals., The method of Smith and Brasseur (1990) could then be used to deduce species-dependent diffusion coefficients from the model-derived $\mathcal{F}$. Here the advantage of isentropic coordinates in needing only one diffusion coefficient (for adiabatic planetary waves) would come into play. The inclusion of more interactions will make the model more sensitive to errors in each parametrisation, but will also make it more useful in predicting the future climatological state of the atmosphere.

9.2 Estimate of the planetary wave PV flux

If this model is to be capable of simulating the real atmosphere it must be fed realistic values for $\mathcal{F}$, the wave-forcing of $\overline{u}$, which is well-approximated by the horizontal eddy flux of PV. These values might come either from an internal model parametrisation or from a calculation using observational data. An accurate estimate of the real stratospheric $\mathcal{F}$ is, however, not available, and hence there is no way of verifying a $\mathcal{F}$ produced by an internal model parametrisation. Therefore, the auxilliary purpose of this thesis was to derive such an estimate for the year from July 1980 to June 1981. Three different methods, to be described now, were used.

9.2.1 The IHD in the diabatic circulation - Chapter 4

This method is a variation of a previously used method (Shine 1989, Yang et al. 1990), in which the acceleration of $\overline{u}$ due to the estimated diabatic circulation is

subtracted from the observed acceleration, the residual being $\mathcal{F}$. This method suffers from being very sensitive in the stratosphere to the diabatic circulation, to the extent that even the sign of the deduced $\mathcal{F}$ is uncertain (Shine 1989). The method used here estimates only the inter-hemispheric difference in the diabatic circulation between the NH and the SH at the same seasonal time (the inter-hemispheric difference, or IHD) and hence arrives at a value for the IHD in $\mathcal{F}$.

The main reason for supposing that the IHD in the diabatic circulation could be more accurately estimated than the full circulation lies in the fact that a systematic error in the full diabatic circulation which occurs to the same extent in both hemispheres would be removed on taking the inter-hemispheric difference. Since the IHD in $\overline{\mathrm{T}}$ is small in summer, and is fairly well anti-correlated with the IHD in $\overline{\mathrm{T}}_t$ in late spring (when $\mathcal{F}$ is expected to be small) the IHD in the diabatic heating was approximated by a relaxation of the IHD in $\overline{\mathrm{T}}$ towards zero. Although simple, this method appeared to give a reliable estimate of the IHD in $\mathcal{F}$ and was robust in the sense that the overall shape did not depend on the size of the relaxation time constant used. The IHD in $\mathcal{F}$ contains less information than just $\mathcal{F}$, but by assuming that $\mathcal{F}$ is always negative in the stratosphere (see below) and that $\mathcal{F}$ is large in the NH in winter and in the SH in spring, we were able to estimate absolute values for $\mathcal{F}$.

In the NH, $\mathcal{F}$ was significant from November till March, reaching a peak of about -6 ms^{-1}day^{-1} in Jan/Feb, while in the SH it was significant from August till October, being at least -3 ms^{-1}day^{-1} in Aug/Sep. The use of shorter, and perhaps more realistic, relaxation time constants led to a 50% increase in these values. In both hemispheres the peak in $\mathcal{F}$ moves polewards over its life-span.

9.2.2 Direct calculation, approximate winds - Chapter 7

The second method calculates the horizontal eddy flux of PV directly from approximate horizontal winds (which were geostrophic with added higher-order corrections) obtained from satellite geopotential height data. The eddy flux of PV is assumed to be a good approximation to $\mathcal{F}$. Positive regions were found, similar to the positive regions found by previous authors in their calculations of the eddy flux of quasi-geostrophic potential vorticity, and the Eliassen-Palm flux divergence, and known to be due at least in part to inaccurate estimates of the horizontal winds. The spurious nature of these positive PV fluxes was further supported by the ex-

pectation that breaking or thermally-dissipating planetary waves will produce a negative PV flux, while the monthly-mean PV flux due to reversible vacillations of the vortex will be small. Moreover, these positive regions, when applied to the model, produced a $\overline{u}$ field with an unrealistically strong horizontal shear. These positive PV fluxes were therefore assumed to be due to a systematic positive error, which was then seen to be correlated with the observed regions of large westerly $\overline{u}$. The real PV flux was then assumed to be negative in the monthly average.

Setting the positive fluxes to zero and subsequent smoothing of the PV flux left large negative values in the NH in January and February (reaching -8 ms^{-1}day^{-1}) and in the SH from September till November (reaching -4 ms^{-1}day^{-1}). Compared with the results of the first method, there is a lack of PV flux in the NH in November, December and March, and in the SH in August, but an excess in the SH in November, although the peak values are similar. These processed fluxes were not considered to be accurate estimates of the real PV flux, but it was decided to investigate their effect on the model, and hence gain some idea of their similarity to the real PV fluxes. Comparison of the model fields (on using these fluxes) with observed $\overline{u}$ and $\overline{T}$ fields, indicated a lack of negative $\mathcal{F}$ in the NH in November and December, and in the SH from June until August, but that the values for the other months were quite realistic. The model also revealed that stratospheric planetary waves probably play a large part in the annual cycle of the temperature of the equatorial tropopause and stratopause, and that a large negative PV flux in one hemisphere can lead to a significant cross-equatorial flow in the upper stratosphere and a deceleration of $\overline{u}$ in the opposite hemisphere.

9.2.3 Dissipation of PV anomalies - Chapter 8

The third method estimates the PV flux due to dissipation of PV anomalies. Dissipation was expected to be more important than parcel dispersion in producing the PV flux in the mid- and high-latitude stratosphere for the following reason. The diffusion coefficient required for estimating the horizontal eddy flux of a tracer (which is due only to parcel dispersion) in the stratosphere is small in comparison with that needed for PV, therefore dispersion does not appear to be a major cause of the PV flux. Also, parcel dispersion by breaking planetary waves appears to occur mainly near the $\overline{u}=0$ line, which generally lies at low latitudes.

In order to estimate the PV flux due to dissipation of zonal asymmetries of PV,

a dissipative time constant is required. The one used here was that for thermal dissipation of temperature anomalies. The assumption that the two time constants are approximately equal was justified in the following way. It was demonstrated that zonal asymmetries in density (which, in isentropic coordinates, is directly related to temperature) are well anti-correlated with asymmetries in PV where planetary waves are large and it was suggested that this anti-correlation arises through the local maintenance of thermal wind balance. It was assumed that this anti-correlation is maintained as the temperature anomalies are dissipated and hence that the dissipation of temperature anomalies leads to a dissipation of PV anomalies with the same time constant. However, it was shown that as a PV anomaly is being thermally dissipated, a local circulation will arise in order to maintain thermal wind balance. This circulation will slow down the dissipation, effectively lengthening the dissipative time constant, the lengthening being greatest at low latitudes and for tall, thin PV anomalies. This effect was neglected in our estimation of the PV flux, which is therefore expected to be overestimated, by a factor of at least 2 at low latitudes. In order to estimate the dissipative PV flux, values are also needed for the parcel displacement. These were defined using the symmetric vortex state, described more fully below, as the 'undisplaced' reference state. At times when the meridional gradient of PV in the symmetric vortex is small, small errors in the value of PV can result in large errors in the displacement, will add a negative contribution to the estimated PV flux.

The PV fluxes found using this method are always negative and generally larger than those found using the other two methods. In the NH they exceed $-3ms^{-1}day^{-1}$ from October till March, with a peak of $-12ms^{-1}day^{-1}$ in January. From May till September, values north of 50N are less than $1ms^{-1}day^{-1}$ in magnitude, though values equatorward of 50N can exceed $-2ms^{-1}day^{-1}$. From the above discussion, these lower-latitude values are probably seriously over-estimated. In the SH the peak value for each month never falls below $2ms^{-1}day^{-1}$ in magnitude, indicating that the SH has a higher 'background' asymmetry than the NH. The largest PV fluxes, of between -6 and $-8ms^{-1}day^{-1}$, occur in September, October and April. In the other months, the peak values are smaller and occur at about 40S, so are expected to be significantly over-estimated. When these fluxes are applied to the model the westerly jet is virtually destroyed at mid-latitudes, except in the SH winter. The PV flux therefore appears to be much too strong, and can be taken as the upper limit on the real PV flux throughout the year. However, the IHD in the modelled $\overline{T}$ is very similar to observations, and the IHD in the PVF is also very similar to the that found in the first method. There therefore seems to be a

systematic negative error in the PV flux, and this error has a small IHD. This overestimation will probably depend on the time of year since, taking $-8ms^{-1}day^{-1}$ as a good estimate for the peak value in the NH winter (since it produces reasonable model results), the error in this method is $-4ms^{-1}day^{-1}$, while the error in the NH summer can be at most $-2ms^{-1}day^{-1}$ since that is the peak value of the summer PV flux.

9.2.4 Deduction of main features of real PV flux

Using the above three estimates, and bearing in mind the limitations of each, the following statements are made with reasonable confidence about the monthly-mean PV flux occuring in the real atmosphere, for the year from July 1980 to June 1981.

(1) When the PV flux is calculated using winds derived from geopotential height data (Chapter 7), positive fluxes in the monthly mean are due to a systematic error, which is correlated with large westerly $\overline{u}$.

(2) The fluxes, when large, take the form of a single negative peak at mid- to high-latitudes, and at about 40-50km.

(3) The flux is small (less than $1ms^{-1}day^{-1}$ in magnitude) in the NH from April till September, and in the SH from December till March. This is deduced from the results of Chapter 7 (which are taken to be fairly accurate when $\overline{u}$ is small - see point (1)) rather than those of Chapter 8 (the dissipative flux), which have a large negative 'background' value due probably to errors in the displacement. A small value in the SH in July is also expected from the small amplitude of the stationary waves then.

(4) The large values in the NH therefore occur from October till March, with the largest values of about $-8ms^{-1}day^{-1}$ (which gives good model results) occuring in January and February. In Chapter 8, large values occur during these months, but in Chapter 7 the values from October to December and in March are obscured by the systematic error.

(5) The large values occur in the SH in April (Chapter 8), June (model results from Chapter 7) and from August till November. The peak value, of about $-4ms^{-1}day^{-1}$, is expected to occur in September/October, with the April value being only slightly smaller (by about $1ms^{-1}day^{-1}$ - Chapter 8).

(6) The peak values generally move polewards from about 35 to 65 degrees (Chapters 4 and 8) from winter to summer.

(7) There are probably values of about -2ms^{-1}day^{-1} in the lower stratosphere in summer. These are not obvious from our calculations but are apparent in the GFDL 'SKYHI' 3D model (Mahlman and Umscheid 1987) and would help to produce, in our model, the observed summer easterlies in the lower stratosphere.

The above points are in general agreement with the few previous published estimates of ∇.F (excluding those calculated using geostrophic winds, which are taken to be highly inaccurate), which should be similar to $\mathcal{F}$ if the residual circulation and the meridional circulation in isentropic coordinates are similar. Using satellite data, Randel (1987) deduced a peak value of about -8ms^{-1}day^{-1} at 75N and 40km in January from a direct calculation using approximate winds of a higher order than geostrophic. However, at lower latitudes in the NH in January and in the SH in August, he also found a peak of about -5ms^{-1}day^{-1}, as well as a small positive region. On the other hand, Mahlman and Umscheid (1987), using the high resolution 3D 'SKYHI' model, found a single negative peak of over -4ms^{-1}day^{-1} in the high latitude stratosphere, and no positive regions, in the NH average for January.

Yang et al. (1990) published an estimate of $\mathcal{F}$ for the four mid-seasonal months, and found values of about -4ms^{-1}day^{-1} in the NH in January and -3ms^{-1}day^{-1} in the SH in October, at mid-latitudes. However, these values will suffer from the same uncertainties as those of Shine (1989) - see section 9.2.1.

9.2.5 Ways of improving the estimate of PV flux

The above work suggests several ways of obtaining a better estimate of the real stratospheric PV flux.

Dissipation of PV anomalies appears to be the dominant process in producing a PV flux in the mid- and high-latitude stratosphere. Estimation of the dissipative PV flux might therefore provide a more accurate estimate of the total PV flux than a direct calculation using approximate winds, which is very sensitive to the geopotential height data used and to the method of approximating the winds. The fact that the IHD in the PV flux estimated in this way is very similar to that

estimated in Chapter 4 is also encouraging. This method does, however, require an accurate estimate of the dissipative time-scale for PV anomalies, which depends on the shape of the anomalies and their latitude, and on suitable values for parcel displacements.

If the systematic errors in the dissipative PV flux, and the calculated PV flux (Chapter 7) could be quantified they could then be subtracted out. It might be possible to do this empirically, by noting the connection between the positive fluxes and $\overline{u}$, for the method of Chapter 7. For the dissipative flux of Chapter 8 there is the relation between the over-estimation of the dissipative flux and the lengthening of the effective time-constant, as well as the large error in the displacements due to smaller errors in the value of PV when the meridional gradient of PV is small.

As noted in Chapter 4, rather than estimating the IHD in $\mathcal{F}$ by parametrising the IHD in the diabatic heating, a better way would be to estimate $\mathcal{F}$ using as accurate an estimate of the full diabatic circulation as possible, and then simply to take the inter-hemispheric difference. This should remove a large part of the effect of systematic errors in the heating rate.

Finally, since 3D models such as the 'SKYHI' model (Mahlman and Umscheid 1987) can simulate the real winds and temperatures quite well, it would be useful to have 'climatologies' of the $\mathcal{F}$ from such models.

9.3 The symmetric vortex

Although it seems to stand on the edge of the main theme of this thesis, the symmetric vortex state of Chapter 3 has several features useful to 2D modelling. The symmetric vortex was obtained from the observed atmospheric state (whose winds were calculated as described in Chapter 2) by distorting the contours of PV till they coincided with latitude circles, while ensuring that PV in the symmetric vortex decreased monotonically from the north pole to the south pole. The mass of each air parcel was conserved by the symmetrisation process, and the symmetric vortex was further constrained to be in thermal wind balance. With suitable boundary conditions, the above constraints were sufficient to uniquely determine the symmetric vortex. The following paragraphs will describe the main properties and features of the symmetric vortex.

The symmetric vortex state should be more useful than the zonally-averaged state for comparisons with 2D models, in that a model fed with the actual monthly-averaged PV flux should be expected to lie closer to the observed monthly-mean symmetric vortex state than the observed monthly-average zonal-mean state.

In the NH, over the year from July 1980 to June 1981, the symmetric vortex differs significantly from the zonal-mean state from late autumn to early spring. The biggest difference is in February 1981, when the peak in $\tilde{u}$ is over 35 m/s stronger than the peak in $\overline{u}$ and $\tilde{T}$ is 7K lower than $\overline{T}$ at the pole. In the summer, the difference is small (less than 5 m/s and 2K).

In the SH, the biggest differences between the symmetric vortex and zonal-mean states occur in early spring and summer, $(\tilde{u}-\overline{u})$ being greatest in December 1980, when it exceeds 15 m/s, and $|(\tilde{T}-\overline{T})|$ being greatest in October, when it is less than -4K at the pole.

The symmetric vortex was used as a reference state from which to define the displacements used in estimating the dissipative flux in Chapter 8. It could also be used to estimate K_{yy}, using definitions which involve displacements. These displacements would be more useful, however, if their sensitivity to small errors in PV, when meridional gradients of PV are small, could be reduced.

That the symmetric vortex is capable of revealing, more clearly than the zonally-averaged state, the effect of irreversible PV fluxes was shown in Chapter 4. It might therefore prove to be a useful tool in analysing the disturbed stratosphere, by removing the blurring effect of travelling Rossby waves and vortex vacillations, revealing the sharp meridional gradients of PV which are important to planetary wave propogation.

Appendix A

A.1 Table of Symbols

Below is a list of the symbols used in the thesis, with the page numbers on which they first appear, or where they are defined.

Symbol	Description	Page
χ	general quantity	8
u, v, w	eastward, northward and vertical wind components	8
S	source of χ	8
x, y, z, t	coordinates	8
$\overline{\chi}$	zonal-mean of χ	9
χ'	departure from zonal-mean	9
θ	potential temperature	27
ψ	(a) potential for defining residual circulation	13
	(b) stream function for isentropic model	78
$\overline{v}^*, \overline{w}^*$	residual circulation	13
$\overline{v}^T, \overline{w}^T$	transport circulation	17
$\overline{v}^L, \overline{w}^L$	Lagrangian-mean circulation	14
Q	(a) diabatic heating	14
	(b) internal energy	75
T	temperature	14
η, ξ	northward and vertical displacements	15
$\underline{\underline{\mathrm{K}}}$	transport matrix	15
K_{yy}	element of $\underline{\underline{\mathrm{K}}}$	15
Φ	geopotential	26

p	pressure	26
M	(a) Montgomery potential	26
	(b) mass function for symmetric vortex	46
z^*	(a) geopotential height	26
	(b) geometric height	88
p_0	reference pressure 1000mb	27
c_p	specific heat (constant p) 1005 J kg^{-1} K^{-1}	26
R	gas constant for dry air 287.05 J kg^{-1} K^{-1}	27
κ	R/c_p=2.88	27
ϕ	latitude	27
ρ	density in isentropic coords	27
Ω	angular velocity of earth	28
f	Coriolis parameter $2\Omega\sin\phi$	27
τ	angular momentum/unit mass	28
a	radius of earth 6371 km	28
Π	Ertel's potential vorticity	28
ζ	relative vorticity	28
$\mathcal{F}$	eddy forcing of $\overline{u}$ in isentropic coords ($ms^{-1}day^{-1}$)	30
$\hat{\chi}$	density-weighted zonal mean $\overline{\chi\rho}/\overline{\rho}$	30
χ^*	departure from $\hat{\chi}$ i.e. $\chi - \hat{\chi}$	30
$\overline{V}$, $\overline{W}$	mass weighted velocities used in model	30
∇.F	divergence of the Eliassen-Palm flux	22
$\tilde{u}$	zonal wind in symmetric vortex	45
$\tilde{T}$	temperature of symmetric vortex	45
F, G	(a) used in symmetric vortex equations	45,47
	(b) used in model equations	76,77
χ^D	the IHD of χ	57
H	diabatic heating rate	75
η	hybrid vertical coordinate used in model	76
A, B, C, D, E	used in model equations	77
k_n	height of nth model layer	80
c_n	cosine of latitude in model	81
N, M	used in model	81
Q, R, X, Y, Z	used in model	82
α	inverse of radiative time constant	61, 126, 140
δ, Δ	vertical and horizontal box heights	145

A.2 Note on approximate heights used in graphs

The approximate height scale used in the graphs of model fields in Chapters 5, 6, 7 and 8 is defined as

$$Z_P \equiv 7\ln(1000/p) \text{ km} \quad , \quad \text{where p is in mb.} \tag{A.1}$$

How well this approximates the actual height is shown in Tables A.1 and A.2, which contain Z_P and Z for January over the equator, and July at 80S, using pressure and temperature data from Barnett and Corney (1985a). Over the equator (Table

Z (km)	T (K)	Z_P (km)	Z_θ (km)
80	206.5	79.7	75.1
75	209.9	74.1	69.9
70	218.3	68.7	65.4
65	235.1	63.5	62.0
60	252.9	58.7	59.0
55	264.6	54.2	55.6
50	267.8	49.8	51.5
45	263.8	45.4	46.7
40	255.2	40.9	41.4
35	242.5	36.1	35.4
30	227.2	31.1	28.8
25	217.5	25.8	22.4
20	204.3	20.2	15.3

Table A.1: *Approximate heights (and temperature) for January over equator*

A.1), Z_P differs from Z by at most 1.5 km. However, near the south pole during winter (Table A.2), Z_P can differ from Z by up to 7 km. From the hydrostatic approximation, it can be shown that, in an isothermal atmosphere, $Z = Z_P T/239$, so that if T=239K, Z_P and Z are equal.

Also shown in tables A.1 and A.2 is the approximate height

$$Z_\theta \equiv \frac{7}{\kappa}\ln(\theta/250) \text{ km} \ , \tag{A.2}$$

which is used for the graphs in isentropic coordinates in Chapters 4 and 8. From the definition of θ it can be shown that

$$Z_\theta = \frac{7}{\kappa}\ln(T/250) + Z_P \ , \tag{A.3}$$

Z (km)	T (K)	Z_P (km)	Z_θ (km)
80	219.9	83.9	80.8
75	224.0	78.7	76.0
70	231.7	73.5	71.7
65	243.6	68.6	68.0
60	258.7	63.9	64.7
55	270.8	59.4	61.3
50	271.9	55.1	57.1
45	261.1	50.7	51.7
40	248.2	46.1	45.9
35	234.0	41.1	39.5
30	209.2	35.8	31.5
25	179.2	29.6	21.5
20	186.5	23.1	16.0

Table A.2: *Approximate heights (and temperature) for July at 80S*

so that, at any point, if T>250K then $Z_\theta > Z_P$, which is apparent in the tables. If T=200K then $(Z_\theta - Z_P) \simeq 5.4$ km, and if T=300K then $(Z_\theta - Z_P) \simeq -4.4$ km. Thus the difference between Z_θ and Z might be expected to be up to about 12K. However, from tables A.1 and A.2, it appears that Z_θ is, on the whole, no worse an estimate of Z than Z_P is, differing from Z by up to 1.5 km in table A.1 (except below 30km and above 60km) and by up to 7 km in table A.2.

It was remarked several times in the thesis that a box height $\kappa/2$ units, when using $\eta(\equiv \ln(\theta/250))$ as a vertical coordinate, corresponded to about 3.5 km. Differentiating eqn. A.3 w.r.t. Z and using the hydrostatic approximation gives

$$\frac{d\eta}{dZ} = \frac{d\ln T}{dZ} + \frac{\kappa}{7}\frac{239}{T} = \frac{\kappa}{7}\frac{239}{T}(1 + \frac{T_Z}{9.8}) \; . \qquad (A.4)$$

(As a check on eqn.A.4, a neutrally stable atmosphere ($d\eta/dZ=0$) will have the dry adiabatic lapse rate of −9.8 K/km). In an isothermal atmosphere with T=239K, $d\eta/dZ=\kappa/7$ and a box height of $\kappa/2$ units *does* correspond to 3.5 km. However, in the troposphere, where T∼270K and $T_Z \sim -7$ K/km, $d\eta/dZ \sim \kappa/28$ while in the stratosphere, T∼240K and $T_Z \sim 2$ K/km, so $d\eta/dZ \sim \kappa/6$. Therefore the approximation of $d\eta/dZ$ by $\kappa/7$ is useful as a rough guide in the stratosphere, but not in the troposphere.

Bibliography

Andrews D.G. (1983) A Finite-Amplitude Eliassen-Palm Theorem in Isentropic Coordinates. *Jour. Atm. Sci.* Vol 40 pp1877-1883

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

Andrews D.G. and McIntyre M.E. (1976) Planetary Waves in Horizontal and Vertical Shear: The Generalised Eliassen-Palm relation and the Mean Zonal Acceleration. *Jour. Atm. Sci.* Vol 32 pp893-898

Barnett J.J. (1974) The mean meridional temperature behaviour of the stratosphere from November 1970 to November 1971 derived from measurements by the selective chopper radiometer on Nimbus IV. *Quart. J. R. Met. Soc.* Vol 100 pp 505-530.

Barnett J.J. and Corney M.(1985a) Middle atmosphere reference model derived from satellite data *Handbook for MAP* Vol. 16 pp 47-85

Barnett J.J. and Corney M.(1985b) Planetary waves *Handbook for MAP* Vol. 16 pp 86-137.

Boville B.A. (1986) Wave-mean flow interaction in a general circulation model of the troposphere and stratosphere. *Jour. Atm. Sci.* Vol. 43 pp 1711-1723.

Brasseur G., Hitchman M.H., Walters S., Dymek M., Falise E., and Pirre M. (1990) An interactive chemical dynamical radiative two-dimensional model of the middle atmosphere *J. Geophys. Res.* Vol 95 pp 5639-5655.

Butchart N. and Remsberg E.E. (1986) The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface. *Jour. Atm. Sci.* Vol 43 pp1319-1339.

Callis L.B., Boughner R.E. and Lambeth J.D.(1987) The stratosphere: climatologies of the radiative heating and cooling rates and the diabatically diagnosed net circulation fields *J. Geophys. Res.* Vol 92D pp 5585-5607

Clough S.A., Grahame N.S. and O'Neill A. (1985) Potential vorticity in the stratosphere derived from data using satellites. *Quart. J. R. Met. Soc.* Vol 111 pp335-358

Crane A.J., Haigh J.D., Pyle J.A. and Rogers C.F. (1980) Mean Meridional Circulations of the Stratosphere and Mesosphere. *Pure and Applied Geoph.* Vol 118 pp307-328

Danielson E.F. (1959) The laminar structure of the atmosphere and its relationship to the concept of a tropopause. *Arch. Met. Geophys. Biokl.* A11 pp292-332

Dickinson R.E. (1973) Methods of parametrisation for infrared cooling between altitudes of 30 and 70km *J. Geophys. Res.* Vol 78 pp4451-4457

Dopplick T.G. (1979) Radiative Heating of the Global Atmosphere: Corrigendum *Jour. Atm. Sci.* Vol. 36 pp 1812-1817.

Dunkerton T. (1978) On the Mean Meridional Mass Motions of the Stratosphere and Mesosphere. *Jour. Atm. Sci.* Vol. 35 pp2325-2333.

Dunkerton T., Hsu C.P.F., and McIntyre M.E., (1981) Some Eulerian and Lagrangian diagnostics for a model stratospheric warming. *Jour. Atm. Sci.* Vol. 38 pp 819-843.

Dutton J.A. (1976) The Ceaseless Wind. *McGraw Hill*

Edmon H.J., Hoskins B.J. and McIntyre M.E. (1981) Eliassen-Palm cross sections for the troposphere. *Jour. Atm. Sci.* Vol. 37 pp 2600-2616.

Eliassen A. (1951) Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophys. Norv.* Vol. 5 No.2 pp19-60.

Fritts D.C. and Dunkerton T.J. (1985) Fluxes of heat and constituents due to convectively unstable gravity waves. *Jour. Atm. Sci.* Vol. 42 pp549-556.

Garcia R.R. and Solomon S.(1983) A numerical model of the zonally averaged dynamical and chemical structure of the middle atmosphere *J. Geophys. Res.* Vol 88C pp 1379-1400

Garcia R.R. and Solomon S.(1985) The effect of breaking gravity waves on the dynamics and chemical composition of the mesoephere and lower thermosphere. *J. Geophys. Res.* Vol 90 No. D2 pp 3850-3868.

Geisler J.E. and Dickinson R.E. (1976) The five-day wave on a sphere with realistic zonal winds. *Jour. Atm. Sci.* Vol. 33 pp632-641.

Geller M.A.,Wu M.F. and Gelman M.E. (1984) Troposphere-stratosphere (surface-55km) monthly general circulation statistics for the Northern hemisphere - interannual variations. *Jour. Atm. Sci.* Vol. 41 pp 1726-1744.

Gillie J.C. and Lyjak L.V.(1986) Radiative heating and cooling rates in the middle atmosphere.*Jour. Atm. Sci.* Vol. 43 pp 2215-2229

Gray L.J. and Pyle J.A. (1987) Two-dimensional model studies of equatorial dynamics and tracer distributions *Quart. J. R. Met. Soc.* Vol 113 pp 635-651.

Haigh J.D.(1984) Radiative heating in the lower stratosphere and the distribution of ozone in a two-dimensional model. *Quart. J. R. Met. Soc.* Vol 110 pp 167-185

Haigh J.D. and Pyle J.A.(1982) Ozone perturbation experiments in a two-dimensional circulation model *Quart. J. R. Met. Soc.* Vol 108 pp 551-574

Harwood R.S. and Pyle J.A. (1977) Studies of the Ozone Budget using a Zonal Mean Circulation Model and Linearised Photochemistry. *Quart. J. R. Met. Soc.* Vol 103 pp319-343

Harwood R.S. and Pyle J.A. (1980) The dynamical behaviour of a two-dimensional model of the stratosphere *Quart. J. R. Met. Soc.* Vol 106 pp395-420

Haynes P.H. and McIntyre M.E.(1987) On the evolution of Vorticity and Potential Vorticity in the presence of diabatic heating and frictional or other forces. *Jour. Atm. Sci.* Vol. 44 pp 828-840

Hirooka T. and Hirota I. (1985) Normal mode Rossby waves observed in the upper stratosphere. Part II: second antisymmetric and symmetric modes of zonal wavenumbers 1 and 2. *Jour. Atm. Sci.* Vol. 42 pp536-548.

Hitchman M.H. and Brasseur G. (1988) Rossby wave activity in a two-dimensional model: Closure for wave driving and meridional eddy diffusivity. *J. Geophys. Res.* Vol 93 No. D8 pp 9405-9417

Holton J.R.(1983) The influence of gravity wave breaking on the general circulation of the middle atmosphere*Jour. Atm. Sci.* Vol. 40 pp 2497-2507

Holton J.R. and Wehrbein W.M. (1980) The role of forced planetary waves in the annual cycle of the zonal mean circulation of the middle atmosphere. *Jour. Atm. Sci.* Vol. 37 pp1968-1983

Hoskins B.J., McIntyre M.E. and Robertson A.W. (1985) On the use and significance of isentropic potential vorticity maps. *Quart. J. R. Met. Soc.* Vol. 111 pp877-946.

Houghton J.T. (1986) The physics of atmospheres *Cambridge University press*

Hunt B.G., and Manabe S., (1968) Experiments with a Stratospheric General Circulation Model. II: Large-scale Diffusion of Tracers in the Stratosphere. *Mon. Wea. Rev.* Vol. 96 pp503-539

Jackman C.H., Newman P.A., Guthrie P.D. and Schoeberl M.R. (1988) Effect of computed horizontal diffusion coefficients on two-dimensional N_2O model disrtibutions *J. Geophys. Res.* Vol. 93 pp 5213-5219.

Jackson D.R., Harwood R.S. and Renshaw E. (1990) Tests of a scheme for regression retrieval and time-space interpolation of stratospheric temperature from satellite measurements. *Quart. J. R. Met. Soc.* Vol. 116 pp 1449-1470.

Kida H., (1983) General Correlation of air parcels and transport charecteristics derived from a hemispheric GCM: I. A determination of advective mass flow in the lower stratosphere. J. Met. Soc. Japan Vol 61 pp171-187.

Kiehl J.T. and Solomon S.(1986) On the radiative balance of the stratosphere*Jour. Atm. Sci.* Vol. 43 pp 1525-1534

Ko M.K.W., Tung K.K, Weinstein D.K., Sze N.D. (1985) A zonal-mean model of the stratospheric tracer transport in isentropic coordinates: Numerical simulations for nitrous oxide and nitric acid. *J. Geophys. Res.* Vol. 90 pp2313-2329

Kohno J. (1984) Stratospheric ozone transport due to transient large-amplitude planetary waves. J. Met. Soc. Japan Vol 62 pp413-439

Li D. (1991) Planetary wave activity in the stratosphere *Ph.D. thesis, Edinburgh University.*

Lindzen R.S. (1981) Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.* Vol. 86 No. C10 pp 9707-9714

Lorentz E.N. (1955) Available potential energy and the maintenance of the general circulation *Tellus* Vol. 7 pp157-167

Lyjak L.V. (1987) Diffusion coefficients calculated from satellite data. *Transport Processes in the Middle Atmosphere. D.Reidel. Ed. Visconti and Garcia* pp343-352

Mahlman J.D. (1969) Heat balance and mean meridional circulations during the sudden warming of January 1958. *Mon. Wea. Rev.* Vol 97 pp534-540

Mahlman J.D. and Umscheid L.J. (1987) Comprehensive modelling of the middle atmosphere: The influence of horizontal resolution. *Transport Processes in the Middle Atmosphere. D.Reidel. Ed. Visconti and Garcia* pp 251-266.

Marks C.J.(1989) Some Features of the Climatology of the Middle Atmosphere Revealed by Nimbus 5 and 6*Jour. Atm. Sci.* Vol. 46 pp2485-2508

Matsuno T., (1980) Lagrangian Motion of Air Parcels in the Presence of Planetary Waves. *Pure and Applied Geoph.,* Vol 118 pp189-216

McIntyre M.E., (1980) An Introduction to the Generalised Lagrangian-Mean Description of Wave, Mean-Flow Interaction. *Pure and Applied Geoph.,* Vol 118 pp153-176

McIntyre M.E., (1982) How Well do we Understand the Dynamics of Stratospheric Warmings? J. Met. Soc. Japan Vol 60 pp37-65

McIntyre M.E., (1987) Dynamics and tracer transport in the middle atmosphere: An overview of some recent developments. *Transport Processes in the Middle Atmosphere. D.Reidel. Ed. Visconti and Garcia* pp 267-296.

McIntyre M.E. and Norton W.A., (1990) Dissipative wave-mean interactions and the transport of vorticity or potential vorticity. *Jour. Fluid Mech.* Vol. 212 pp 403-435.

McIntyre M.E. and Palmer T.N. (1983) Breaking planetary waves in the stratosphere. *Nature* Vol 305 pp593-600.

McPeters R.D., Heath P.F. and Bhartia P.K. (1984) Average ozone profiles from the Nimbus 7 SBUV instrument.*J. Geophys. Res.* Vol 89 pp 5199-5214

Murgatroyd R.J. and Singleton J. (1961) Possible meridional circulations in the stratosphere and mesosphere. *Quart. J. R. Met. Soc.* Vol 87 pp125-135.

Newell R.E., Kidson J.W., Vincent D.G.and Boer G.J.(1972) The General Circulation of the Tropical Atmosphere

Newman P.A., Schoeberl M.R. and Plumb R.A. (1986) Horizontal mixing coefficients for two-dimensional models calculated from National Meteorological Society data. *J. Geophys. Res.* Vol 91 pp7919-7924.

Newman P.A., Schoeberl M.R. and Plumb R.A., Rosenfield J.E. (1988) Mixing rates calculated from potential vorticity. *J. Geophys. Res.* Vol 93

No. D5 pp 5221-5240

O'Neill A. and Pope V.D. (1987) The seasonal evolution of the stratosphere in the northern hemisphere. *Transport Processes in the Middle Atmosphere. D.Reidel. Ed. Visconti and Garcia* pp57-72

Palmen E. and Newton C.W., (1969) Atmospheric Circulation Systems. *Academic Press*

Palmen E. and Vourela L.A., (1963) On the Mean Meridional Circulations in the Northern Hemisphere during the Winter Season. *Quart. J. R. Met. Soc.* Vol 89 pp131-138

Pawson S. and Harwood R.S. (1989) Monthly-mean diabatic circulations in the stratosphere. *Quart. J. R. Met. Soc.* Vol 115 pp807-840.

Pawson S., Harwood R.S. and Haigh J.D. (1991) A study of the radiative dissipation of planetary waves using satellite data *To be published*

Pick D.R. and Brownscombe J.L., (1981) Early results based on the stratospheric channels of TOVS on the TIROS-N series of operational satellites. Int. Counc. Sci. Unions, Comm. Space Res., *Adv. Space Res.* Vol. 1 No. 4 pp 247-260.

Plumb, R.A., and Mahlman J.D. (1987) The zonally averaged transport characteristics of the GFDL general circulation/transport model. *Jour. Atm. Sci.* Vol. 44 pp298-327

Pyle J.A. (1980) A calculation of the possible depletion of ozone by chlorofluorocarbons using a tw0-dimensional model *Pure Appl. Geophys.* Vol. 118 pp 355-377

Pyle J.A. and Rogers C.F. (1980) Stratospheric transport by stationary planetary waves - the importance of chemical processes. *Quart. J. R. Met. Soc.* Vol 106 pp421-446

Randel W.J. (1987) On the evaluation of winds from geopotential height data in the stratosphere *Jour. Atm. Sci.* Vol. 44 pp3097-3120

Reed R.J. and German K.E. (1965) A contribution to the problem of stratospheric diffusion by large-scale mixing. *Mon. Wea. Rev.* Vol. 93 pp313-321.

Robinson W. A. (1986) The application of the quasi-geostrophic Eliassen-Palm flux to the analysis of stratospheric data. *Jour. Atm. Sci.* Vol. 43 pp 1017-1023.

Rodgers C.D. (1976) Evidence for the 5-day wave in the upper stratosphere. *Jour. Atm. Sci.* Vol. 33 pp710-711

Rogers C.F. and Pyle J.A. (1980) Stratospheric tracer transport: a modified diabatic circulation model. *Quart. J. R. Met. Soc.* Vol. 110 pp219-237.

Rosenfield E., Schoerberl M.R. and Geller M.A.(1987) A computation of the stratospheric residual circulation using an accurate radiative transfer model*Jour. Atm. Sci.* Vol. 44 pp 859-876

Salby M.L. and Garcia R.R. (1987) Vacillations induced by interference of stationary and travelling planetary waves. *Transport Processes in the Middle Atmosphere. D.Reidel. Ed. Visconti and Garcia* pp353-370

Schneider H.R., Ko M.K., Sze N.D., Shi G. and Wang W.(1989) An evaluation of the role of eddy diffusion in stratospheric interactive two-dimensional models*Jour. Atm. Sci.* Vol. 46 pp 2079-2093

Shine K.P. (1987) The middle atmosphere in the absence of dynamical heat fluxes. *Quart. J. R. Met. Soc.* Vol. 113 pp 603-633.

Shine K.P. (1989) Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation. *Quart. J. R. Met. Soc.* Vol. 115 pp265-292.

Shiotani M. and Hirota I. (1985) Planetary wave-mean flow interaction in the stratosphere: a comparison between northern and southern hemispheres. *Quart. J. R. Met. Soc.* Vol. 111 pp309-334

Smith A.K. and Brasseur G.P. (1990) The dependence of constituent transport on chemistry in a two-dimensional model of the middle atmosphere *J. Geophys. Res.* Vol 95 No. D9 pp 13749-13764.

Smith A.K., Lyjak L.V. and Gille J.C. (1988) The eddy transport of nonconserved trace scecies derived from satellite data. *J. Geophys. Res.* Vol 93 pp11,103-11,122

Tanaka H.(1986) A slowly Varying Model of the Lower Stratospheric Zonal Wind Minimum Induced by Mesoscale Mountain Wave Breakdown *Jour. Atm. Sci.* Vol. 43 pp1881-1892

Tung K.K. (1982) On the Two-Dimensional Transport of Stratospheric Trace Gases in Isentropic Coordinates. em Jour. Atm. Sci. Vol. 39 pp2330-2355.

Tung K.K. (1984) Modelling of Tracer Transport in the Middle Atmosphere. *Dynamics of the Middle Atmosphere. Ed J.R.Holton and T.Matsuno. D.Reidel Publishing Company.*

Tung K.K. (1987) A coupled model of zonally-averaged dynamics, radiation and chemistry. *Transport Processes in the Middle Atmosphere. D.Reidel.*

Ed. Visconti and Garcia pp183-198

Tung K.K. (1986) Nongeostrophic Theory of Zonally Averaged Circulation. Part 1: Formulation *Jour. Atm. Sci.* Vol. 43 pp2600-2618

Tung K.K. and Yang H. (1988) Dynamical Component of Seasonal and Year-to-Year Changes in Antarctic and Global Ozone. *J. Geophys. Res.* Vol 93 pp12,537-12,559

White A.A. (1979) A Note on the Energetics of Certain Zonal Average Models *Mon. Wea. Rev.* Vol. 107 pp347-8

World Meteorological Organization (1986) Global ozone research and monitoring project Report no. 16

Yang H., Tung K.K. and Olaguer E. (1990) Nongeostrophic theory of zonally averaged circulation. Part II: Eliassen-Palm flux divergence and isentropic mixing coefficient. *Jour. Atm. Sci.* Vol. 47 pp215-241.