# Planetary Waves and the Global Ozone Distribution

GREGORY ELTON BODEKER

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the
Space Physics Research Institute
of the
Department of Physics,
University of Natal

Durban
November 3, 1994

# Preface

The work described in this thesis was carried out in the Department of Physics, University of Natal, Durban under the supervision of Professor M.W.J. Scourfield and Dr. J.P.S. Rash and at the National Center for Atmospheric Research (NCAR), Boulder, Colorado, under the supervision of Dr. W.J. Randel, from April 1991 to February 1994.

These studies represent original work by the author and have not been submitted in any form to another university. Where use was made of the work of others, it has been duly acknowledged in the text.

# Acknowledgements

The author wishes to thank the following people and organisations:

My supervisor, Professor Malcolm W.J. Scourfield for introducing me to this field of study. His support and guidance proved invaluable.

Dr. William Randel for his help and guidance with the work performed at NCAR.

The Space Physics Research Institute for financial assistance for various conference attendances.

The Foundation for Research Development for financial support.

The University of Natal for financial assistance to attend the SASAS conference in 1993.

The National Center for Atmospheric Research for financial support during my working visit in January and February 1993.

The National Institute for Water and Atmospheric Research, Lauder, New Zealand, for the helpful comments and suggestions during my visit in June and July 1993.

Terry Holloway for computing assistance, for helping with the TEXing of this thesis and for printing out the final version of this thesis.

Kathy Mosher for computing assistance during my stay at NCAR and for continuing to help after my return to South Africa.

Michael Gallias and Jonathan Doull for their help with image processing and graphical output.

Judy Stephenson, Barbara Duigan and Dr. Erhard Mravlag for many stimulating discussions.

Dr. Anne Thompson for valuable suggestions and comments on many aspects of this work.

Karin Kreher and my mother for help with the proof reading.

Kathryn Volker for help with the proof reading and for a continuous supply of good CDs.

Technical staff of the Physics Department for assisting with hardware problems.

The South African Weather Bureau for providing the SANAE upper air meteorological data.

Dr. P. McPeters and Dr. A. Krueger (NASA/GSFC); members of the TOMS Nimbus Experiment and ozone processing teams; and the National Space Science Data Center through the World Data Center-A for rockets and satellites, for access to the TOMS ozone data.

# Abstract

Planetary waves are known to play an important role in determining the annual variability in the severity of Antarctic ozone depletion. In this thesis, data obtained from the TOMS experiment on board the Nimbus-7 satellite are used to create global maps of total column ozone for each day from 1979 to 1992. Planetary wave morphology within these distributions is examined using spatial spectral analysis. The wave powers obtained are compared with a measure of the depth and area of the Antarctic ozone hole.

A theoretical overview is given of middle atmosphere dynamics and how it influences the global distribution of ozone. Planetary waves play a predominant role in this dynamical distribution as they are responsible for the eddy transport of heat, momentum and long-lived tracers. A discussion of the equations of motion governing planetary wave propagation and their physical interpretation is given. The use of derived meteorological parameters, such as potential vorticity, in the study of planetary waves is developed. The theoretical tools, together with the equations of motion, are used to examine the excitation mechanisms for planetary waves. This theoretical analysis is also used to show that the strength of the westerly jet within the polar vortex determines which planetary wave modes can be expected in the geopotential height fields. In general only wavenumber 1 and 2 propagate during the Southern Hemisphere winter and the strength of these waves strongly modulates the severity of the Antarctic ozone depletion. Furthermore, the timing of the breakdown of the polar vortex is connected with the temporal variation of planetary wave power. A review of the current research in the field of planetary waves and the global distribution of ozone is given.

To quantify planetary wave activity, spatial spectral analysis (Fourier Transforms) of midlatitude zonal profiles of total column ozone are used to calculate wave powers in the ozone distribution. Powers at wavenumber 1 to 6 are calculated for each day from 1979 to 1992 for both the Northern and Southern Hemispheres. The powers are seen to decrease with wavenumber throughout the year and the amplitude of wave modes 3 and higher is suppressed during the winter in accordance with the theoretical discussion. Southern Hemisphere wave powers are compared with a daily depleted mass of ozone over the Antarctic from 1979 to 1992. It is found that during years of high (low) planetary wave activity Antarctic

ozone depletion is very weak (severe). Finally, the planetary wave morphology observed in the Southern Hemisphere total column ozone distribution is compared with geopotential height waves reported in the literature. Comparison of Northern Hemisphere wave powers and Arctic ozone depletion shows little or no correlation.

To examine whether planetary waves may effect the interannual variability in Antarctic vortex temperatures, winter averaged 100 hPa upper air temperatures from SANAE (70°18' S, 2°21' W) are compared with derived planetary wave powers. It is found that during years of high (low) planetary wave activity the SANAE 100 hPa temperatures are above (below) the mean. Analysis of daily total column ozone and upper air temperatures at SANAE, during the winters of 1987 and 1988, shows that wave 1 forcing can significantly influence the day-to-day variation in these quantities.

A statistical model of total ozone variation over the entire globe, from 1979 to 1992, has been developed. This model incorporates a long-term linear trend, an annual variation, a quasi-biennial oscillation, a solar cycle variation and a semi-annual variation. Since monthly average total column ozone data are used in this model, short term planetary wave influences are masked, and the variations in global ozone are determined primarily by the 5 signals discussed above. However, analysis of the model coefficients indicates that planetary wave activity may significantly contribute to zonal asymmetry in global total ozone trends. Differences between model results and measured data are compared with planetary wave activity. The timing of the breakdown of the Antarctic circumpolar vortex causes large interannual differences in monthly average total column ozone for the months of October and November in the Southern Hemisphere.

This analysis is made more specific for South Africa by examining the variation in monthly average total column ozone over the five South African cities of Pretoria, Bloemfontein, Durban, Port Elizabeth and Cape Town for the years 1979 to 1992. A model of surface erythemal irradiance as a function of total column ozone, time of the year and other meteorological parameters is developed. Total ozone data for the 5 cities, used within the context of the statistical model and the UV irradiance model, allows the prediction of future UV levels over South Africa to the year 2000.

Planetary wave motion owes its existence to the conservation of potential vor-

ticity. However, to make use of this conservation principle as a diagnostic for planetary wave propagation, it is necessary to know the conditions under which it is conserved. A formalism developed by Plumb and Ko (1992) has been used to suggest a technique whereby the 'lifetime' of potential vorticity may be determined. Use was made of data from the NCAR CCM2 model to test this hypothesis. The concentrations of long-lived tracers ($CH_4$ and $N_2O$) extracted from the CCM2 also show the effects of planetary wave breaking on the meridional distribution of these species. This study has important implications for airborne polar chemistry campaigns.

Papers have been submitted on certain portions of this work, viz:

1. Bodeker G.E., M.W.J. Scourfield, M.D. Barker, Total Column Ozone above South Africa: 1987 to 1990, *South African Journal of Science*, 88, 222-224, **1992**

2. Duigan B.L., G.E. Bodeker and M.W.J. Scourfield, The effect of changes in total column ozone on surface UV irradiance and erythemal burn times. *Proceedings of the National Association of Clean Air conference, Brits, South Africa*, **1993**.

3. Bodeker G.E. and M.W.J. Scourfield, Planetary waves in total ozone and their relation to Antarctic ozone depletion, *Geophysical Research Letters*, submitted 1994

4. Bodeker G.E. and R.L. McKenzie, Erythemal UV at 45°S: longitudinal and secular variability, *Weather and Climate*, 13, 17-21, **1994**

5. Bodeker G.E. and M.W.J. Scourfield, Trends in total column ozone over South Africa and its influence on surface UVB radiation, *Special SASAS '93 issue of South African Journal of Science*, submitted 1994.

6. Thompson A.M., R.D. Diab, G.E. Bodeker, M. Zunckel, G. Coetzee, C.B. Archer, D.P. McNamara, K.E. Pickering, J.B. Combrink, E. Mravlag, F. Sokolic, J. Fishman and D. Nganga, Total ozone over Southern Africa during SAFARI-92/TRACE-A, *Journal of Geophysical Research*, submitted 1994

7. G.E. Bodeker and M.W.J. Scourfield, Long-term Ozone and Temperature Correlations Above SANAE, Antarctica. *Proceedings of the Quadrennial*

*Ozone Symposium, Charlottesville, U.S.A.*, **1994**.

In addition, various parts of the work have been read at conferences viz:

1. Bodeker G.E. and M.W.J. Scourfield, Long-term Ozone and Temperature Correlations Above SANAE, Antarctica. *Poster presented at the Quadrennial Ozone Symposium, Charlottesville, U.S.A.*, **1992**

2. Scourfield M.W.J., G.E. Bodeker and M.D. Barker, South African Ozone Layer: 1979 to 1991. *Paper presented at the 37th Annual Conference of the South African Institute of Physics, Wits University*, Paper E8, **1992**

3. Bodeker G.E., M.W.J. Scourfield, Day-to-day ozone and temperature correlations above SANAE, Antarctica 1980 – 1990. *Paper presented at the 9th annual conference of the South African Society for Atmospheric Sciences, Pretoria*, **1992**

4. Bodeker G.E. and M.W.J. Scourfield, Long-term changes in total column ozone aver Okaukuejo: 1979 to 1992. *Poster presented at the TRACE-A/SAFARI data workshop, Stellenbosch, South Africa*, **1993**

5. Bodeker G.E., M.W.J. Scourfield, B.L. Duigan and E. Mravlag, Surface UV irradiance and erythemal burn times compared with total column ozone and cloud cover. *Poster presented at the TRACE-A/SAFARI data workshop, Stellenbosch, South Africa*, **1993**

6. Duigan B.L., G.E. Bodeker, M.W.J. Scourfield and B. Stefañski jr., Effects of variations in the vertical ozone profile on surface solar irradiance. *Poster presented at the TRACE-A/SAFARI data workshop, Stellenbosch, South Africa*, **1993**

7. Bodeker G.E., M.W.J. Scourfield and B. Stefañski jr., Trends in total column ozone over South Africa and its influence on surface UVB radiation. *Poster presented at the 10th annual conference of the South African Society for Atmospheric Sciences, MidRand*, **1993**

8. Bodeker G.E., W.J. Randel and M.W.J. Scourfield, Potential vorticity as a tracer for long-lived atmospheric constituents. *Paper presented at the 10th annual conference of the South African Society for Atmospheric Sciences, MidRand*, **1993**

# Contents

# List of Figures

# List of Tables

# Chapter 1

# INTRODUCTION

The annual formation of the Antarctic ozone hole is now a well documented phenomenon (Farman et al., 1985; Schoeberl et al., 1986; Krueger et al., 1988; Schoeberl et al., 1989; Stolarski et al., 1990; Newman et al., 1991; Krueger et al., 1992; Gleason et al., 1993). The persistent increase in the severity of the ozone depletion has spurred international collaboration to understand the chemistry and dynamics associated with the spatial formation and temporal evolution of the Antarctic ozone hole. A knowledge of the chemical reactions involved in stratospheric ozone depletion has been achieved through laboratory analysis, in situ observations of trace gases and the refinement of atmospheric models. Remote measurements of trace species are commonly made using differential absorption spectroscopy, both from the ground and from satellites. In situ measurements are generally performed using balloon-borne payloads or equipment based on aircraft. Interpretation of the data obtained has formed the basis of recent ozone research, providing new insights into the chemistry involved. An overview of the chemistry associated with polar ozone depletion is presented in section 1.1.

During the formative years of investigation into Antarctic ozone depletion, it was proposed that dynamical perturbations to the Antarctic stratosphere could be responsible for the observed ozone loss (Mahlman and Fels, 1986; Tung et al., 1986). Following the return of solar illumination to the polar stratosphere early in spring, solar heating in this unusually cold environment causes radiative warming. This warming would create large scale upwelling bringing ozone poor air from the troposphere into the stratosphere. However, the vertical structure of the depletion observed in ozone profiles obtained from ozonesonde launches performed in the

Antarctic (Solomon et al., 1986; Deshler et al., 1990), indicated that the ozone loss occurred over a limited altitude range. The 1989 ozone depletion shown in Figure 1.1 occurs almost entirely between 15 and 20 km suggestive of local, structured depletion.

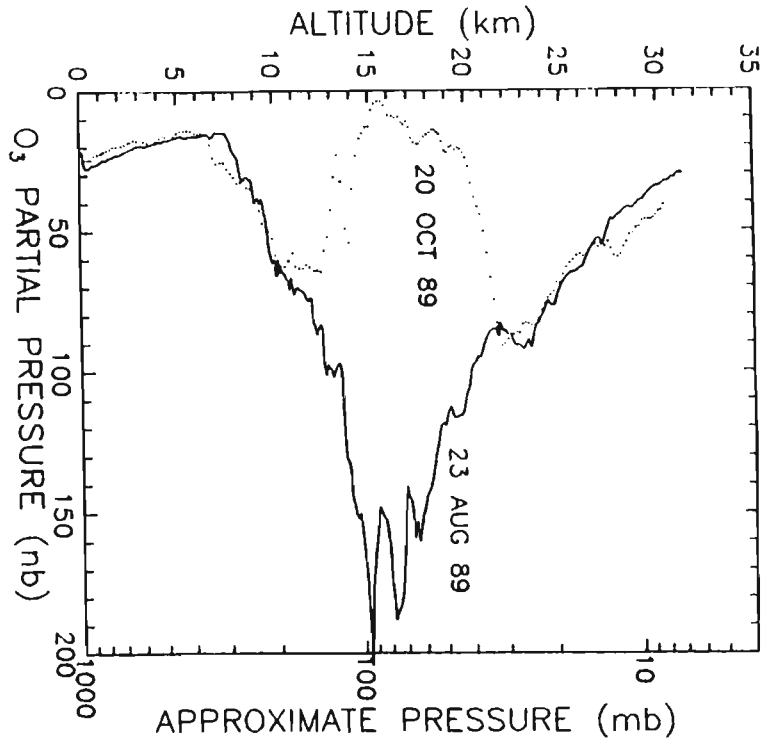


Figure 1.1: Comparison of ozone partial pressure at the height of the 1989 depletion on 20 October, with the initial sounding on 23 August 1989. From Deshler et al., 1990.

Together with observations of flow within the polar vortex (discussed below) it became apparent that large scale upwelling was not responsible for the observed depletion. Observations of background aerosols and other tracers of atmospheric dynamics, such as $N_2O$ (Podolske et al., 1989), indicated downward rather than upward motion.

In spite of the failure of the 'dynamical theory' to explain the springtime depletion of ozone over the Antarctic, it was realized that the unique meteorological conditions found within the Antarctic polar vortex were a strict prerequisite for the formation of the Antarctic ozone hole (Schoeberl and Hartmann, 1991). The isolation of the wintertime Antarctic stratosphere by the circumpolar vortex, a jet of strong westerly winds, results in extremely low upper air temperatures, per-

mitting the formation of polar stratospheric clouds (PSCs). These clouds play a crucial role in preconditioning the Antarctic stratosphere for ozone depletion (Toon and Turco, 1991).

This need to understand the meteorological attributes of the Antarctic ozone hole has created renewed interest in middle atmosphere dynamics. Large computer models incorporating atmospheric dynamics, stratospheric chemistry and radiative transfer, have been developed to aid this research effort. Satellite derived data sets are often used to initialize model runs and provide boundary constraints. The development of such models and the creation of large global meteorological data sets have provided powerful tools for the understanding of stratospheric chemistry.

The renewed interest in atmospheric dynamics and modelling has also spawned new diagnostic tools for the analysis of stratospheric chemistry. The calculation of back-trajectories makes it possible to obtain a detailed history of an air parcel on which measurements of trace gas concentrations have been made. Some data products derived from meteorological variables act as quasi-conserved atmospheric tracers under certain conditions eg. potential temperature and potential vorticity (see Chapter 2). The chemistry of the polar stratosphere is therefore closely linked with its dynamics. The application of these techniques to the study of the dynamical behaviour of the Antarctic ozone hole has shown that planetary wave activity can have a pronounced effect on the temporal evolution and severity of the ozone depletion. These waves, a natural oscillation in rotating fluids, disrupt the circumpolar vortex and significantly influence the distribution of ozone and other chemical species. An overview of the dynamics associated with ozone depletion is given in section 1.2, while section 1.3 is devoted to an introductory discussion of the importance of planetary waves in determining the global total column ozone distribution.

## 1.1 An overview of the chemistry of polar ozone depletion

The chlorofluorocarbon (CFC) theory has gained the most credibility in describing the chemical depletion of ozone over the Antarctic (Solomon et al., 1986; McElroy et al., 1986). Although the CFC theory concentrated initially on the

ozone depleting potential of chlorine, it now incorporates the effects of bromine containing halons, as discussed later. It had been suggested (Rowland and Molina, 1975) that widely used CFC compounds, although stable in the troposphere, would undergo photolysis in the UV environment of the stratosphere to release chlorine radicals, for example:

$$CF_2Cl_2 + h\nu \rightarrow Cl + CF_2Cl \tag{1.1}$$

These chlorine radicals would then participate in the following catalytic cycle that destroys ozone:

$$Cl + O_3 \rightarrow ClO + O_2 \tag{1.2}$$

$$ClO + O \rightarrow Cl + O_2 \tag{1.3}$$

However, these homogeneous gas phase reactions occur at heights between 35 and 45 km, while the maximum in ozone concentration lies between 15 and 25 km. Since photochemical models at that time included only homogeneous chemistry, and therefore only considered ozone depletion above 35 km, they indicated ozone depletions of 5% over 100 years and failed to predict the occurrence of the Antarctic ozone hole (Wuebbles et al., 1983).

Following the development of a theory incorporating heterogeneous chlorine chemistry taking place on the surface of clouds in the Antarctic stratosphere, a more accurate assessment of ozone depletion was provided. It was already known that the extremely low winter temperatures of the Antarctic stratosphere could lead to the formation of PSCs in greater quantities than anywhere else on Earth. Following the discovery of the Antarctic ozone hole (Farman et al., 1985) it was suggested (Solomon et al., 1986) that these clouds could participate in ozone destruction reactions through heterogeneous processes such as:

$$HCl(s) + ClONO_2(g) \rightarrow Cl_2(g) + HNO_3(s) \tag{1.4}$$

where (s) refers to solid and (g) refers to gaseous species. This reaction converts chlorine from the inactive compounds HCl and $ClONO_2$ into the more reactive species $Cl_2$. Following the end of the polar night, photodissociation of $Cl_2$ leads to the release of free chlorine radicals (Cl) into the stratosphere. This atomic chlorine is then involved in a catalytic cycle that destroys ozone, as discussed below.

In comparison to the altitude range of homogeneous gas phase ozone destruction (35 to 45 km), PSCs form between 10 and 22 km (Hofmann and Deshler, 1991).

The heterogeneous chemistry associated with these clouds releases free chlorine at that altitude range where ozone destruction is most effective since it incorporates the maximum in the vertical ozone profile. Other heterogeneous reactions such as

$$ClONO_2(g) + H_2O(s) \rightarrow HNO_3(s) + HOCl(g) \quad (1.5)$$

$$N_2O_5(g) + HCl(s) \rightarrow HNO_3(s) + ClNO_2(g) \quad (1.6)$$

$$N_2O_5(g) + H_2O(s) \rightarrow 2HNO_3(s) \quad (1.7)$$

associated with the formation of PSCs further convert Cl from the stable reservoirs of $ClONO_2$ and HCl into the more reactive forms, HOCl and $ClNO_2$.

There are two types of PSCs (Toon and Turco, 1991):

- Type I clouds consist of nitric acid trihydrate ($HNO_3 \cdot 3H_2O$), also known as NAT, forming at a temperature of 195 K (-78 °C).
- Type II clouds consist of pure water ice forming at 190 K (-83 °C).

The fact that type I clouds form at higher temperatures than the pure water ice clouds, results in a greater frequency of occurrence and greater spatial distribution of these clouds than initially suspected. Not only do PSCs release free chlorine into the stratosphere but they also convert reactive nitrogen (in the form of the $NO_2$ contained in the $ClONO_2$) to $HNO_3$. All four PSC reactions listed above (1.4 to 1.7) result in the formation of a solid, stable $HNO_3$ particulate which may precipitate out of the stratosphere. The resultant removal of reactive nitrogen through this denitrification has important implications for chlorine chemistry as a result of the reaction:

$$ClO + NO_2 + M \rightarrow ClONO_2 + M \quad (1.8)$$

where M is any third body. This reaction is the primary process for the conversion of active chlorine into its inactive form of $ClONO_2$. Furthermore, it is the availability of $NO_2$ which determines the rate of this reaction with the result that the $NO_2$ concentration limits the time for which chlorine is available to destroy ozone. The $NO_2$ distribution thus plays a central role in determining the extent and duration of the ozone depletion. Sources of $NO_x$ (N, NO, $NO_2$, $NO_3$, $N_2O_5$, $HO_2NO_2$) in the stratosphere include emissions from the Earth's surface by bacteria in the soil and by overuse of nitrogen fertilizers. Other contributions are

major ocean upwellings and the burning of fossil fuels. $NO_x$ may be removed from the stratosphere via PSC sedimentation as described above or it may be washed out of the atmosphere in the form of $HNO_3$ (Brasseur and Solomon, 1986).

Following the conversion of inert chlorine reservoirs to active chlorine species through chemical pathways involving PSCs, photolysis leads to the release of free chlorine atoms into the stratosphere. The following chlorine catalytic cycles are then believed to be responsible for ozone destruction (Solomon, 1990):

$$2 \times (Cl + O_3 \rightarrow ClO + O_2) \quad (1.9)$$

$$ClO + ClO + M \rightarrow Cl_2O_2 + M \quad (1.10)$$

$$Cl_2O_2 + h\nu \rightarrow Cl + ClO_2 \quad (1.11)$$

$$ClO_2 + M \rightarrow Cl + O_2 + M \quad (1.12)$$

$$\text{Net} : 2O_3 + h\nu \rightarrow 3O_2 \quad (1.13)$$

Note that in this reaction sequence chlorine acts catalytically, being consumed in reaction 1.9 and released again in reactions 1.11 and 1.12. A single chlorine radical may destroy up to 100 000 ozone molecules (Stolarski, 1988) in this way before being removed from the stratosphere through the $NO_2$ interference reaction (reaction 1.8). It may be expected that if more $NO_2$ is available, the number of ozone molecules destroyed per chlorine atom will be less than 100 000. The ozone creation and homogeneous destruction reactions, together with the $NO_2$ interference reactions, are summarized schematically in Figure 1.2. The complete heterogeneous ozone destruction sequences are omitted for clarity.

At the same time as the chlorine catalyzed heterogeneous reaction sequences were being investigated, it was also suggested (McElroy et al., 1986; Solomon, 1990) that reactions involving the coupling between chlorine and bromine could contribute to ozone depletion:

$$Cl + O_3 \rightarrow ClO + O_2 \quad (1.14)$$

$$Br + O_3 \rightarrow BrO + O_2 \quad (1.15)$$

$$ClO + BrO \rightarrow Br + ClO_2 \quad (1.16)$$

$$ClO + BrO \rightarrow BrCl + O_2 \quad (1.17)$$

$$ClO_2 + M \rightarrow Cl + O_2 + M \quad (1.18)$$

$$BrCl + h\nu \rightarrow Br + Cl \quad (1.19)$$

$$\text{Net} : 2O_3 \rightarrow 3O_2 \quad (1.20)$$

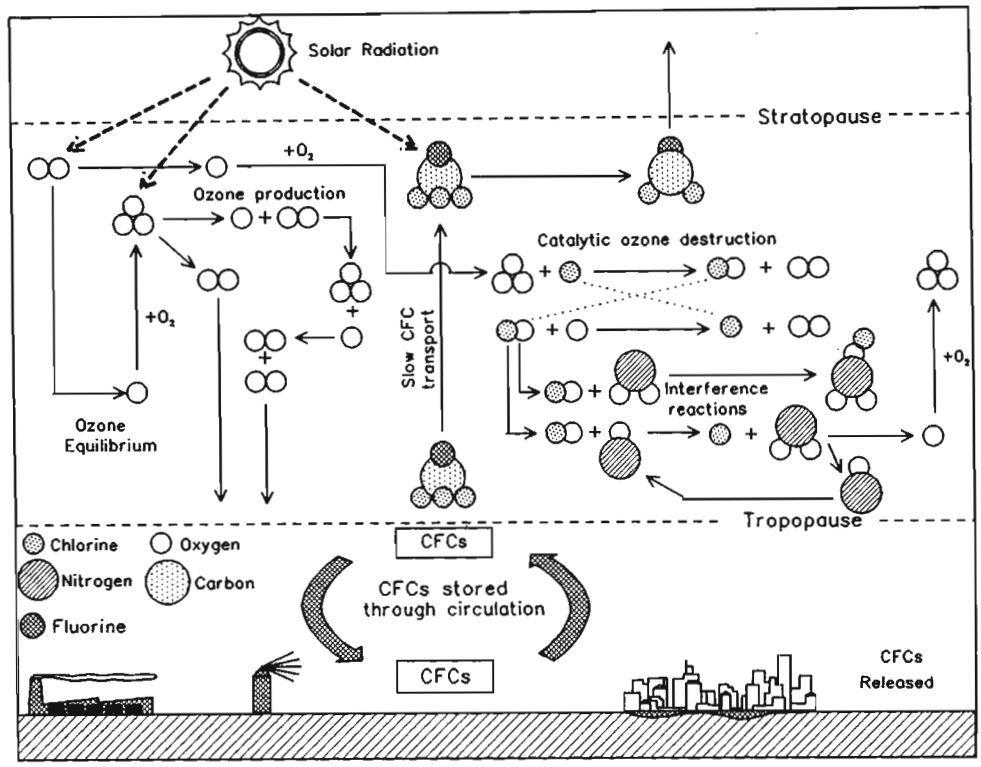


Figure 1.2: Schematic diagram of stratospheric ozone chemistry including ozone creation, homogeneous chlorine catalyzed ozone depletion and interference reactions.

Although the primary source of bromine is believed to be natural from surface releases of $CH_3Br$, anthropogenic emissions of bromine-containing halons are increasing (WMO report no. 20, 1990). Global concentrations of atmospheric bromine are about 10 p.p.t.v. (parts per trillion ($10^{12}$) by volume), although recent north polar measurements have reported values as high as 600 p.p.t.v. (Wahner and Schiller, 1992). Calculations have suggested that 20% to 30% of the observed decline in ozone is caused by the coupling of chlorine and bromine chemistry via the reactions discussed above (Solomon, 1990).

## 1.2 An overview of the dynamics of polar ozone depletion

It has been indicated above that the annual evolution of the Southern Hemisphere circulation plays a supportive role for chlorine catalyzed ozone depletion

over the Antarctic by creating an isolated vortex inside which temperatures fall low enough for the formation of PSCs. This is not the only way in which the Southern Hemisphere general circulation contributes to the distribution of ozone. The Antarctic circumpolar vortex consists of a stratospheric jet of winds flowing parallel to latitude circles in an eastward direction around the Antarctic. This mean circulation results from parcels of air moving poleward in response to a meridional pressure gradient force and then being deflected towards the east by the Coriolis force (see Chapter 2). Known as the geostrophic wind, this zonal flow (along latitude circles) tends to homogenize the ozone distribution in the zonal direction, but limits meridional transport of ozone from its source regions near the equator towards the pole. Any perturbations to the zonally symmetric flow are therefore expected to contribute to the meridional transport of ozone. The mechanisms facilitating meridional transport of ozone are very effective since although ozone production maximizes at low latitudes, the largest column abundances appear at high latitudes. This observation confirms that the global distribution of ozone cannot be accounted for in terms of photochemical processes alone.

The photochemical lifetime of ozone, which measures how quickly its concentration adjusts to changes in the photochemical environment, is determined by reactions involving its creation and destruction. Under certain conditions, ozone has a long photochemical replacement time and behaves like a quasi-conserved trace gas, being transported passively (advected) by the circulation of the atmosphere. In Figure 1.3 both the ozone photochemical replacement time and its meridional distribution are shown.

Assuming a vertical velocity of the order of $0.05\,\mathrm{cm\,s^{-1}}$, the advective transport time scale for the lower stratosphere is about 100 days for a constituent with a 5 km vertical scale height. Therefore, the distribution of a species whose photochemical lifetime exceeds 100 days, will be dominated by transport rather than photochemistry (Solomon et al., 1985). With a lifetime of hours in the upper equatorial stratosphere, ozone concentrations are expected to be close to the equilibrium state (Gray and Pyle, 1989). However, during the winter months, poleward of 50° and below 30 km, the photochemical replacement time for ozone exceeds 100 days and the ozone concentration can be far from its chemical equilibrium value since dynamics exerts a major influence. Most of the total column ozone at high latitudes is located in this dynamically controlled region, showing large local variability induced by planetary and cyclone waves (Dobson et al.,

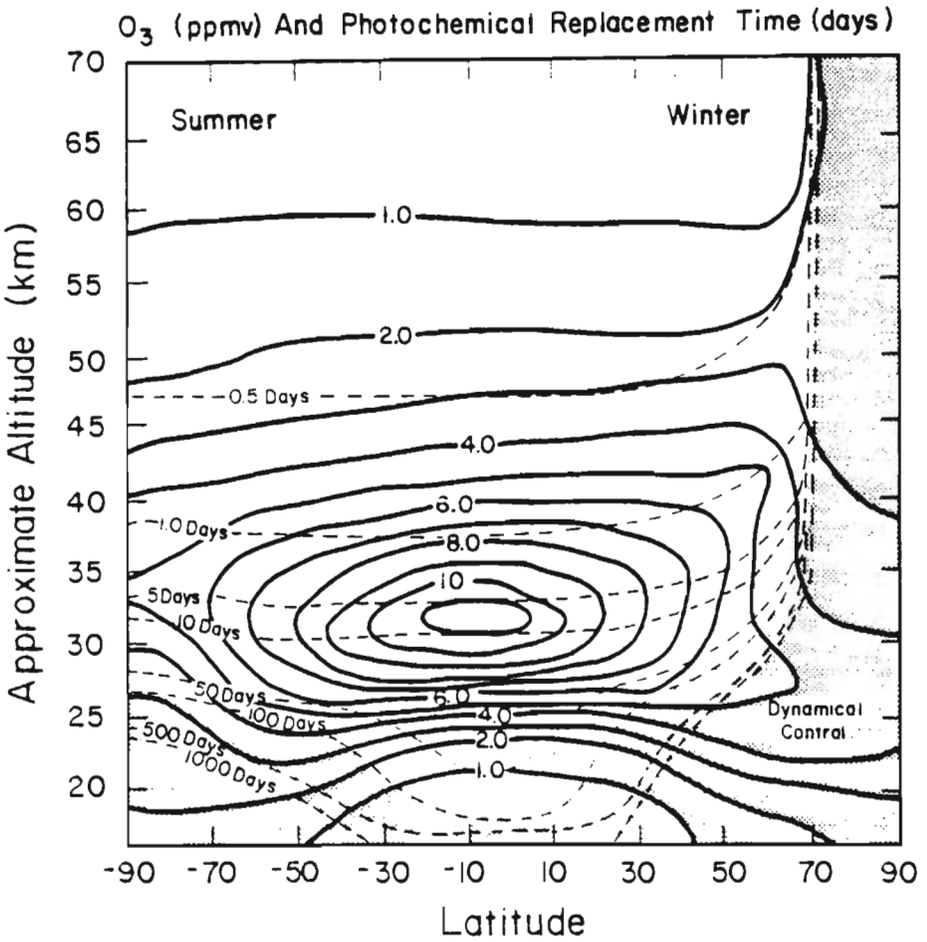


Figure 1.3: Computed photochemical lifetime of ozone in days (dashed) and ozone mixing ratio distribution in p.p.m.v. (solid) for the end of December. From the model of Garcia and Solomon (1983) and Solomon and Garcia (1984) as cited in Solomon et al., 1985.

1929; Schoeberl and Krueger, 1983). Even in the chemically controlled region, ozone concentrations are not independent of transport processes (Solomon et al., 1985). Although ozone is short lived above 25 to 30 km in the sunlit atmosphere at midlatitudes, and above 20 km in the tropics, its concentration depends on $NO_x$, $H_2O$ and other trace species that are long-lived enough to be strongly influenced by transport. During the winter months, as the Southern Hemisphere polar vortex forms with its associated perturbation to the general circulation cells, both the vortex isolation and dynamical transport play important roles in determining the ozone distribution.

Early in the year in the Southern Hemisphere, upper air temperatures above

the Antarctic are high, resulting in a small meridional pressure gradient and a weak westerly geostrophic wind. The Antarctic stratosphere is connected to the stratosphere at lower latitudes through adiabatic transport where air flows from the equator to the pole along isentropic surfaces (surfaces of constant potential temperature - see Chapter 2) descending in altitude to higher pressure (Salby and Garcia, 1990). The resultant compression of the air increases ozone densities, leading to significant increases in total column ozone. The net effect is that during this time of the year, ozone abundances above the Antarctic are higher than over the tropical regions. As the winter approaches, cooling above the Antarctic continent steepens the meridional temperature gradient creating an increased pressure gradient and a stronger geostrophic wind (Andrews et al., 1987). This process creates a polar vortex, or polar night jet, decoupling the Antarctic stratosphere from the rest of the Southern Hemisphere as discussed above. The radiative processes cooling the Antarctic stratosphere early in the winter result in symmetric overturning of the air, and together with non-local wave driving, creates downward motion in the polar regions and weak rising motion away from the Antarctic continent. The downward motion over the pole warms the air through adiabatic compression, partially offsetting the radiative cooling which would otherwise result in temperatures of -90 °C or lower (Schoeberl and Hartmann, 1991).

During the winter and spring, the circumpolar flow may be disrupted by global scale waves, or planetary waves. These waves, propagating vertically out of the troposphere, can produce total column ozone levels in excess of 400 DU (Salby and Garcia, 1990). As will be shown in Chapter 2, planetary waves are large scale Rossby waves, a type of wave motion possible in a rotating fluid medium. When they amplify in the stratosphere, planetary waves force the vortex center away from the pole, causing air to flow across latitude circles and allowing tracer constituents to be transported meridionally (eddy transport) over periods as short as one day. Antarctic stratospheric chemistry is therefore strongly influenced by planetary wave activity. Early studies (Kohno, 1984) reported that when a large zonal wave number 1 (wave 1 - see Chapter 2) planetary wave develops in the winter, an ozone wave is induced in the midlatitudes followed by an increase in total column ozone at high latitudes and a small decrease over the tropics. The importance of the horizontal and vertical transport associated with planetary waves in determining the dynamical distribution of ozone is discussed in greater detail in section 1.3.

The observed differences between the reductions of ozone over the Antarctic and over the Arctic, also indicate the importance of atmospheric dynamics in ozone depletion studies. There is abundant direct evidence for perturbed chemistry in the Arctic similar to that of the Antarctic, but the required cold temperatures are neither as widespread nor as long-lasting as in the spring season in Antarctica. The result is that ozone depletion in the Arctic has been 5% to 8% whereas that above the Antarctic has been as much as 50% (WMO report no. 20, 1990). These differences between the two polar regions are determined dynamically since their photochemical environments are similar due to similar solar illumination. Although it has been suggested that meridional transport by planetary waves may provide a dynamic mechanism for the hemispheric differences in the ozone distribution (Kohno, 1984), this may not be true for winter and spring conditions when the vortex is well formed. This is discussed in greater detail later. Furthermore, using the CCM2 model (see Chapter 7), Boville (1993) has shown that the difference between simulations of the Southern and Northern Hemisphere winter is due almost entirely to the parameterization of stationary gravity waves generated by flow over orography and not to differences in planetary wave activity. Gravity waves are discussed further in Chapter 2.

Differences in land mass distribution between the Northern and Southern Hemispheres result in significant differences in planetary wave activity which indirectly influences the isolation of the Northern and Southern Hemisphere winter vortices. In the Northern Hemisphere, mountain ranges in Asia and North America induce upward propagation of large-amplitude Rossby waves into the stratosphere (Schoeberl and Hartmann, 1991). The continent of Antarctica, however, is almost centered on the South Pole and is surrounded by ocean. Only the Andes mountain range and perhaps the Southern Alps of New Zealand weakly disturb the zonal symmetry of the surface conditions.

The Arctic circumpolar vortex is neither as large, as strong, nor as long-lived, as that in the Antarctic. It is well known that the propagation of planetary waves depends on the strength of the stratospheric westerly jet (Charney and Drazin, 1961). Previous analyses (Schoeberl and Geller, 1976) reported that differences in the intensity of the polar night jet between $60\,\mathrm{m\,s^{-1}}$ and $80\,\mathrm{m\,s^{-1}}$ cause very different planetary wave responses in the altitude regions both below and above the maximum velocity within the stratospheric jet. At the time of the final warming, zonal winds are usually 30 to $40\,\mathrm{m\,s^{-1}}$ in the Southern Hemisphere,

whereas in the Northern Hemisphere zonal winds are almost always less than $20\,m\,s^{-1}$, with the result that Northern Hemisphere planetary wave activity is enhanced. Increased planetary wave induced heat transport into the Arctic vortex results in Northern Hemisphere winter temperatures being 10 to 15 K warmer than Antarctic winter temperatures. This limits the probability of PSC formation and heterogeneous ozone depletion chemistry.

Occasionally, during the winter, the planetary wave field amplifies, the geostrophic wind weakens and polar temperatures increase. In the most severe cases this results in a reversal of the meridional temperature gradient and is known as a 'stratospheric sudden warming' (SSW). During a sudden warming, stratospheric temperatures in the polar night increase by as much as 50 K in a few days, becoming higher than those in the sunlit tropics (Salby and Garcia, 1990). At the same time, the zonal mean flow reverses direction, and ozone concentrations increase rapidly at high latitudes (Leovy et al., 1985). Although SSWs are seen to occur over the Arctic, they seldom occur over the Antarctic due to the suppressed Southern Hemisphere planetary wave activity. The breakdown of the polar vortex is associated with a final warming which usually develops in March in the Northern Hemisphere (spring equinox) but may be as late as early December (much later than the spring equinox) in the Southern Hemisphere (Schoeberl and Hartmann, 1991).

It has been noted (Newman, 1986) that in the Southern Hemisphere the final warming and the vortex breakdown are two distinct events, the warming preceding the vortex breakdown by 15 to 40 days. These events are nearly coincident in the Northern Hemisphere. The cause of this time lag is related to the thermal wind equation relating the temperature gradient to the vertical geostrophic wind shear (see Chapter 2). The final warming occurs when the wind maximum (zero temperature gradient) descends past the 30 hPa level, while the vortex breaks down as the zero wind line nears this level (Newman, 1986). Since the descent of the zero wind line is rapid over the Arctic, the final warming and vortex breakdown are nearly coincident. In the Southern Hemisphere the wind speeds are much stronger at the time of the final warming (see above) and this process is much slower, with the result that the vortex disappears 15 to 40 days after the final warming.

## 1.3 Planetary waves and the ozone distribution

The radiative-chemical-dynamical models of Antarctic ozone depletion are usually 2 dimensional models in the meridional plane. Since species concentrations are mostly zonally symmetric, computational overheads can be reduced and model accuracy maintained by restricting the model to 2 dimensions. However, deviations from the zonal means, or eddies, must still be accounted for. The eddy fluxes of quasi-conservative tracers are written (Reed and German, 1965) in terms of the zonal mean gradients of the species using so-called K-theory, to be discussed in Chapter 2. The predominant contribution to these eddy motions in the stratosphere are vertically propagating quasi-stationary planetary waves. Planetary waves generated in the troposphere by mountain ranges and thermal contrasts provide wave forcing to the stratosphere (Hartmann and Garcia, 1979; Randel, 1987). Since this propagation from the troposphere into the stratosphere occurs only when westerly winds are present (Charney and Drazin, 1961), the waves are confined to the winter hemisphere (Holton, 1992) where the required wind regimes are found.

Early studies (Pyle and Rogers, 1980) suggested that both the thermally driven mean circulation and diffusion at higher latitudes are responsible for ozone transport and that the eddy transport of chemical species due to planetary waves is the most important mechanism other than in cases of SSWs. Analyses of satellite data over limited time periods (Ghazi et al., 1976) showed increases in ozone at high latitudes during periods of strong wave activity. Hartmann and Garcia (1979) used a one-dimensional time dependent mechanistic model to show that a particular phase relationship between wave motion and perturbed ozone gives rise to enhanced poleward eddy transport in the middle and upper stratosphere.

Planetary waves may affect the total column ozone distribution directly through meridional and vertical transport or indirectly by modifying the properties of the Antarctic vortex. There is disagreement currently as to whether the Antarctic vortex acts as a 'flowing processor' or as an isolated material entity (Randel, 1993 a). The long-standing belief is that the vortex isolates the Antarctic stratosphere from the rest of the Southern Hemisphere stratosphere, creating a containment vessel inside which chemical depletion of ozone occurs. When discussing the polar vortex, it is also important to specify the altitude range over which it is

effective. Schoeberl and Hartmann (1991) suggest that the most isolated region of the vortex begins just above the tropopause (~10 km in the Antarctic) where the influence of tropospheric weather systems decreases and ends at about 30 km where irreversible meridional transport through planetary wave 'breaking' intensifies.

### 1.3.1 Planetary wave breaking

When planetary waves propagate, parcels of air are caused to move in a wave-like motion backward and forward across latitude circles. Provided dissipation is unimportant, the oscillatory motion of the air is completely reversible. However, dissipative processes prevent the air parcels from returning to their original latitudes, resulting in a net redistribution of ozone and other long-lived tracers. When the wave field amplifies, excursions of the airstream across latitude circles steepen to the point that the flow becomes dynamically unstable. The waves then 'break' (McIntyre and Palmer, 1984) and secondary vortices develop through instability. As they amplify, these unstable eddies stir bodies of air down to small dimensions where nonconservative processes are efficient. Dissipation then results in irreversible mixing of air and a net transport of constituents in the meridional plane. It is the poleward transport of heat during a wave-breaking event that causes an SSW. The overall frequency of wave breaking events also determines the effective isolation of the vortex (Schoeberl and Hartmann, 1991).

### 1.3.2 Direct transport of ozone by planetary waves

During the winter when planetary wave activity maximizes, the isolation of the polar vortex prevents planetary wave induced transport into the polar regions. Aircraft measurements of trace gases have shown that horizontal transport across the vortex boundary is insignificant (Randel, 1993 a). Although planetary waves are unable to transport trace gases deep into the polar vortex, they are able to 'erode' material away from the vortex edge. In the Southern Hemisphere the vortex edge is at approximately 65°S although this may vary considerably depending on planetary wave perturbations. The erosion results in the steepening of ozone gradients at the vortex edge and the irreversible mixing of air on the equatorward edge of the vortex with air at midlatitudes.

Planetary wave breaking in the midlatitudes creates a 'surf zone' where the strongest quasi-horizontal and irreversible tracer transport takes place (McIntyre and Palmer, 1984). This well mixed region, typically between 20°S and 50°S (Randel et al., 1993) is characterized by relatively flat meridional tracer gradients. During the early winter of each year, radiative cooling within the vortex results in diabatic descent of the air. Since air outside the vortex is warmer and radiative cooling rates are large, it might be expected that diabatic descent will be greater in this region. However, the lateral mixing by planetary waves in the surf zone transports heat poleward, offsetting the radiative cooling. The result is that air within the vortex is displaced downward by 2 to 3 km compared to air exterior to the vortex on the same pressure surface (Schoeberl and Hartmann, 1991).

As mentioned earlier, planetary waves can also modify the distribution of ozone by moving air vertically, especially in the lower stratosphere, where isentropic surfaces slope downward toward the pole. Air flowing poleward along these surfaces descends to higher pressure, compressing the air and increasing ozone concentrations and column abundances substantially. The relationship between ozone and isentropic pressure is equally strong for negative anomalies (Salby and Callaghan, 1993). The Antarctic has the highest mean elevation of all continents, resulting in surface pressures of approximately 700 hPa in the interior. Total column ozone levels below 200 DU over the south pole are attended by low pressures as a result of the high elevation of the continent which forces isentropic surfaces upward.

The final breakdown of the polar vortex occurs during the late winter or spring in both hemispheres and is associated with planetary wave activity. With the return of solar illumination to the Antarctic stratosphere, the meridional temperature gradient weakens resulting in a decrease in the strength of the geostrophic wind. During this final warming, or shortly thereafter in the case of the Southern Hemisphere, planetary wave amplitudes increase explosively, and the vortex shatters into smaller fragments that drift to midlatitudes (Schoeberl and Hartmann, 1991). Following this event, planetary waves mix ozone poor air from over the pole with ozone rich air equatorward of the vortex edge.

### 1.3.3 Indirect effects of planetary waves on ozone hole dynamics

Planetary waves may also influence the ozone distribution indirectly by perturbing the unique meteorological conditions within the vortex that are necessary for chlorine catalyzed ozone depletion. The amplitude of upward propagating planetary waves increases with altitude as does the magnitude of their heat transport (Schoeberl and Hartmann, 1991). During the winter, when stratospheric heating through UV absorption by ozone is reduced, heat transport above about 25 to 30 km results in an increase in vortex temperature with the result that PSCs are unable to form. Enhanced planetary wave activity will therefore reduce the altitude range over which PSCs form and will limit ozone destruction. The erosion of the vortex edge (see above) by planetary waves also decreases the areal extent of the vortex, reducing the spatial occurrence of PSCs.

Planetary wave induced shifts in the position of the vortex can affect polar chemistry and the concomitant ozone destruction. A strong winter planetary wave 1 perturbation shifts the center of the ozone hole off the pole, usually towards the South American continent. This has important implications for $NO_x$ chemistry since air parcel trajectories within the vortex are no longer confined to the polar night but experience short periods of sunlight (Solomon et al., 1993). Many models of $NO_x$ chemistry in the Antarctic fail to incorporate these effects since they assume that air contained inside the vortex experiences no illumination during the entire winter. The importance of $NO_2$ in reacting with ClO to convert reactive chlorine to $ClONO_2$, limiting ozone depletion, has been discussed above. When air parcels within the vortex are periodically illuminated as they circumnavigate the pole, the following reactions:

$$N_2O_5 + h\nu \rightarrow NO_2 + NO_3 \tag{1.21}$$

$$NO_3 + h\nu \rightarrow NO_2 + O \tag{1.22}$$

lead to a shift in the equilibrium towards $NO_2$ thus limiting ozone depletion (see equations 1.4 and 1.8). If the model of the vortex as a 'flowing processor' is accurate, then eddy transport of $NO_x$ into the top of the vortex by planetary waves may also contribute to the shift in the $NO_x$ equilibrium. Since planetary waves may alter the trajectory of an air parcel, the parcel may also move through regions of chemical sources and sinks in addition to regions of varying sunlight. Therefore the photochemistry of a species can have profound effects on its transport

by planetary scale waves (Pyle and Rogers, 1980)

The mechanisms directly influencing the global ozone distribution through transport and the processes responsible for indirect influences on ozone through perturbation of the Antarctic polar vortex are summarized in Figure 1.4.

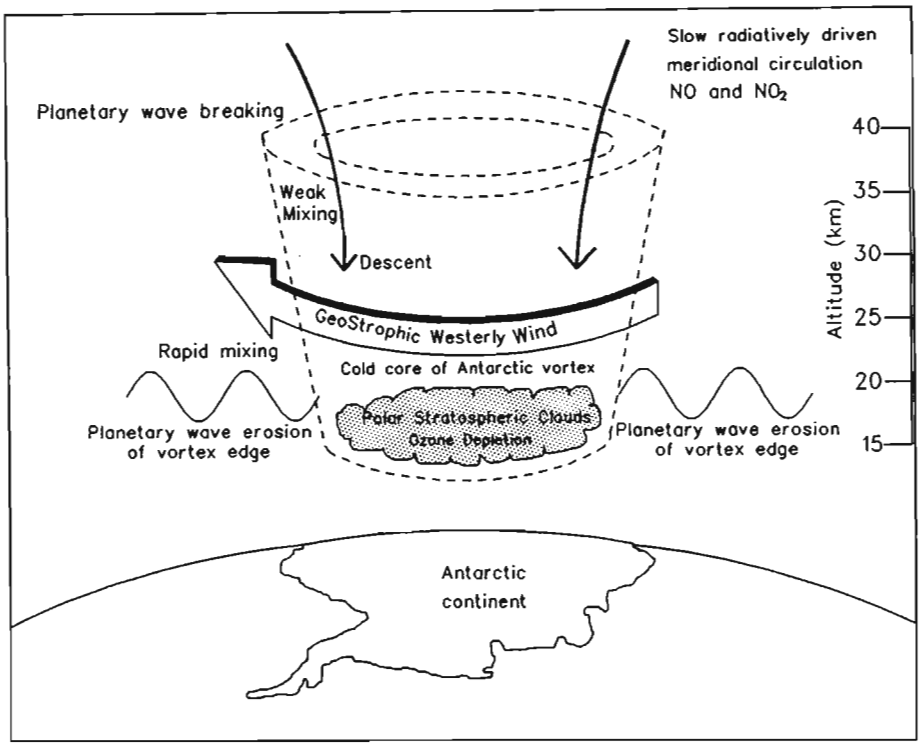


Figure 1.4: **Direct and indirect processes influencing the global distribution of total column ozone.**

Results from a numerical model designed to examine the relationship between planetary wave dynamics and polar ozone chemistry indicate that the interannual variations in the depth of the ozone hole are more likely to result from the variations in springtime planetary wave amplitudes than from the preceding wintertime ozone amounts (Austin and Butchart, 1992). Six experiments were performed where only planetary wave amplitudes were changed. Results indicated that high total column ozone over the pole always occurs for the largest wave amplitudes. It has also been reported (Schoeberl and Hartmann, 1991) that year-to-year fluctuations in the vortex strength and temperature appear to dominate over the slow increase in available stratospheric Cl in determining the interannual variability of the severity of the ozone hole. Shiotani and Gille (1987) stated that year-to-year variations of the ozone mixing ratio at high latitudes are related to variations of the wave activity during the winter and spring.

## 1.4 The scope of this work

Previous analyses (Schoeberl et al., 1986) have reported eastward propagating planetary waves observed in the Southern Hemisphere ozone distribution as seen by the Total Ozone Mapping Spectrometer (TOMS). In this thesis, data obtained from the TOMS experiment on board the Nimbus-7 satellite are used to create global maps of total column ozone for each day from 1979 to 1992. Planetary wave morphology within these distributions is examined using spatial spectral analysis (Chapter 4). The planetary wave powers obtained are compared with a measure of the depth and spatial extent of the Antarctic ozone hole.

A statistical model of ozone variation over the entire globe, from 1979 to 1992, is developed (Chapter 5). This model incorporates a long-term linear trend, an annual variation, a quasi-biennial oscillation, a solar cycle variation and a semi-annual variation. Model coefficients are examined for zonal asymmetries, possibly caused by planetary wave perturbations. Differences between model predicted global ozone and actual data are compared with planetary wave activity. The model is then applied to a study of long-term variability in total column ozone above five South African cities. Forward predictions of the monthly average total column ozone are then used as inputs to a surface UV irradiance model to allow prediction of the future UV environment for these cities (Chapter 6).

The use of potential vorticity conservation as a diagnostic for planetary wave propagation requires a knowledge of the conditions under which potential vorticity is conserved. In Chapter 7, data from the NCAR CCM2 model are used to suggest a method whereby potential vorticity 'lifetimes' may be estimated under different atmospheric conditions. This analysis also briefly examines the isolation of the Antarctic vortex, tracer gradient steepening across the vortex edge by planetary wave erosion, and wave mixing in the midlatitude 'surf zone'.

An overview of the mathematical theory associated with atmospheric waves is given in Chapter 2. Interpretation of the mathematical background provides insight into empirical results presented in later chapters. Since extensive use is made of the TOMS data base, Chapter 3 has been devoted to a discussion of the TOMS/SBUV instrument and the problems associated with total column ozone retrievals.

# Chapter 2

# BACKGROUND

## 2.1 Introduction

This chapter provides a theoretical foundation for atmospheric dynamics and planetary wave theory. To be brief, most of the discussion below will be of a general, qualitative nature. Fundamental equations are presented in detail, while derived equations are expanded only when particularly relevant to planetary wave propagation and its influence on the global ozone distribution. The material presented in this chapter was obtained in part from Pedlosky (1979), Brasseur and Solomon (1986), Andrews et al. (1987) and Holton (1992), where more in depth discussions of atmospheric dynamics may be found.

The fluid mechanical laws of conservation of mass, momentum and energy are embodied in a fundamental set of equations that may then be used to derive a mathematical description of the motions of the atmosphere. For the atmospheric motions of interest, the primary forces acting on an air parcel are the pressure gradient force, the gravitational force and friction. When applying Newton's second law to these forces in the rotating reference frame of the Earth, the centrifugal force and Coriolis force must be included among the forces acting on the air parcel. These forces, known as 'apparent forces', are discussed together with the three primary forces in section 2.2.

Scale analysis of the horizontal components of Newton's second law (the horizontal momentum equation) leads to the derivation of the geostrophic wind equations. For large scale motions away from the equator they provide an approxima-

tion to the actual horizontal wind field and account for the existence of the polar night jet. The thermal wind equation shows the effect of horizontal temperature gradients on the vertical shear in the geostrophic wind, as discussed in section 2.4.2.

Potential temperature and potential vorticity may be derived from observed meteorological variables and provide constraints on atmospheric flow. It has been mentioned in Chapter 1 that adiabatic, frictionless flow occurs along constant potential temperature surfaces, accounting for the observed increase in total column ozone with latitude. Conservation of potential vorticity, the atmospheric equivalent of angular momentum, provides the restoring force for planetary wave propagation. It addition, it may be used as a tracer of atmospheric flow. These two quasi-conserved atmospheric tracers are discussed in sections 2.5 and 2.6. The conservation of potential vorticity leads to the formulation of the potential vorticity equation. This prognostic (predictive) equation can be used as a starting point for the theory describing the behaviour of large-scale planetary waves as detailed in sections 2.7 and 2.8. Application of the theory to Rossby waves in the zonal flow shows how the strength of the vortex winds modulate planetary wave propagation (section 2.8.3).

A number of theoretical tools have been developed to investigate the effects of planetary waves on eddy transports and the interactions of planetary waves with the zonal flow. Some of these are reviewed in section 2.9. The Eliassen-Palm (EP) flux is an important diagnostic for the wave contribution to the eddy heat flux or eddy momentum flux. Planetary waves are responsible for the EP flux divergence which is a measure of the deceleration of zonal winds through wave transience and dissipation. It is therefore the primary dynamical term limiting the strength of the polar vortex. An outstanding question for 3-dimensional atmospheric modelling is the importance of wave drag, which provides a mechanism for the interaction of waves with the mean flow. This effect is usually parameterized in terms of a Rayleigh friction coefficient. In 2-dimensional atmospheric modelling, K-theory provides a formalism for eddy transport processes. A discussion of this theory is given in section 2.9.4 in preparation for work presented in Chapter 7.

Many previous studies have investigated the effect of planetary waves on the ozone distribution and transport of trace species. A review of this past research is delayed until later chapters where the work of this thesis is presented.

## 2.2 Fundamental equations

### 2.2.1 The pressure gradient force

Consider a parcel of air of mass $m$ and density $\rho$ in the presence of a pressure gradient $\nabla p$. By considering the balance of forces on the parcel it can be shown that the pressure gradient force per unit mass is given by:

$$\frac{\mathbf{F}}{m} = -\frac{1}{\rho}\nabla p \tag{2.1}$$

In a non-rotating coordinate system, parcels of air may be expected to accelerate down negative pressure gradients. However, in a rotating coordinate system, such flow is deflected by the Coriolis force (section 2.2.5).

### 2.2.2 The gravitational force

In addition to the pressure gradient force, a parcel of air of mass $m$ will experience the Earth's gravitational attraction given by:

$$\frac{\mathbf{F}_g}{m} = \mathbf{g} = -\frac{GM}{r^2}\left(\frac{\mathbf{r}}{r}\right) \tag{2.2}$$

where $G$ is the gravitational constant, M is the mass of the Earth, $\mathbf{r}$ is a vector from the center of the Earth to the air parcel and $\mathbf{g}$ is the gravitational acceleration vector of magnitude $9.8\,\mathrm{m\,s^{-2}}$. Since this equation is usually applied within 100 km of the surface of the Earth, the variation of $\mathbf{g}$ with altitude may be ignored.

### 2.2.3 Frictional forces

Frictional forces may be understood in terms of transport of momentum across plates of fluid undergoing shear stress. In fact, the transport of momentum down the velocity gradient, per unit area, per unit time, is simply the shearing stress. Similarly, random molecular motions transport heat down mean temperature gradients and trace constituents down mean mixing ratio gradients. The frictional force, a manifestation of viscosity within the fluid, is given by:

$$\frac{\mathbf{F}}{m} = \nu\nabla^2\mathbf{U} \tag{2.3}$$

where $\mathbf{U}$ is the wind vector field and $\nu$ is the kinematic viscosity coefficient with a sea level value of $\nu = 1.46 \times 10^{-5}\ \mathrm{m^2 s^{-1}}$. For the atmosphere below 100 km, $\nu$

is so small that molecular viscosity is negligible except within a few centimeters of the Earth's surface. Above this surface molecular boundary layer, momentum is transferred primarily by turbulent eddy motions.

### 2.2.4 The centrifugal force

When applying fluid dynamical theory to a description of the motions of the atmosphere, it is convenient to apply the fundamental equations in a reference frame that is rotating with the Earth. When using these laws in this noninertial reference frame of the rotating Earth, the acceleration of the coordinates must be considered. The change in coordinate systems requires the inclusion of the 'apparent' centrifugal and Coriolis forces in the statement of Newton's second law. The forces are termed 'apparent' since they do not arise from any direct physical process but only from the rotation of the coordinate axes.

Consider a parcel of air of mass $m$ on the surface of the rotating Earth. In an inertial reference frame fixed in space, the parcel will be seen to require a centripetal force to accelerate it towards the center of the Earth. However, in a noninertial reference frame, the centripetal force must be balanced to maintain the stationary position of the parcel. The required force, known as the centrifugal force, is given by:

$$\frac{\mathbf{F}}{m} = \Omega^2 \mathbf{R} \tag{2.4}$$

where $\Omega$ is the angular speed of the rotation of the Earth ($\Omega = 7.292 \times 10^{-5}$ rad s$^{-1}$) and $\mathbf{R}$ is the position vector from the axis of the Earth's rotation to the parcel. Since the centrifugal force is not directed towards the center of the Earth, it might be expected that it would have a small equatorward component. However, when forming, the Earth's geological structure was subject to the same forces, with the result that the Earth is not spherical, but spheroid in shape and the resultant force is everywhere perpendicular to the surface. The net effect of the centrifugal force is to reduce the apparent weight of an air parcel. For this reason the centrifugal force is usually incorporated into the expression for the Earth's gravity and is not treated explicitly. Gravity can be represented in terms of the gradient of a potential function $\Phi$, called the geopotential, by defining:

$$\frac{d\Phi}{dz} = g \tag{2.5}$$

### 2.2.5 The Coriolis force

If a parcel of air moves with respect to the surface of the Earth, a second 'apparent force', the Coriolis force, must also be incorporated into the expression of Newton's second law. Consider the motion of a control volume of air, of unit mass, that is free to move on a horizontal frictionless surface on the rotating Earth. If the air packet is in the Southern Hemisphere and set in motion in an eastward direction along a latitude circle it will experience a greater centrifugal force which will now have both a vertical and equatorward component. As shown in the previous section, the vertical component will cause a minor change in the apparent weight of the parcel, while the horizontal component will cause an equatorward deflection of the parcel. When moving equatorward, in order to conserve total angular momentum, the parcel must deflect westward. The net effect of the Coriolis force is anti-cyclonic (anti-clockwise) rotation in the Southern Hemisphere and anti-cyclonic (clockwise) rotation in the Northern Hemisphere. The Coriolis deflection is negligible for motions with time scales that are short compared to the rotation period of the Earth but is essential for an understanding of synoptic scale disturbances. The mathematical formulation of the Coriolis force per unit mass is (Holton, 1992):

$$\mathbf{F}_{Co} = 2\,\boldsymbol{\Omega}\times\mathbf{U} \tag{2.6}$$

It will be shown in section 2.8 that planetary waves owe their existence to the latitudinal variation of the Coriolis force.

### 2.2.6 The hydrostatic equation

In the absence of atmospheric motion, the vertical dependence of pressure is given by the hydrostatic equation:

$$\frac{dp}{dz} = -\rho g \tag{2.7}$$

The gravity force is balanced by the vertical component of the pressure gradient force. The constraint of no dynamic activity is weak and the condition of hydrostatic balance provides a good approximation to the real atmosphere. The hydrostatic equation may also be expressed in terms of geopotential height rather than geometric height. Using equation 2.5 and making substitutions from the ideal gas law:

$$d\Phi = gdz = -\rho dp = -(RT/p)dp = -RTd\ln p \tag{2.8}$$

Integrating this equation gives:

$$\Phi(z_1) - \Phi(z_2) = -R \int_{p_1}^{p_2} T d \ln p \tag{2.9}$$

Defining the geopotential height as $Z \equiv \Phi(z)/g_0$ (where $g_0 = 9.80665\ \mathrm{m\,s^{-2}}$) we obtain:

$$Z_T \equiv Z_1 - Z_2 = \frac{R}{g_0} \int_{p_1}^{p_2} T d \ln p \tag{2.10}$$

where $Z_T$ is the thickness of the layer between pressure levels $p_1$ and $p_2$ and the $T$ subscript denotes its temperature dependence.

## 2.3 The conservation laws

Atmospheric motions are governed by three underlying physical principles: conservation of mass, conservation of momentum and conservation of energy. In deriving the mathematical expressions for these laws, two coordinate systems may be used viz:

- The *Eulerian* frame of reference where the control volume is fixed with respect to the coordinate axes.
- The *Lagrangian* reference frame which moves with the fluid flow, always containing the same fluid particles.

Although, in the derivation of the fundamental laws, the two formalisms lead to the same result, one may be more elegant or provide a more physically intuitive approach than the other. In more advance theoretical analyses, it can be shown that Eulerian zonal averaging in the presence of waves tends to mask eddy transport effects, resulting in discrepancies between the net meridional transport and the circulation derived from eddy diffusion theory (section 2.9.4). This property of Eulerian zonal averaging, know as the non-interaction theorem, is reviewed by Brasseur and Solomon (1986) and discussed in detail by Matsuno (1980). The difficulties associated with eddy-mean flow cancellation led to the development of Lagrangian mean averaging (Andrews and McIntyre, 1976) where horizontal averages are calculated not along latitude circles, but along a path that follows the wave trajectory.

### 2.3.1 Conservation of momentum - Newton's second law

Assuming that the only true forces acting on the atmosphere are the pressure gradient force, gravitation and friction, Newton's second law may be obtained by adding the expressions for these effects, derived in section 2.2, to obtain:

$$\frac{D\mathbf{U}}{Dt} = -2\boldsymbol{\Omega}\times\mathbf{U} - \frac{1}{\rho}\nabla p + \mathbf{g} + \mathbf{F}_r \tag{2.11}$$

where $\mathbf{F}_r$ is the frictional force and the centrifugal force has been combined with gravitation in the gravity term $\mathbf{g}$. The operator $D/Dt$ is the total derivative given by:

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla \tag{2.12}$$

Equation 2.12 states that the total time rate of change of some quantity is given by the local rate of change plus an advective term giving the contribution to the local change through transport. For purposes of theoretical analysis and numerical modelling, it is convenient to transform equation 2.11 to spherical coordinates ($\lambda$, $\phi$, $z$), where $\lambda$ is the longitude, $\phi$ is the latitude, and $z$ is the vertical distance above the Earth's surface. Newton's second law may then be stated in component form as:

$$\frac{Du}{Dt} = \frac{\tan\phi}{a}uv - \frac{uw}{a} + 2\Omega v\sin\phi - 2\Omega w\cos\phi - \frac{1}{\rho a\cos\phi}\frac{\partial p}{\partial\lambda} + F_\lambda \tag{2.13}$$

$$\frac{Dv}{Dt} = -\frac{\tan\phi}{a}u^2 - \frac{uw}{a} - 2\Omega u\sin\phi - \frac{1}{\rho a}\frac{\partial p}{\partial\phi} + F_\phi \tag{2.14}$$

$$\frac{Dw}{Dt} = \frac{1}{a}u^2 + \frac{1}{a}v^2 + 2\Omega u\cos\phi - g - \frac{1}{\rho}\frac{\partial p}{\partial z} + F_z \tag{2.15}$$

where $u$, $v$ and $w$ are the eastward, northward and vertical wind velocity components, and $a$ is the radius of the Earth. This form of the momentum equation is fundamental to most work in dynamic meteorology, although many simplifying assumptions are made in its application. The quadratic terms proportional to $1/a$ are called the curvature terms and are negligible for midlatitude synoptic scale motions.

### 2.3.2 Conservation of mass - the continuity equation

The law of mass conservation, embodied in the continuity equation, may be derived in an Eulerian reference frame by considering the mass budget within a

fixed control volume. The continuity equation can then be stated in its mass divergence form as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0 \tag{2.16}$$

The equation states that the local time rate of change of density is equal in magnitude and opposite in sign to the mass divergence. Applying the relationship for the total derivative (equation 2.12), this equation can be manipulated into the velocity divergence form of the continuity equation:

$$\frac{1}{\rho}\frac{D\rho}{Dt} + \nabla \cdot \mathbf{U} = 0 \tag{2.17}$$

It is clear that for incompressible fluids, $\nabla \cdot \mathbf{U} = 0$.

### 2.3.3 Conservation of energy - the first law of thermodynamics

The first law of thermodynamics is usually derived by considering a system in thermodynamic equilibrium ie. a system that is initially at rest and after exchanging heat with its surroundings and performing work on its surroundings, returns to its rest state. The law stipulates that the change in internal energy of the system is equal to the difference between the heat added to the system and the work done by the system. By considering the rate of change of internal energy in a Lagrangian control volume, the first law of thermodynamics can be derived as:

$$c_v \frac{DT}{Dt} + p\frac{D\alpha}{Dt} = J \tag{2.18}$$

where $c_v$ is the specific heat at constant volume (for dry air $= 717\,\mathrm{J\,kg^{-1}\,K^{-1}}$), $\alpha = 1/\rho$ and $J$ is the heating rate per unit mass owing to radiation, conduction, and latent heat release.

## 2.4 Geostrophic balance and the thermal wind equation

### 2.4.1 The geostrophic wind

Scale analysis of the horizontal momentum equations (2.13 and 2.14) indicates that the Coriolis force and the pressure gradient force are in approximate balance, giving the geostrophic relationships (Holton, 1992):

$$fv \approx \frac{1}{\rho a \cos\phi}\frac{\partial p}{\partial \lambda} \tag{2.19}$$

$$= \frac{1}{\rho}\frac{\partial p}{\partial x} \tag{2.20}$$

$$fu \approx -\frac{1}{\rho a}\frac{\partial p}{\partial \phi} \tag{2.21}$$

$$= -\frac{1}{\rho}\frac{\partial p}{\partial y} \tag{2.22}$$

where $f \equiv 2\Omega \sin\phi$ is called the Coriolis parameter. A horizontal velocity field satisfying these equations is given by:

$$\mathbf{V}_g \equiv \mathbf{k} \times \frac{1}{\rho f}\nabla p \tag{2.23}$$

where $\mathbf{V}_g$ is know as the geostrophic wind and $\mathbf{k}$ is the local vertical unit vector. Application of this equation to the large meridional pressure gradient during the Southern Hemisphere winter (see Chapter 1) suggests a source for the zonal wind constituting the circumpolar vortex. As polar temperatures fall with the onset of winter, the meridional pressure gradient steepens ($\nabla p$ increases), resulting in an increase in the strength of the westerly wind given by equation 2.23 (recall that $f$ is negative in the Southern Hemisphere). This equation may provide a 10 to 15% accurate estimate of the strength of the vortex. The factor of $1/f$ in this equation results in division by zero at the equator, and hence it may only be applied in the middle and high latitudes.

### 2.4.2 The thermal wind equation

Not only does the steepening of the meridional temperature gradient result in an increase in the meridional pressure gradient, and hence an increase the strength of the geostrophic wind, but it also creates vertical shear in the geostrophic wind.

This is crucial for planetary wave propagation which not only requires a westerly wind but also depends on the variation of the polar night jet (geostrophic wind) intensity with altitude (section 2.8.3). The vertical shear in the geostrophic wind is given by the thermal wind equation, derived below.

By choosing pressure as the vertical coordinate (isobaric coordinate frame) and expressing the geostrophic wind equation 2.23 in this coordinate system (Holton, 1992)

$$\mathbf{V}_g = \mathbf{k} \times \frac{1}{f} \nabla_p \Phi \tag{2.24}$$

it can be seen that the geostrophic wind is proportional to the geopotential gradient on an isobaric surface. To see how a meridional temperature gradient affects $\nabla_p \Phi$, consider two isobaric surfaces separated by pressure $\delta p$. Equation 2.8 gives the thickness between the layers as:

$$\delta z \approx -\frac{RT}{g} \delta \ln p \tag{2.25}$$

Since $\delta \ln p$ is constant, a meridional gradient in temperature implies a meridional gradient in $\delta z$. Hence the steep meridional temperature gradient induced by the wintertime stratospheric radiative cooling creates a poleward decrease in $\delta z$ and a steepening of the geopotential surfaces with height. As shown in equation 2.24 this produces a zonal wind that increases in strength with height.

The explicit equations for the rate of change of the geostrophic wind with height may be derived by expanding equation 2.24 in component form as:

$$v_g = \frac{1}{f} \frac{\partial \Phi}{\partial x} \tag{2.26}$$

and

$$u_g = -\frac{1}{f} \frac{\partial \Phi}{\partial y} \tag{2.27}$$

Differentiating these equations with respect to pressure and applying equation 2.8 we obtain:

$$p \frac{\partial v_g}{\partial p} \equiv \frac{\partial v_g}{\partial \ln p} = -\frac{R}{f} \left( \frac{\partial T}{\partial x} \right)_p \tag{2.28}$$

and

$$p \frac{\partial u_g}{\partial p} \equiv \frac{\partial u_g}{\partial \ln p} = \frac{R}{f} \left( \frac{\partial T}{\partial y} \right)_p \tag{2.29}$$

Expressed in vectorial form, these equations constitute the thermal wind equation, which states that the vertical shear of the horizontal wind field is proportional to the horizontal temperature gradient. Strictly speaking, the thermal

wind refers to the vector difference between the geostrophic winds at two levels. Designating the thermal wind vector by $\mathbf{V}_T$, the thermal wind equation can be integrated from pressure level $p_0$ to $p_1$ to obtain:

$$\mathbf{V}_T \equiv \mathbf{V}_g(p_1) - \mathbf{V}_g(p_0) = -\frac{R}{f}\int_{p_0}^{p_1} (\mathbf{k}\times\nabla_p T)\, d\ln p \tag{2.30}$$

Letting $<T>$ be the mean temperature in the layer bounded by $p_0$ and $p_1$, the zonal and meridional components of the thermal wind are given by:

$$u_T = -\frac{R}{f}\left(\frac{\partial <T>}{\partial y}\right)_p \ln\left(\frac{p_0}{p_1}\right) \tag{2.31}$$

and

$$v_T = \frac{R}{f}\left(\frac{\partial <T>}{\partial x}\right)_p \ln\left(\frac{p_0}{p_1}\right) \tag{2.32}$$

These equations show that the polar vortex strength increases (decreases) with altitude in the presence of an equatorward (poleward) temperature gradient in both the Northern and Southern Hemisphere. Furthermore, asymmetries in the zonal temperature distribution caused by land-sea heating differences, convective heating in the tropical troposphere or localized planetary wave breaking events, will result in a non-zero $(\partial < T > /\partial x)_p$ term which will produce a meridional component to the geostrophic wind. This will advect material across latitude lines contributing to eddy transport. This phenomenon is discussed further in section 2.6.4.

## 2.5 Potential temperature

Taking the total derivative of the ideal gas law and using $c_p = c_v$ + R, where $c_p$ is the specific heat at constant pressure (1004 $\mathrm{J\,kg^{-1}\,K^{-1}}$), the first law of thermodynamics (equation 2.18) can be stated as:

$$c_p\frac{DT}{Dt} - \alpha\frac{Dp}{Dt} = J \tag{2.33}$$

The first two terms in this equation show the inverse relationship between temperature and pressure for a fluid undergoing adiabatic (J = 0) expansive cooling or compressive heating. In the stratosphere and mesosphere, the value of J is determined by heating through the absorption of UV by ozone and cooling by infrared emissions of ozone, $CO_2$ and water vapor. Dividing equation 2.33 by

temperature and using substitutions from the ideal gas law, the entropy form of the first law of thermodynamics may be stated as:

$$c_p \frac{D \ln T}{Dt} - R \frac{D \ln p}{Dt} = \frac{J}{T} \equiv \frac{Ds}{Dt} \tag{2.34}$$

where $s$ is the entropy. This equation gives the rate of change of entropy per unit mass, following the motion, for a thermodynamically reversible process (Holton, 1992). For an ideal gas undergoing an adiabatic process (a reversible process where no heat is exchanged with the surroundings) the first law of thermodynamics becomes:

$$c_p \, D \ln T \; - \; R \, D \ln p = 0 \tag{2.35}$$

Integrating this equation from a state at ($p$, $T$) to a state at ($p_s$, $\theta$), it can be shown that:

$$\theta = T(p_s/p)^{R/c_p} \tag{2.36}$$

The temperature $\theta$, known as the potential temperature, is that which a parcel of dry air at pressure $p$ and temperature $T$ would have if it were expanded or compressed adiabatically to a standard pressure $p_s$, usually taken to be the surface pressure. It can be shown (Holton, 1992) that since this value is conserved for dry, frictionless, adiabatic motion, most advective transport in the stratosphere occurs along surfaces of constant potential temperature (isentropic surfaces). It was mentioned in Chapter 1 that since these surfaces tend to decrease in altitude towards the polar regions, equatorial air moving poleward undergoes compression, substantially enhancing total column ozone values. Meridional cross-sections of potential temperature obtained from the NCAR CCM2 model are shown in Chapter 7.

### 2.5.1 Buoyancy oscillations

In an atmosphere where potential temperature increases with height, an air parcel undergoing adiabatic displacement from its equilibrium position will be positively (negatively) buoyant when displaced vertically downward (upward), so that it will tend to return to its initial altitude. The atmosphere is then described as stably stratified, as in the case of the stratosphere. Adiabatic oscillations of a fluid parcel about its equilibrium level are then referred to as buoyancy oscillations. The characteristic frequency of these oscillations is given by (Holton, 1992):

$$N^2 = g \frac{d \ln \theta_0}{dz} \tag{2.37}$$

where N is the Brunt-Väisälä frequency and $\theta_0$ is the potential temperature of the basic state. Note that if $d \ln \theta_0/dz$ is negative (turbulent atmosphere with $\theta$ decreasing with height) the value of $N$ will be pure imaginary. This has important implications for the growth or decay of gravity wave amplitudes with height, as discussed in sections 2.8.1 and 2.8.2.

## 2.6 Circulation and vorticity

In addition to potential temperature, potential vorticity also acts as a quasi-conserved tracer of atmospheric motion. The conservation of potential vorticity is the atmospheric equivalent of the conservation of angular momentum. Together with the variation of the Coriolis force (and hence the planetary vorticity) with latitude, potential vorticity conservation provides the mechanism for planetary wave propagation. This will be discussed further in section 2.6.4 where planetary wave forcing by differential zonal heating and by topographic forcing is discussed.

### 2.6.1 Vorticity - the curl of the wind field

The vorticity of a wind field is given by the curl of the wind velocity viz.:

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{2.38}$$

where $\zeta$ is known as the relative vorticity. This equation may be derived (Holton, 1992) by considering the circulation of a closed loop of fluid parcels in a wind field, and allowing the area enclosed by the loop to tend to zero. It is assumed that the horizontal components of the wind velocity are sufficiently larger than the vertical component that only the vertical component of the vorticity need be considered.

### 2.6.2 The Kelvin Circulation Theorem

Kelvin's Circulation Theorem provides an expression for the time rate of change of the circulation of a velocity field, providing the starting point for a statement of the conservation of potential vorticity. The equation is derived by taking the line integral of Newton's second law (equation 2.11) for a closed chain of fluid

parcels in an absolute coordinate system (inertial reference frame) and ignoring viscous forces to obtain:

$$\oint \frac{D_a \mathbf{U}_a}{Dt} \cdot d\mathbf{l} = -\oint \frac{\nabla p \cdot d\mathbf{l}}{\rho} - \oint \nabla \Phi \cdot d\mathbf{l} \tag{2.39}$$

where the $a$ subscript denotes that the integral is taken in an absolute coordinate system, which also accounts for the omission of the $2\Omega \times \mathbf{U}$ term. The gravitational term has been replaced by the gradient of the geopotential. Since

$$\frac{D_a \mathbf{U}_a}{Dt} \cdot d\mathbf{l} = \frac{D_a}{Dt}(\mathbf{U}_a \cdot d\mathbf{l}) - \mathbf{U}_a \cdot d\mathbf{U}_a \tag{2.40}$$

equation 2.39 may be written as:

$$\oint \frac{D_a}{Dt}(\mathbf{U}_a \cdot d\mathbf{l}) - \oint \mathbf{U}_a \cdot d\mathbf{U}_a = -\oint \frac{1}{\rho}(\nabla p \cdot d\mathbf{l}) - \oint \nabla \Phi \cdot d\mathbf{l} \tag{2.41}$$

Since the line integral of a perfect differential around a closed loop is zero:

$$\oint \nabla \cdot d\mathbf{l} = 0 \tag{2.42}$$

and

$$\oint \mathbf{U}_a \cdot d\mathbf{U}_a = 0 \tag{2.43}$$

equation 2.41 reduces to the circulation theorem:

$$\frac{DC_a}{Dt} = \frac{D}{Dt} \oint \mathbf{U}_a \cdot d\mathbf{l} = -\oint \frac{1}{\rho} dp \tag{2.44}$$

where $C_a$ is the circulation of the velocity field. The term on the right of equation 2.44 is known as the solenoidal term. In a barotropic fluid, where the density is a function of pressure only, the solenoidal term vanishes and the absolute circulation is conserved following the motion. This result is the Kelvin Circulation Theorem. For the analysis of atmospheric dynamics, it is more convenient to use relative rather than absolute circulation. In this noninertial reference frame, equation 2.44 becomes the Bjerknes circulation theorem (Holton, 1992):

$$\frac{DC}{Dt} = -\oint \frac{dp}{\rho} - 2\Omega \frac{DA_e}{Dt} \tag{2.45}$$

where $A_e$ is the projection of the area of the circulation onto the Earth's equatorial plane.

### 2.6.3 Potential vorticity

Using the ideal gas law together with the definition for potential temperature (equation 2.36) it can be shown that:

$$\rho = p^{c_v/c_p} \frac{(p_s)^{R/c_p}}{R\theta} \tag{2.46}$$

Therefore on an isentropic surface ($\theta$ = const) the solenoidal term in the Kelvin circulation theorem becomes:

$$\oint \frac{dp}{\rho} \propto \oint dp^{1-c_v/c_p} = 0 \tag{2.47}$$

This is equivalent to the characteristic of a barotropic fluid (see above) and for adiabatic flow satisfies Kelvin's circulation theorem, which may be stated in the Bjerknes form as:

$$\frac{D}{Dt}(C + 2\Omega\,\delta A \sin\phi) = 0 \tag{2.48}$$

Since the vertical component of the vorticity is given by the circulation divided by the area of the loop, as the area tends to zero we obtain:

$$\delta A(\zeta_\theta + f) = const \tag{2.49}$$

where $\zeta_\theta$ is the vertical component of relative vorticity on the isentropic surface and $f$ is the Coriolis parameter (see section 2.4.1). Consider a parcel of air confined between potential temperature surfaces $\theta_0$ and $\theta_0 + \delta\theta$, which are separated by a pressure $-\delta p$. Since the mass of the parcel is conserved it can be shown that:

$$\delta A = -\frac{\delta M g}{\delta p} = \left(-\frac{\delta\theta}{\delta p}\right)\left(\frac{\delta M g}{\delta\theta}\right) = \text{const} \times g\left(-\frac{\delta\theta}{\delta p}\right) \tag{2.50}$$

Since the surfaces considered are isentropic, $\delta\theta$ is constant and $\delta A$ can be eliminated in equation 2.49. Taking $\delta p \rightarrow 0$ equation 2.50 reduces to:

$$P \equiv (\zeta_\theta + f)\left(-g\frac{\partial\theta}{\partial p}\right) = const \tag{2.51}$$

where P is the isentropic coordinate form of Ertel's potential vorticity in units of $\mathrm{K\,kg^{-1}\,m^2\,s^{-1}}$.

### 2.6.4 Potential vorticity and planetary wave propagation

Assuming the atmosphere to be a homogeneous incompressible fluid, the mass of a parcel of air is given by:

$$M = \rho\,\delta A\,\delta z \tag{2.52}$$

Since M and $\rho$ are constant, this may be substituted into equation 2.49 to obtain:

$$(\zeta + f)/\delta z = constant \tag{2.53}$$

where $\zeta$ is evaluated at fixed altitude. This conservation condition imposes strong constraints on the large scale atmospheric flow. Together with the thermal wind equations (2.31 and 2.32), it can be used to demonstrate the generation of planetary scale waves by zonal temperature variations and topographic forcing. The required temperature variations may result from land-sea differences in heating or by convective heating in the tropical troposphere, where large quantities of latent heat are released in cumulus cloud systems (Salby and Hartmann, 1990). These two forcing mechanisms are discussed further below.

### Forcing through differential heating

Consider purely eastward geostrophic flow ($\zeta = 0$), between 50° and 60°S, approaching an area of colder temperatures. Analysis at constant altitude allows the use of equation 2.53 in constraining the flow. Air parcels approaching the region of cold air encounter a temperature gradient which causes vertical shear in the geostrophic wind as given by the thermal wind equations (2.31 and 2.32). The situation is shown schematically in Figure 2.1.

Latitude is plotted on the ordinate while longitude is plotted on the abscissa. Actual longitude values are irrelevant for this argument and are excluded from Figure 2.1 for clarity. The dashed curve on the left of the figure shows the change in planetary vorticity ($f$) in units of $10^{-4}$ rad s$^{-1}$. Note that the value of $f$ is negative, and becomes more negative with increasing latitude. The temperature gradient shown in the figure creates an anti-cyclonic turning in the zonal wind with altitude, resulting in equatorward flow at some higher level. However, as the air at this higher level streams equatorward, it encounters increasing (less negative) planetary vorticity and therefore must decrease its relative vorticity ie. $\zeta < 0$. This results in the creation of the circulation cell labelled **a** in Figure 2.1. This cyclonic circulation turns the zonal wind poleward such that air parcels stream across latitude lines past the equilibrium position ($\zeta = 0$). The flow now encounters decreasing (more negative) planetary vorticity and must therefore increase its relative vorticity as dictated by equation 2.53 ie. $\zeta > 0$. This results in the formation of the circulation cell labelled **b** in Figure 2.1. The net result

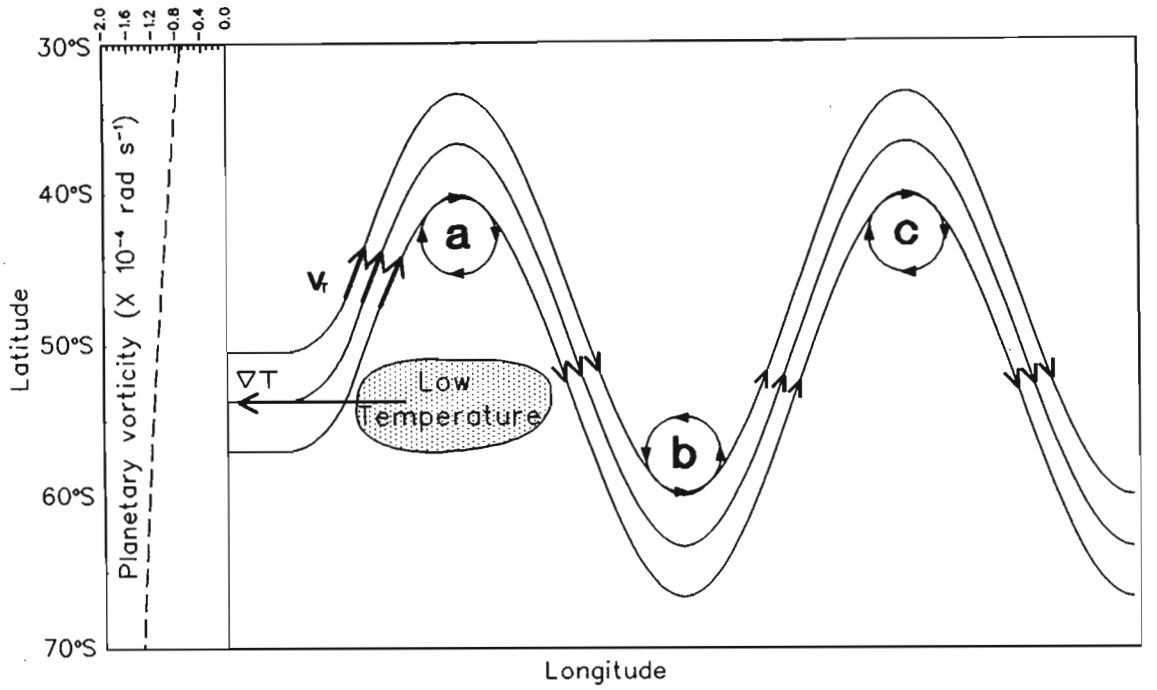


Figure 2.1: **Perturbations to the zonal wind by changes in local temperature as stipulated by the thermal wind equations and potential vorticity conservation arguments.**

of the initial perturbation to the zonal wind speed and direction is the wavelike behaviour shown in Figure 2.1, thus creating a planetary scale wave.

### Topographic forcing

Consider again uniform zonal flow ($\zeta = 0$) in an eastward direction, approaching a mountain range as depicted in Figure 2.2.

The lower half of the figure shows the flow in the (x, z) plane, together with the expected changes in $\zeta$ while the upper half of the figure shows the resultant perturbation to the zonal flow in the (x, y) plane.

If the flow is adiabatic, it will be bounded below by some isentropic surface denoted $\theta_0$ and above by $\theta_0 + \delta\theta$. These surfaces of constant potential temperature are also distorted by the presence of the mountain range, as shown in Figure 2.2. Equation 2.51 can now be used to determine the required perturbations to the flow such that potential vorticity is conserved. As a result of the vertical displacement of the upper level isentropic surface upstream of the mountain range, the air column is stretched, causing $\partial\theta/\partial p$ to decrease. To conserve potential

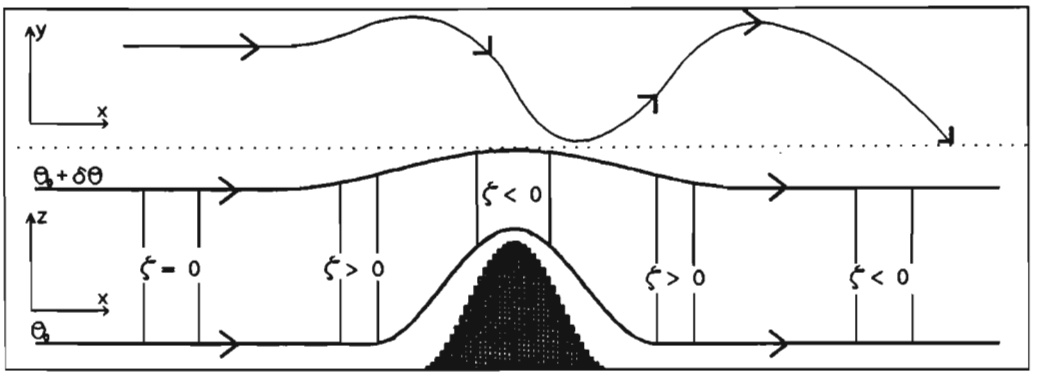


Figure 2.2: **Perturbations to zonal flow arising from flow over changing topography. Adapted from Holton (1992).**

vorticity, $\zeta$ must increase (see equation 2.51). This results in the equatorward turning of the zonal wind just upstream of the barrier as shown in the upper half of Figure 2.2. Of course as the air parcels stream equatorward they encounter increasing planetary vorticity which reduces the change in $\zeta$ required for potential vorticity conservation. Later, as the airstream crosses the mountain range, it is compressed causing $\partial\theta/\partial p$ to increase and $\zeta$ to become negative. This causes poleward turning of the zonal wind as shown in the upper half of Figure 2.2. Using the potential vorticity arguments from the previous section, a wavelike pattern in the zonal flow will be created downstream of the mountain range.

It can be further shown (Holton, 1992) that for the case of easterly flow over a mountain range, the disturbance to the zonal flow damps out away from the barrier and no wavelike flow pattern forms. Hence only westerly flow permits topographically forced planetary waves. These two qualitative examples suggest that a valid starting point for the investigation of planetary scale wave motion, is to derive an expression for the time dependence of the absolute vorticity. This theoretical development is presented in the following section.

## 2.7 The potential vorticity equation

Use of the fundamental momentum equations (2.13 to 2.15) in the modelling of atmospheric dynamics is computationally prohibitive. Fortunately, equations describing many types of atmospheric motion can be derived by replacing these 3D partial differential equations with a potential vorticity equation. Furthermore, in many instances, the physical interpretation of atmospheric dynamics is elucidated

when examined in light of vorticity conservation.

To derive a conservation expression for absolute vorticity, we begin with the fundamental horizontal momentum equations (2.13 and 2.14) and retain acceleration terms. Retaining these terms introduces the time dependence which allows the development of prognostic equations. Ignoring friction and curvature terms (terms containing $1/a$), and expanding the total time derivative, the horizontal momentum equations can be stated as:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} - fv = -\frac{1}{\rho}\frac{\partial p}{\partial x} \tag{2.54}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} + fu = -\frac{1}{\rho}\frac{\partial p}{\partial y} \tag{2.55}$$

Taking $\partial/\partial y$ of equation 2.54 and $\partial/\partial x$ of equation 2.55 and subtracting the first from the second we obtain:

$$\frac{\partial \zeta}{\partial t} + u\frac{\partial \zeta}{\partial x} + v\frac{\partial \zeta}{\partial y} + w\frac{\partial \zeta}{\partial z} + (\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)$$

$$+ \left(\frac{\partial w}{\partial x}\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}\frac{\partial u}{\partial z}\right) + v\frac{\partial f}{\partial y} = \frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial p}{\partial y}\frac{\partial p}{\partial x}\right) \tag{2.56}$$

where we have made use of:

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{2.57}$$

The first four terms constitute the total derivative of $\zeta$ and since the Coriolis parameter ($f$) depends only on $y$, the equation may be rewritten in the form:

$$\frac{D}{Dt}(\zeta + f) = -(\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \quad \text{[Divergence term]} \tag{2.58}$$

$$-\left(\frac{\partial w}{\partial x}\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}\frac{\partial u}{\partial z}\right) \quad \text{[Tilting/Twisting term]} \tag{2.59}$$

$$+\frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y}\frac{\partial p}{\partial x}\right) \quad \text{[Solenoidal term]} \tag{2.60}$$

This equation, describing the time rate of change of absolute vorticity, is made up of three terms viz.:

The divergence term: Consider a chain of fluid parcels in a wind field with positive divergence. Since the area enclosed by the chain will increase with time, the absolute vorticity will decrease (see section 2.6.1). This is the fluid analog of the change in angular momentum resulting from a change in the moment of inertia of a solid body when angular momentum is conserved (Holton, 1992).

The tilting or twisting term: This term allows for changes in the vertical vorticity resulting from the tilting of horizontal components of vorticity into the vertical by a nonuniform vertical wind field.

The solenoidal term: This is the microscopic equivalent of the solenoidal term in the Kelvin circulation theorem (equation 2.44). Using $\alpha = 1/\rho$, the solenoidal term may be expressed as (Holton, 1992):

$$-\left(\frac{\partial \alpha}{\partial x}\frac{\partial p}{\partial y} - \frac{\partial \alpha}{\partial y}\frac{\partial p}{\partial x}\right) = -(\nabla\alpha \times \nabla p)\cdot \mathbf{k} \tag{2.61}$$

### 2.7.1 The vorticity equation in isobaric coordinates

The vorticity equation (2.60) can be expressed in a more useful way in the isobaric coordinate system where the vertical coordinate ($z$) is replaced by pressure ($p$). Recall that the approximate horizontal momentum equations (2.54 and 2.55) can be written in vector form as:

$$\frac{D\mathbf{V}}{Dt} + f\mathbf{k}\times\mathbf{V} = -\frac{1}{\rho}\nabla p \tag{2.62}$$

where $\mathbf{V}$ is the horizontal velocity vector with zero vertical component. The vertical coordinate transformation from altitude to pressure is facilitated using the identities (Holton, 1992)

$$-\frac{1}{\rho}\left(\frac{\partial p}{\partial x}\right)_z = -\left(\frac{\partial \Phi}{\partial x}\right)_p \tag{2.63}$$

and

$$-\frac{1}{\rho}\left(\frac{\partial p}{\partial y}\right)_z = -\left(\frac{\partial \Phi}{\partial y}\right)_p \tag{2.64}$$

allowing equation 2.62 to be expressed as:

$$\frac{D\mathbf{V}}{Dt} + f\mathbf{k}\times\mathbf{V} = -\nabla_p\Phi \tag{2.65}$$

where $\nabla_p$ is the horizontal gradient operator applied at constant pressure. By operating on equation 2.65 with the vector operator $\mathbf{k}\cdot\nabla\times$ (the $p$ subscript on $\nabla$ has been omitted for clarity) and using $\zeta = \mathbf{k}\cdot(\nabla\times\mathbf{V})$, the vorticity equation in isobaric coordinates can be derived as (Holton, 1992):

$$\frac{\partial \zeta}{\partial t} = -\mathbf{V}\cdot\nabla(\zeta + f) - \omega\frac{\partial \zeta}{\partial p} - (\zeta + f)\nabla\cdot\mathbf{V} + \mathbf{k}\cdot\left(\frac{\partial \mathbf{V}}{\partial p}\times\nabla\omega\right) \tag{2.66}$$

where $\omega = Dp/Dt$ is the pressure change following the motion, playing the same role in isobaric coordinates as $w = Dz/Dt$ in height coordinates. Scale analysis of the vorticity equation (Holton, 1992) shows that to a first approximation, valid for synoptic scale systems:

$$\frac{D_h(\zeta + f)}{Dt} = -f\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \tag{2.67}$$

where

$$\frac{D_h}{Dt} \equiv \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} \tag{2.68}$$

Equation 2.67 shows that for planetary scale systems, the time rate of change of absolute vorticity, following the horizontal motion, is given by the creation (destruction) of vorticity owing to convergence (divergence) of the horizontal wind.

### 2.7.2 The barotropic (Rossby) potential vorticity equation

In the barotropic model of the atmosphere, fluid parcels are incompressible and the continuity equation (2.17) reduces to $\nabla \cdot \mathbf{U} = 0$. This allows the divergence of the horizontal wind field in equation 2.67 to be replaced by $-\partial w/\partial z$ to obtain:

$$\frac{D_h(\zeta + f)}{Dt} = (\zeta + f)\left(\frac{\partial w}{\partial z}\right) \tag{2.69}$$

where the relative vorticity in the divergence term has been retained (Holton, 1992). For an ideal gas in a barotropic atmosphere, isobaric surfaces are also isothermal. Hence the temperature derivative terms in the thermal wind equations (2.28 and 2.29) are zero and the geostrophic wind is independent of height. Approximating the wind field by the geostrophic wind ($u_g$, $v_g$) and integrating from the lower boundary ($z_1$) of the atmosphere to the upper boundary ($z_2$), equation 2.69 becomes:

$$h\frac{D_h(\zeta_g + f)}{Dt} = (\zeta_g + f)[w(z_2) - w(z_1)] \tag{2.70}$$

where $h$ is the depth of the atmosphere and $\zeta_g$ denotes that the vorticity is derived from the geostrophic wind. Since $w \equiv Dz/Dt$ and $h$ is independent of height, this equation may be expressed as:

$$\frac{1}{\zeta_g + f}\frac{D_h(\zeta_g + f)}{Dt} = \frac{1}{h}\frac{D_h h}{Dt} \tag{2.71}$$

or

$$\frac{D_h \ln(\zeta_g + f)}{Dt} = \frac{D_h \ln h}{Dt} \tag{2.72}$$

which implies that:

$$\frac{D_h}{Dt}\left(\frac{\zeta_g + f}{h}\right) = 0 \tag{2.73}$$

which is equivalent to equation 2.53 ie:

$$\frac{\zeta_g + f}{h} = const \tag{2.74}$$

The barotropic potential vorticity is conserved following the motion. If the wind field is purely horizontal $D_h h/Dt = 0$ and the barotropic vorticity equation can be stated as:

$$\frac{D_h(\zeta_g + f)}{Dt} = 0 \tag{2.75}$$

expressing the conservation of absolute vorticity following the horizontal motion.

### 2.7.3 The quasi-geostrophic vorticity equation

Although the geostrophic wind equations provide a good approximation to the observed horizontal wind field for circulation systems greater than about 1000 km in size, they cannot be used prognostically since they contain no time derivatives. However, it is possible to derive a set of prognostic equations which reduces to the geostrophic balance as a first order approximation, but also includes terms accounting for ageostrophic flow. These equations, known as the quasi-geostrophic equations, may then be used to derive the quasi-geostrophic vorticity equation. This equation provides a method for predicting the evolution of the geopotential height field, given its initial 3-dimensional distribution.

For large-scale, low frequency perturbations, the requirements of hydrostatic and geostrophic balance constrain the baroclinic motions of the atmosphere. Under these conditions, the structure and evolution of the 3-dimensional velocity field are determined, to a good approximation, by the distribution of geopotential height on isobaric surfaces. The equations that express these relationships are the quasi-geostrophic equations from which the quasi-geostrophic vorticity equation may be derived (Holton, 1992):

$$\frac{\partial \zeta_g}{\partial t} = -\mathbf{V}_g \cdot \nabla(\zeta_g + f) + f_0 \frac{\partial \omega}{\partial p} \tag{2.76}$$

This equation will be interpreted in section 2.8.3 to provide a qualitative description of Rossby wave propagation.

## 2.8 Wave theory

In this section, the elementary equations describing wave propagation are developed. The perturbation technique is presented and then used to derive the wave propagation equations for internal gravity waves and Rossby waves. Parts of the derivation below have been adapted from Holton (1992) and Pedlosky (1979). Rossby wave equations are derived in detail in preparation for Chapter 4 where it is shown that Rossby waves may be identified in the global total column ozone distribution.

### 2.8.1 Fundamental concepts

#### The perturbation technique

To derive atmospheric wave equations in detail, the non-linear quasi-geostrophic potential vorticity equation must be solved numerically. However, such numerical analysis provides no insight or intuitive understanding of the processes involved in wave generation and propagation. To provide a theoretical formalism for wave behaviour (as opposed to numerical computation), the fundamental partial differential equations of mass, momentum and energy conservation need to be simplified. This is accomplished by linearizing the equations using the perturbation technique.

In the perturbation method, all field variables are divided into two parts, a basic state which is independent of longitude and time, and the perturbed state which is the local deviation of the field from the basic state. The basic state is denoted with a zero subscript while the perturbed state is denoted with an $e$ subscript (excess). Field variables could then be expressed as:

$$\begin{aligned} u(x,t) &= u_0 + u_e(x,t) \\ p(x,t) &= p_0 + p_e(x,t) \\ \rho(x,t) &= \rho_0 + \rho_e(x,t) \end{aligned}$$

For the meridional components (eg. $v$), the perturbed quantity ($v_e$) is the eddy contribution. These eddy fluctuations can then be described in terms of wave motions.

The notation used in this analysis differs from that of the two references listed

above where the prime notation is reserved for denoting first order spatial derivatives. The perturbation technique requires that:

1. The basic state fields satisfy the fundamental equations when the perturbations are set to zero.

2. The perturbation fields are sufficiently small that products of perturbation terms may be neglected.

### Wavenumbers

In Chapter 4, a Fourier transform technique is used to extract wave powers from planetary wave structures observed in the global TOMS ozone distributions. A discussion of the Fourier techniques used is given in Appendix A. The Fourier transform of a function $h(x)$ is given by:

$$H(k_x) = \int_{-\infty}^{\infty} h(x)\, e^{2\pi i k_s x} dx \tag{2.77}$$

where $k_s = 2\pi s/L$ is the zonal wavenumber (units $\mathrm{m}^{-1}$), $L$ is the distance around a latitude circle, and $s$, the planetary wavenumber, is an integer specifying the number of waves around the zone. In the analysis below, and in the following chapters, the term 'wave 1' refers to a wave with a planetary wavenumber of 1.

### Phase velocity, group velocity and dispersion

The phase speed of a wave is the speed at which an observer must move in order to observe a constant phase on the wave eg. to remain at a peak, or at a trough of the propagating waveform. The phase speed is given by $\nu/k$ where $\nu$ is the wave frequency and $k$ is its wavenumber.

Waves for which the wave phase speed is dependent on (independent of) wavenumber are termed dispersive (nondispersive). The formula relating the frequency $\nu$ and the wavenumber $k$ is called the dispersion relation.

For dispersive waves, the shape of the waveform does not remain constant as the wave propagates since different frequency components will move at different speeds. When waves are dispersive, the velocity of the waveform (the group velocity) will differ from the phase velocity. The group velocity is given by $\partial\nu/\partial k$.

### Exponential growth and decay of waves

Consider a wave equation of the form $A = e^{i\nu t}$. Suppose that $\nu$ itself has real and imaginary components ie. $\nu = \nu_r + i\nu_i$. The wave equation then takes the form:

$$A = e^{-\nu_i t} e^{i\nu_r t} \tag{2.78}$$

In addition to a sinusoidally varying part ($e^{i\nu_r t}$), the wave amplitude also has an ($e^{-\nu_i t}$) exponential envelope. If $\nu_i$ is positive, the wave will decay with time and if $\nu_i$ is negative the wave amplitude will increase exponentially with time.

## 2.8.2 Internal gravity waves

### Pure internal gravity waves

The internal gravity wave equations outlined below were derived independently of the two references listed above (Holton, 1992; Pedlosky, 1979) and will therefore differ in appearance but not in interpretation.

In this section, internal gravity waves are investigated and the dispersion relation is derived. For these waves, the restoring force is provided by **g**. The waves can exist only when the atmosphere is stably stratified so that fluid parcels displaced vertically will undergo buoyancy oscillations (section 2.5.1). Recall that the frequency of these oscillations is given by:

$$N^2 = g \frac{d \ln \theta_0}{dz} \tag{2.79}$$

We make use of the Boussinesq approximation in which the density is treated as a constant except when it is coupled with gravity in the buoyancy term of the vertical momentum equation (Holton, 1992). In this approximation, the atmosphere is treated as being incompressible and local density variations are assumed to be small perturbations to the basic state.

The gravity wave equations are derived in the (x, z) space shown in Figure 2.3. Consider the field variables at two altitudes separated by a small distance $\varepsilon$. Hydrostatic pressure balance (equation 2.7) requires:

$$\frac{dp_0(z)}{dz} = -\rho_0(z) g \tag{2.80}$$

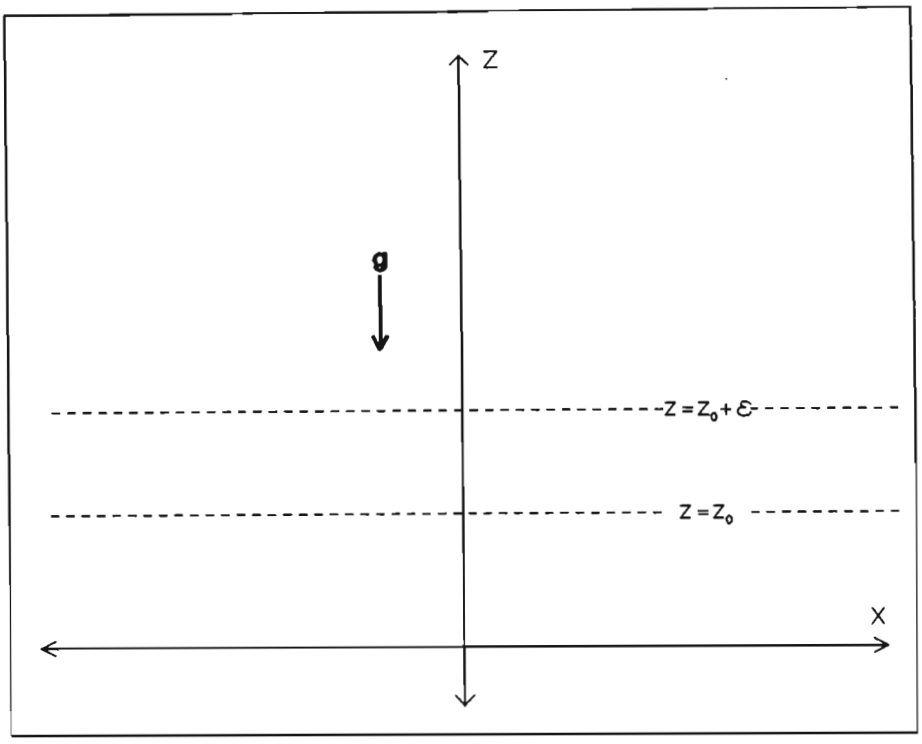


Figure 2.3: The (x, z) space in which the gravity wave equations are derived.

where the 0 subscript denotes the basic state values. The density and pressure equations at the upper boundary are then:

$$\rho_0(z+\varepsilon) = \rho_0(z) + \varepsilon\left(\frac{d\rho_0}{dz}\right) \tag{2.81}$$

$$\begin{aligned} p_0(z+\varepsilon) &= p_0(z) + \varepsilon\left(\frac{dp_0}{dz}\right) \\ &= p_0(z) - \rho_0 g\varepsilon \end{aligned} \tag{2.82}$$

In moving from $z = z_0$ to $z = z_0 + \varepsilon$, the fluid element experiences the pressure drop adiabatically, so that

$$\delta\rho = \frac{\delta p}{c_0^2} \tag{2.83}$$

where $c_0^2$, the speed of sound squared, is given by $\gamma p_0/\rho_0$ and $\gamma = c_p/c_v$. From equation 2.82, the pressure difference between the two levels is $-\rho_0 g\varepsilon$. Therefore:

$$\delta\rho = -\rho_0 g\varepsilon/c_0^2 \tag{2.84}$$

and the density of the fluid element is:

$$\rho = \rho_0 - \rho_0 g\varepsilon/c_0^2 \tag{2.85}$$

To calculate the restoring force on the fluid element at $z+\varepsilon$, consider the excess density of the fluid element over its surroundings (equation 2.85 – equation 2.81):

$$\rho_e = -\varepsilon\left(\frac{\rho_0}{c_o^2}g + \frac{dp_0}{dz}\right) \tag{2.86}$$

the excess weight is then $g\rho_e$ and the application of Newton's second law in the vertical gives:

$$\frac{\partial^2\varepsilon}{\partial t^2} = -\varepsilon g\left(\frac{\rho_0'}{\rho_0} + \frac{g}{c_0^2}\right) \tag{2.87}$$

where $\rho_0' = \partial\rho_0/\partial z$. Solutions to this equation are of the form $\varepsilon = Ae^{iNt}$ where:

$$N^2 = g\left(\frac{\rho_0'}{\rho_0} + \frac{g}{c_0^2}\right) \tag{2.88}$$

and are oscillatory for $N^2 > 0$. It may be shown that $N^2$ is equivalent to the Brunt-Väisälä frequency defined in section 2.5.1. For $N^2 > 0$, the atmosphere is stably stratified, while for $N^2 < 0$, the atmosphere is unstable and small perturbations amplify exponentially.

Making use of the linearization method, the wave equations for internal gravity waves in 3 dimensions are derived to provide a more detailed understanding of the wave propagation. Consider perturbations to a stable incompressible atmosphere where the continuity equation 2.16 becomes:

$$\nabla\cdot(\rho_0\mathbf{U}) = 0 \tag{2.89}$$

The linearized momentum equation may be shown to be:

$$\rho_0\frac{\partial\mathbf{U}}{\partial t} + \nabla p_e = \rho_e\mathbf{g} \tag{2.90}$$

Taking the divergence of equation 2.90 and noting that $\mathbf{g}$ is in the negative z direction, we obtain:

$$\frac{\partial}{\partial t}(\nabla(\rho_0\mathbf{U})) + \nabla^2 p_e = -g\frac{\partial\rho_e}{dz} \tag{2.91}$$

and using equation 2.89 this becomes:

$$\nabla^2 p_e = -g\frac{\partial\rho_e}{\partial z} \tag{2.92}$$

The z component of equation 2.90 can be expressed as:

$$\frac{\partial q}{\partial t} + \frac{\partial p_e}{\partial z} = -g\rho_e \tag{2.93}$$

where $q = \rho_0 w$ is the vertical mass flux associated with the wave. Taking $\partial/\partial z$ of equation 2.92 and $\nabla^2$ of equation 2.93 it can be shown that:

$$\frac{\partial}{\partial z}\nabla^2 p_e = -g\frac{\partial^2 \rho_e}{\partial z^2} \tag{2.94}$$

$$\nabla^2\frac{\partial q}{\partial t} + \nabla^2\frac{\partial p_e}{\partial z} = -g\left(\frac{\partial^2 \rho_e}{\partial x^2} + \frac{\partial^2 \rho_e}{\partial y^2} + \frac{\partial^2 \rho_e}{\partial z^2}\right) \tag{2.95}$$

$$\Rightarrow \nabla^2\frac{\partial q}{\partial t} = -g\left(\frac{\partial^2 \rho_e}{\partial x^2} + \frac{\partial^2 \rho_e}{\partial y^2}\right) \tag{2.96}$$

Recall from above (equation 2.87) that Newton's second law for gravity waves can be stated as:

$$\rho_0\frac{\partial^2 \varepsilon}{\partial t^2} = g\rho_e \tag{2.97}$$

from which the wave equation may be obtained:

$$\frac{\partial^2 \varepsilon}{\partial t^2} = N^2\varepsilon \tag{2.98}$$

Substituting equation 2.98 into equation 2.97 we obtain:

$$\rho_0 N^2 \varepsilon = g\rho_e \tag{2.99}$$

taking partial time derivatives of this equation produces:

$$\begin{aligned} g\frac{\partial \rho_e}{\partial t} &= N^2\rho_0\frac{\partial \varepsilon}{\partial t} \\ &= N^2 q \end{aligned} \tag{2.100}$$

where as before $q = \rho_0 w$ and $w = \partial\varepsilon/\partial t$. This implies:

$$\frac{\partial \rho_e}{\partial t} = \frac{N^2 q}{g} \tag{2.101}$$

Taking the partial time derivative of equation 2.96 and substituting from equation 2.101 we obtain:

$$\begin{aligned} \nabla^2\frac{\partial^2 q}{\partial t^2} &= -g\left[\frac{\partial^2}{\partial x^2}\left(\frac{\partial \rho_e}{\partial t}\right) + \frac{\partial^2}{\partial y^2}\left(\frac{\partial \rho_e}{\partial t}\right)\right] \\ &= -N^2\left[\frac{\partial^2 q}{\partial x^2} + \frac{\partial^2 q}{\partial y^2}\right] \end{aligned} \tag{2.102}$$

This is the wave equation for internal gravity waves. Looking for plane wave solutions of the form

$$A = e^{i(\nu t - \mathbf{k}\cdot\mathbf{r})} \tag{2.103}$$

and substituting equation 2.103 into 2.102, the gravity wave dispersion relation is derived as:

$$(-k_x^2 - k_y^2 - k_z^2)(-\nu^2) = -N^2(-k_x^2 - k_y^2) \tag{2.104}$$

which implies:

$$k_z^2 = (k_x^2 + k_y^2)\left(\frac{N^2}{\nu^2} - 1\right) \tag{2.105}$$

or

$$\frac{k_z}{k_\perp} = \pm\sqrt{\frac{N^2}{\nu^2} - 1} \tag{2.106}$$

where

$$k_\perp = \sqrt{k_x^2 + k_y^2} \tag{2.107}$$

Equation 2.106, the dispersion equation for gravity waves, defines a conical surface of revolution in the ($k_x$, $k_y$, $k_z$) space. Furthermore, it can be shown from ray tracing theory (Lighthill, 1980) that for internal gravity waves, the direction of wave energy propagation is perpendicular to the wave dispersion surface in the ($k_x$, $k_y$, $k_z$) space. This may be illustrated by taking $k_y = 0$ for simplicity, calculating the group velocity $\mathbf{u} = \partial\nu/\partial\mathbf{k}$ and then showing that $\mathbf{u} \cdot \mathbf{k} = 0$. Figure 2.4 shows the dispersion equation surface of revolution in the ($k_x$, $k_y$, $k_z$) space as well as the direction of wave energy propagation.

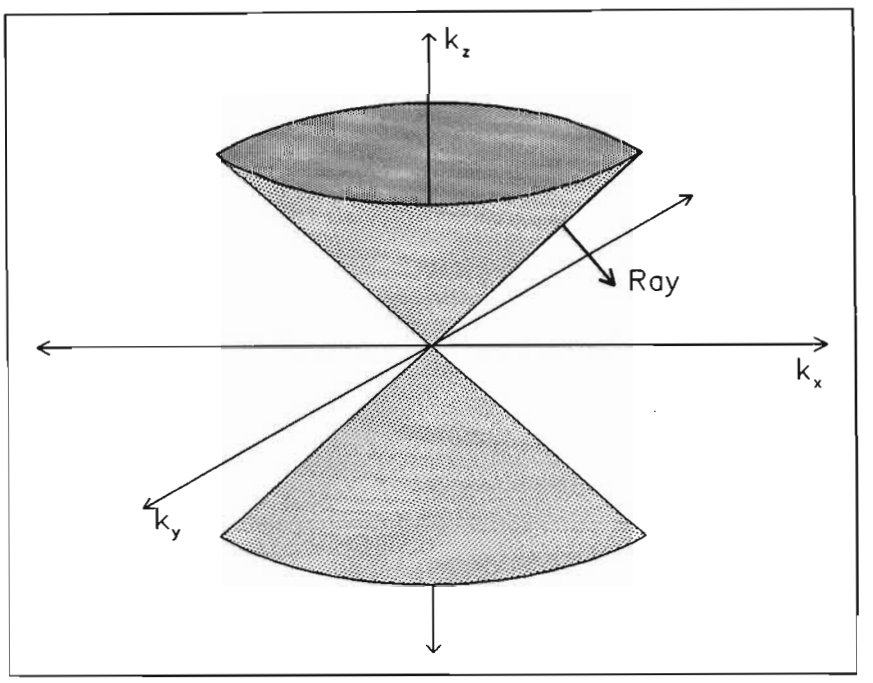


Figure 2.4: The gravity wave dispersion relation diagram. Also shown is the direction of wave energy propagation.

The group velocity vector is parallel to lines of constant phase. Internal gravity waves therefore have the property that group velocity is perpendicular to the direction of phase propagation. Since energy propagates at the group velocity, the wave energy propagates parallel to the wave fronts rather than perpendicular

to them as in the case of acoustic waves. In the atmosphere, internal gravity waves generated in the troposphere by cumulus convection or flow over topography may propagate energy many scale heights into the upper atmosphere, even though individual fluid parcel oscillations may be confined to vertical distances much less than a kilometer (Holton, 1992). Furthermore, it can be shown (Holton, 1992) that the angle of the phase lines to the local vertical depends only on the ratio of the wave frequency to the buoyancy frequency and is independent of wavelength.

### 2.8.3 Rossby waves

The most important wave influencing large scale meteorological processes and the variability of global total column ozone is the Rossby wave, or planetary wave. The influence of these waves on the transport of trace species and on the ozone distribution has been commented on in Chapter 1. In an inviscid fluid, vorticity conservation provides the wave restoring force via the following mechanisms (Holton, 1992):

- In a barotropic fluid, where density depends on pressure alone, the Rossby waves is an absolute vorticity conserving motion that owes its existence to the variation of the Coriolis force with latitude.
- In a baroclinic fluid, where density depends on pressure and temperature, the Rossby wave is a potential vorticity conserving motion that owes its existence to the isentropic gradient of potential vorticity.

Synoptic scale Rossby waves are strongly trapped in the troposphere and decay rapidly with height in the middle atmosphere (Andrews et al., 1987). However, planetary scale Rossby waves do propagate vertically provided their phase speeds are westward (but not too strongly westward) relative to the zonal flow.

#### The $\beta$-plane approximation

Strictly speaking, the solution of the Rossby wave equation is a 3 dimensional problem involving all of the mathematical complexity of the spherical geometry. An approximation (the $\beta$-plane approximation) replaces the complicated spherical

geometry with plane geometry while retaining the underlying dynamics of Rossby wave propagation. The essential characteristics of this approximation are shown in Figure 2.5. We consider the change in $\Omega$ for a small change in latitude ($\delta\phi$)

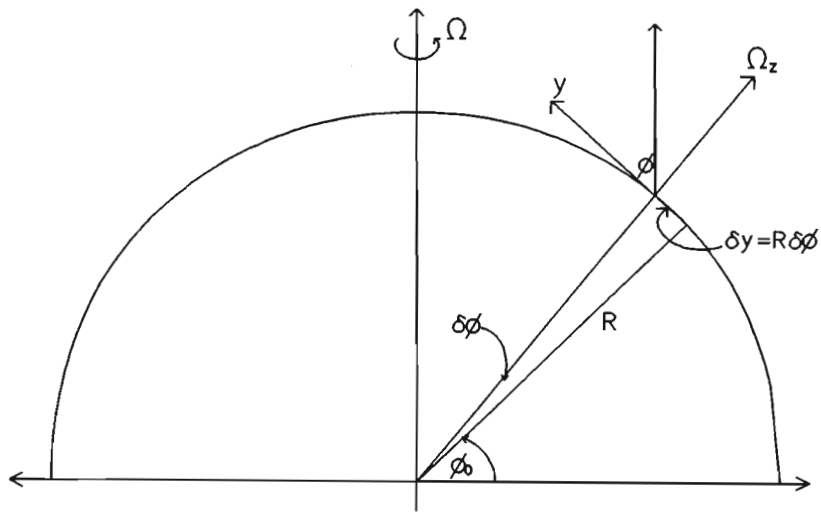


Figure 2.5: The coordinate system for the derivation of the $\beta$-plane equations.

in the (x, y, z) space corresponding to (east, north, zenith). Then:

$$\Omega_z = \Omega \sin(\phi_0 + \delta\phi) \tag{2.108}$$

Expanding as a Taylor polynomial and excluding terms of order $(\delta\phi)^2$ this becomes:

$$\begin{aligned} \Omega_z &= \Omega(\sin\phi_0 + \cos\phi_0\delta\phi) \\ &= \Omega_0 + \frac{\beta}{2}y \end{aligned} \tag{2.109}$$

where $\Omega_0 = \Omega \sin\phi_0$ and

$$\beta = \frac{2\Omega\cos\phi_0}{R} \tag{2.110}$$

Furthermore it is clear that:

$$\beta = \frac{df}{dy} \tag{2.111}$$

the planetary vorticity gradient at $\phi_0$.

### Free barotropic Rossby waves

The dispersion relation for barotropic Rossby waves is derived by considering wave equation solutions of the linearized barotropic vorticity equation (equation

2.75). This equation states that the vertical component of absolute vorticity is conserved following horizontal motion. Making use of the $\beta$-plane approximation, it may be stated as:

$$\left(\frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}\right)\zeta + \beta v = 0 \tag{2.112}$$

Considering a zonal wind only, the linearized parameters can be expressed as:

$$\begin{aligned} u &= u_0 + u_e \\ v &= v_e \\ \zeta &= \zeta_0 + \zeta_e \end{aligned} \tag{2.113}$$

A perturbation streamfunction $\psi_e$ may be defined (Holton, 1992) according to:

$$\begin{aligned} u_e &= -\frac{\partial \psi_e}{\partial y} \\ v_e &= \frac{\partial \psi_e}{\partial x} \end{aligned} \tag{2.114}$$

Substituting into the linearized field equations which are then substituted into equation 2.112 we obtain the linearized barotropic vorticity equation:

$$\left(\frac{\partial}{\partial t} + u_0\frac{\partial}{\partial x}\right)\nabla^2\psi_e + \beta\frac{\partial \psi_e}{\partial x} = 0 \tag{2.115}$$

Assuming wave solutions of the form:

$$\psi_e = Ae^{i(k_x x + k_y y - \nu t)} \tag{2.116}$$

where $k_x$ and $k_y$ are the wavenumbers in the zonal and meridional directions respectively, and substituting into equation 2.115 gives:

$$(-\nu + ku_0)(-k_x^2 - k_y^2) + k_x\beta = 0 \tag{2.117}$$

from which is may be shown:

$$\nu = u_0 k_x - \frac{\beta k_x}{k_x^2 + k_y^2} \tag{2.118}$$

The zonal phase speed relative to the mean wind may then be stated as:

$$c_x - u_0 = -\frac{\beta}{k_x^2 + k_y^2} \tag{2.119}$$

The Rossby wave zonal phase speed is always westward relative to the zonal wind. However, for the Antarctic vortex, the zonal wind is in the eastward direction, and

since the wind speed is generally higher than the phase speed, the waves appear to propagate in the eastward direction. Equation 2.119 also shows that Rossby waves are dispersive due to the inverse square dependence on the horizontal wave number. Phase speeds decrease rapidly with increasing wavenumber (decreasing wavelength). For longer wavelengths the westward Rossby wave phase speed may be equal to the wind speed so that the wave is stationary with respect to the Earth's surface.

Recalling that the group velocity is given by $\partial\nu/\partial k$ equation 2.117 may be used to derive the group velocity:

$$c_{gx} = u_0 - \frac{\beta(k_y^2 - k_x^2)}{(k_x^2 + k_y^2)^2} \tag{2.120}$$

Consider waves which are stationary with respect to the Earth's surface, then setting $c_x = 0$ in equation 2.119 and substituting for $u_0$ in equation 2.120 it can be shown that:

$$c_{gx} = \frac{2\beta k_x^2}{(k_x^2 + k_y^2)^2} \tag{2.121}$$

Which shows that for stationary waves, the group velocity always has an eastward zonal component relative to the Earth. Rossby wave energy is always propagated downstream from its topographic forcing. This is examined in greater detail below.

### Forced topographic Rossby waves

The wave equations derived above were for free barotropic Rossby waves. These waves constitute a natural mode of oscillation of the Earth's atmosphere and although they may be detected by careful observational studies, they generally have small amplitudes. This is because the forcing is quite weak at the large phase speeds characteristic of these waves (Holton, 1992).

More important are topographically forced Rossby waves. It has been mentioned in the previous chapter that hemispheric differences in topographic forcing may account for the differences in wave activity and as a result, differences in the integrity of the Arctic and Antarctic polar vortices.

A dynamical model of topographic Rossby waves again makes use of the barotropic vorticity equation (equation 2.75) for a homogeneous fluid of variable depth (Holton, 1992). We consider the upper boundary of the atmosphere to be fixed

at height $H$ and the lower boundary to be variable, described by $h_T(x,y)$. It is also assumed that the relative vorticity of the geostrophic wind ($\zeta_g$) is very much less than the Coriolis parameter $f_0$ (the quasi-geostrophic approximation). The barotropic vorticity equation can then be written in the form:

$$H\left(\frac{\partial}{\partial t} + \mathbf{V}\cdot\nabla\right)(\zeta_g + f) = -f_0\frac{Dh_T}{Dt} \tag{2.122}$$

Using the linearization method and applying the midlatitude $\beta$-plane approximation, this equation becomes (Holton, 1992):

$$\left(\frac{\partial}{\partial t} + u_0\frac{\partial}{\partial x}\right)\zeta_{eg} + \beta v_{eg} = -\frac{f_0}{H}u_0\frac{\partial h_T}{\partial x} \tag{2.123}$$

where the $eg$ subscript denotes the perturbed field for the geostrophic wind. Simulating mountain ranges with sinusoidally varying topography of the form

$$h_T(x,y) = h_0 e^{ik_x x}\cos k_y y \tag{2.124}$$

and representing the geostrophic wind and vorticity by the perturbation stream function (Holton, 1992)

$$\psi(x,y) = \psi_0 e^{ik_x x}\cos k_y y \tag{2.125}$$

equation 2.123 has a steady-state solution with complex amplitude given by (Holton, 1992):

$$\psi_0 = \frac{f_0 h_0}{H(K^2 - K_s^2)} \tag{2.126}$$

where $K^2 = k_x^2 + k_y^2$ is the horizontal wavenumber and $K_s^2$ is the horizontal wave number required to produce waves with zero phase speed relative to the ground. From equation 2.119 it can be shown that $K_s^2 = \beta/u_0$. Note that if the zonal wind is such that $K^2 = K_s^2$ then the amplitude given by equation 2.126 becomes infinite. This situation in the real atmosphere corresponds to a resonance effect that would usually be damped through various process not discussed here.

### The mechanism for the Rossby wave

The quasi-geostrophic vorticity equation (equation 2.76) shows that the local time rate of change of $\zeta_g$ is given by the advection of the absolute vorticity by the geostrophic wind ($\mathbf{V}_g$) plus the change in vorticity due to stretching or shrinking of the fluid column. Recalling (see equation 2.111) that $\beta$ is the spatial derivative

of the planetary vorticity in the $y$ (northward) direction, the advection of absolute vorticity may be written as:

$$\mathbf{V}_g \cdot \nabla(\zeta_g + f) = \mathbf{V}_g \cdot \nabla\zeta_g + \beta v_g \tag{2.127}$$

The terms on the right of equation 2.127 correspond to the geostrophic advection of relative vorticity and planetary vorticity, respectively. Figure 2.6 shows a theoretical wave-like geostrophic wind as well as the signs of these two terms at different positions along the wave. The figure shows that the two terms are

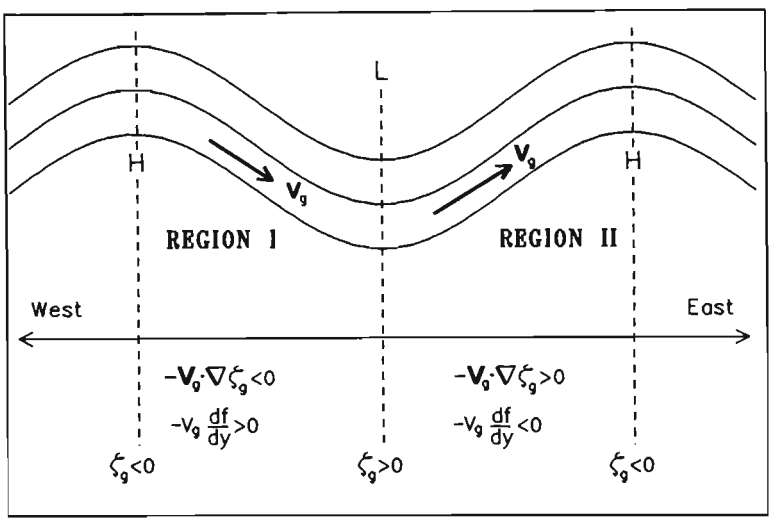


Figure 2.6: Idealized geostrophic wind showing the relative signs of the advection of relative vorticity and planetary vorticity. Adapted from Holton, 1992.

of opposite sign and the total response of the vorticity depends on which type of vorticity advection dominates. It may be shown (Holton, 1992) that the advection of relative vorticity dominates over planetary vorticity for short waves ($k_x < 3000\,\text{km}$), while for long wavelengths ($k_x > 10\,000\,\text{km}$) the planetary vorticity advection dominates. It should be expected that waves of short wavelength will move eastward with the advecting zonal flow while planetary waves of long wavelength will move westward against the mean zonal flow. Waves of intermediate wavelength may be quasi-stationary or move eastward much slower than the geostrophic wind speed.

**Rossby waves in the zonal flow**

To understand the confinement of planetary waves to the winter hemisphere it is necessary to examine the zonal wind conditions under which planetary waves can propagate vertically. For extra-tropical wave activity in the stratosphere, use is made of the midlatitude $\beta$-plane approximation and the quasi-geostrophic potential vorticity equation, which in log pressure coordinates can be written as (Holton, 1992):

$$\left(\frac{\partial}{\partial t} + \mathbf{V}_g \cdot \nabla\right) q = 0 \tag{2.128}$$

where

$$q \equiv \nabla^2\psi + f + \frac{f_0^2}{\rho_0 N^2}\frac{\partial}{\partial z}\left(\rho_0 \frac{\partial \psi}{\partial z}\right) \tag{2.129}$$

$\psi = \Phi/f_0$ is the geostrophic streamfunction and $f_0$ is the equilibrium value of the Coriolis parameter. Using the perturbed fields

$$\begin{aligned} \psi &= -u_0 y + \psi_e \\ q &= q_0 + q_e \end{aligned} \tag{2.130}$$

to linearize equation 2.128, it can be shown that the perturbation field must satisfy (Holton, 1992):

$$\left(\frac{\partial}{\partial t} + u_0\frac{\partial}{\partial x}\right) q_e + \beta\frac{\partial \psi_e}{\partial x} = 0 \tag{2.131}$$

where

$$q_e \equiv \nabla^2\psi_e + \frac{f_0^2}{\rho_0 N^2}\frac{\partial}{\partial z}\left(\rho_0 \frac{\partial \psi_e}{\partial z}\right) \tag{2.132}$$

and $\beta = \partial q_0/\partial y$. We seek wave solutions of the form

$$\psi_e = A(z)e^{i(k_x x + k_y y - k_x ct) + z/2H} \tag{2.133}$$

where the factor $z/2H$ is included to simplify the equation for the vertical dependence (H is the scale height). It can be shown that $A(z)$ must satisfy the wave equation:

$$\frac{d^2A}{dz^2} + m^2 A = 0 \tag{2.134}$$

where

$$m^2 \equiv \frac{N^2}{f_0^2}\left[\frac{\beta}{u_0 - c} - (k_x^2 + k_y^2)\right] - \frac{1}{4H^2} \tag{2.135}$$

where $c$ is the wave phase speed. For vertical propagation it is required that $m^2 > 0$ in which case $m$ is the vertical wave number ($k_z$). For waves of zero

phase speed (stationary waves) it can be seen from equation 2.135 that vertically propagating waves are possible only for zonal flow conditions that satisfy:

$$0 < u_0 < \frac{\beta}{(k_x^2 + k_y^2) + f_0^2/(4N^2H^2)} \equiv U_c \tag{2.136}$$

where $U_c$ is known as the Rossby critical velocity. The vertical propagation of stationary waves can occur only when westerly winds are weaker than a critical value that depends on the horizontal scale of the waves.

## 2.9 Diagnostics of wave activity

### 2.9.1 Eliassen-Palm flux

Atmospheric transport models that exclude wave transience and dissipation tend to considerably overestimate the strength of the zonal wind. The reason for this is that these two processes provide mechanisms for the deceleration of the zonal wind. Assuming that wave transience and dissipation can be neglected in the thermodynamic (but not the zonal momentum) equation, a set of prognostic equations may be derived for atmospheric transport (Brasseur and Solomon, 1986). Scale analysis shows that the dominant term in the stratosphere in the zonal wind momentum equation is the gradient of the Eliassen-Palm flux ($\mathbf{F}$) which has quasi-geostrophically scaled components of:

$$F_\phi = \rho_s a \cos\phi \, \overline{u_e v_e} \tag{2.137}$$

and

$$F_z = \rho_s a \cos\phi f \frac{\overline{v_e \theta_e}}{\overline{\theta_z}} \tag{2.138}$$

where the overbar denotes zonal means and the $e$ subscript denotes eddy quantities. The Eliassen-Palm flux defined by these two equations is identical to the quasi-geostrophic potential vorticity flux due to planetary waves which can be derived from the perturbation form of equation 2.76. $\nabla \cdot \mathbf{F}$ represents the momentum forcing by eddies, and has the important property that it vanishes for steady, conservative waves (Brasseur and Solomon, 1986). However, if the eddies are completely steady and dissipationless, then $\nabla \cdot \mathbf{F} = 0$ and the net mean meridional velocity must be zero. The strength of the mean meridional circulation in a rotating frame of reference depends crucially on the ability of transient, dissipating eddies to act as a momentum sink (or source) to balance the Coriolis torque produced when parcels move in the meridional direction.

### 2.9.2 Wave drag

The zonally averaged momentum budget must be balanced by the convergence or divergence of momentum flux due to wave perturbations. This momentum deposition may be described in terms of wave drag. Wave drag is a macroscopic process that can be brought about by transience (growth and decay of the wave amplitude) or by dissipation (thermal damping and turbulent diffusion). The diffusion that is referred to in wave drag is a rough parameterization of far more complex, non-linear processes that result in a decrease in the spatial and temporal scales of motion. The origin and magnitude of wave drag in the stratosphere is an important question for atmospheric dynamics, but no simple parameterization of wave drag in the stratosphere has yet been proposed.

While the transformed Eulerian circulation provides a convenient framework for a qualitative understanding of transport in the middle atmosphere, the question of the role of transient, dissipating eddies in determining the distributions of the chemical species must also be carefully considered (Brasseur and Solomon, 1986). To understand the effects of these processes on chemical constituents, it is necessary to know what kind of waves are involved and the physical mechanisms responsible for their transience and dissipation. These are presently not well known in the stratosphere.

### 2.9.3 Rayleigh friction

Modelling studies mostly use a 'Rayleigh' or linear friction coefficient to parameterize the effects of wave drag in the zonal momentum equation. A deceleration coefficient $K_r$ is assumed to act linearly on the zonal wind and is used in the zonal momentum equation as follows (Brasseur and Solomon, 1986):

$$\frac{\partial u_0}{\partial t} - f v_0^* = -K_r u_0 \tag{2.139}$$

where $K_r$ is the Rayleigh friction coefficient and $v_0^*$ is the transformed zonal mean meridional wind component.

### 2.9.4 K-Theory

Early models of stratospheric motion neglected eddy transport processes and found that discrepancies existed in the angular momentum budget of the derived circulation. Reed and German (1965) provided the formalism required to include eddy transport of heat. The circulation could then be described in terms of 2-dimensional turbulent diffusion coefficients of the form:

$$\overline{v_e\theta_e} = -\left(K_{yy}\frac{\partial\overline{\theta}}{\partial y} + K_{yz}\frac{\partial\overline{\theta}}{\partial z}\right) - \left(K_{zy}\frac{\partial\overline{\theta}}{\partial y} + K_{zz}\frac{\partial\overline{\theta}}{\partial z}\right) \qquad (2.140)$$

where $K_{yy}$, $K_{zz}$, $K_{yz}$ and $K_{zy}$ represent the eddy diffusion coefficients in the $y$, $z$ and off-diagonal directions. This description assumes that the eddy transport of heat occurs via a diffusive process (if $K_{yz} = K_{zy}$), and that some particular mixing length can be assumed to represent the distance travelled by a parcel in the eddy before it mixes irreversibly with its surroundings (Brasseur and Solomon, 1986). It was also suggested that a similar form of equation 2.140 could be adopted for the eddy transport of chemical constituents by replacing potential temperature with the mixing ratio of the species. It was found, however, that the numerical value of the K fields had to be adjusted arbitrarily for many of the two dimensional models to obtain agreement with measured values.

# Chapter 3

# THE TOMS/SBUV EXPERIMENT

## 3.1 Introduction

The analyses presented in Chapters 4, 5 and 6 of this thesis make extensive use of data obtained from the TOMS experiment carried on the American Nimbus-7 satellite. This experiment was in operation from 20 October 1978 to 6 May 1993, providing daily global maps of total column ozone over the entire surface of the Earth outside the regions of polar night. One of seven other instruments on Nimbus-7 together with TOMS, was the SBUV (Solar Backscatter Ultra Violet) experiment, sharing the TOMS front end optics and some of the supporting electronics. This instrument is able to provide coarse vertical ozone profiles over the altitude range of 25 to 50 km. The Nimbus-7 SBUV was in operation continuously from October 1978 until 13 February 1987 when the dark current chopper went out of synchronization with the measuring optics. Some studies (Wellemeyer et al., 1993) were able to correct for this non-synchronization, extending the data set to June 1990 after which such corrections were no longer admissible.

These two ozone measuring experiments on Nimbus-7 operated in some respects as a single instrument and for this reason both systems will be discussed in this chapter. Both instruments use differential absorption spectroscopy to make remote measurements of ozone. Many of the aspects of the measurement technique are similar to those employed in ground-based spectroscopy. The importance of dark current correction, stray light rejection, accurate wavelength determina-

tion, narrow instrument bandpass etc. are common to both satellite and ground-based spectroscopy. However, many more restrictions such as power consumption, weight, reliability and resilience to the space environment are placed on satellite based instruments. Some of these restrictions severely limit the performance characteristics eg. limited use of cryogenic cooling due to weight and power consumption considerations. In diode array spectroscopy, cryogenic cooling is necessary to suppress dark current resulting from thermal noise. In the discussion of the hardware of the TOMS/SBUV experiment, it will be seen how these limitations have influenced its design.

The TOMS/SBUV is an improved version of the BUV instrument launched on the Nimbus-4 satellite in 1970 (Heath et al., 1975). The BUV instrument sounded ozone only in the nadir direction, ie. looking downward towards the center of the Earth. In addition, under certain orbital conditions, the data were influenced by charged particles in the Earth's radiation belts. The SBUV/TOMS instrument operates over an extended wavelength range, and makes special provision for minimizing the effects of space radiation. The experiment measures the vertical ozone profile, the solar ultraviolet spectrum, and provides a total column ozone map by means of a mechanical scan across the satellite track. The vertical ozone profile is obtained by measuring the backscattered solar radiation in the UV absorption band from 255 to 340 nm using a double monochromator operated in a stepped wavelength scan mode.

A second TOMS instrument was launched on the Russian Meteor-3 satellite on 15 August 1991. However, differences in the satellite orbit parameters prevent Meteor-3 TOMS from making its measurements in the same way as Nimbus-7 TOMS. These differences and their affect on Nimbus-7 and Meteor-3 TOMS data comparisons are discussed further in sections 3.2 and 3.7. In addition to the single TOMS instrument currently in orbit, a number of future TOMS experiments are planned as part of the EOS (Earth Observing System) programme. Instruments are scheduled for launch on Earth Probe-94 in the middle of 1994, the Japanese ADEOS satellite early in 1996 and on Earth Probe-98 in the middle of 1998. With the apparent loss of the NOAA-9 satellite and the SBUV/2 instrument on board, Meteor-3 TOMS is the only instrument providing global ozone coverage until the launch of Earth Probe in 1994.

Once the TOMS/SBUV instrument has measured the required spectra, the data must be extensively processed before total column ozone and vertical ozone pro-

files may be extracted. The data retrieval algorithms for these experiments will be described briefly, including a discussion of their current shortfalls. After the raw data have been reduced, the resultant data set is made available to the user in a gridded format. A discussion of this final product as well as an examination of the TOMS data quality is given below. The primary means of quality control for the TOMS data set is comparison with a network of 39 ground-based Dobson spectrophotometers. Comparisons were made throughout the lifetime of Nimbus-7 TOMS and some of these are discussed in section 3.5.

A number of different versions of the TOMS data have been released by NASA. New versions follow from older versions due to improvements in the TOMS data reduction algorithm and corrections for instrument drifts. Versions 1 to 3 were interim data sets designated during the original development of the overall data reduction algorithm (Herman et al., 1991). Version 4 (V4) was the first operational algorithm used for providing global total column ozone data. The most recent is the version 6 (V6) grid TOMS product discussed in greater detail in section 3.6. Towards the end of 1994 it is planned to reprocess the Nimbus-7 TOMS data with major changes to the reduction algorithm and in-flight calibration determination. This version 7 (V7) product is expected to contain small but significant changes in total column ozone values. Finally a discussion of possible differences between Meteor-3 TOMS data and Nimbus-7 TOMS data is given in section 3.7.

## 3.2 The Nimbus-7 and Meteor-3 satellites

The orbit and performance parameters of the Nimbus-7 TOMS and Meteor-3 TOMS are summarized in Table 3.1. The most important difference between the two satellites is in their precession periods. The orbit of the Nimbus-7 satellite was sun-synchronous such that its equator transit time was always near local noon. It drifted from local noon at the time of launch to 10:45 a.m. in 1993. The result was that the Nimbus-7 TOMS instrument saw the entire globe at least once a day. Because of the long precession period of the Meteor-3 orbit, the equator crossing time goes from noon (solar zenith angle = 0°) to the day-night terminator (solar zenith angle = 90°). Data quality is best when the equator crossing time of the ascending node of the orbit is between 10 a.m. and 2 p.m. or between 10 p.m. and 2 a.m. (which indicates that the equator crossing time of the descending

| Parameter | Nimbus-7 | Meteor-3 |
|---|---|---|
| Altitude | 955 km | 1020 km |
| Orbit inclination | 81° | 82.5° |
| Orbit period | 104.2 min | 109 min |
| Orbit eccentricity | N/A | $< 2 \times 10^{-2}$ |
| Orbit precession | 22.5 hours | 212 days |
| Launch Date | 24 October 1978 | 15 August 1991 |
| Launch Time | N/A | 12:15 Moscow standard time |
| First ozone data | N/A | 22 August 1991 |
| No. orbits per day | 13.82 | 13.21 |
| Field of view | $3° \times 3°$ | $3° \times 3°$ |
| Wavelengths | 312.5, 317.5, 331.2, 339.8, 360.0, 380.0 nm | 312.35, 317.4, 331.13, 339.73, 359.0, 380.16 nm |
| Bandwidth | 1 nm | 1.1 nm |
| Cross-track scan | $\pm$ 51° | $\pm$ 51° |
| No. views per scan | 35 | 35 |
| Nadir field of view | 47 km $\times$ 47 km | 64 km $\times$ 64 km |
| Extreme off nadir | 250 km $\times$ 250 km | 260 km $\times$ 70 km |

Table 3.1: **Orbit and performance parameters for Nimbus-7 TOMS and Meteor-3 TOMS. Where no information is available, entries have been noted (N/A).**

node of the orbit is between 10 a.m. and 2 p.m.). The precessing nature of the Meteor-3 orbit also complicates the calibration of the TOMS instrument on board due to changing view angles and changing temperatures. A schematic diagram of the mode of operation of the Nimbus-7 TOMS experiment is given in Figure 3.1.

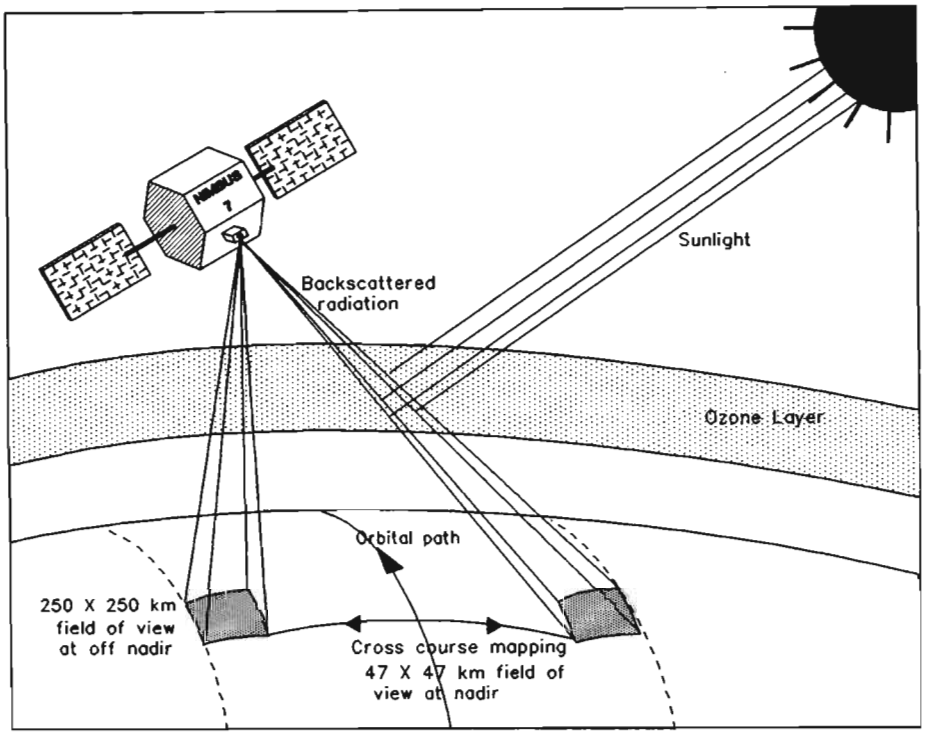


Figure 3.1: The Nimbus-7 satellite and the mode of operation of the TOMS experiment.

## 3.3 The TOMS and SBUV measurement technique

### 3.3.1 The geometry of the satellite-based measurements

Total column ozone was measured by Nimbus-7 TOMS by means of a cross course mapping UV spectrometer which swept at right angles to the direction of spacecraft motion with a 3° by 3° instantaneous field of view (Heath et al., 1975) as shown in Figure 3.1. It made 200 000 total ozone measurements daily in a 47 km × 47 km field of view at nadir and 250 km × 250 km at extreme off-nadir. Of the 6 wavelengths of backscattered UV terrestrial irradiance measured by TOMS (see Table 3.1) only the four shorter bands are sensitive to atmospheric ozone; the two longer wavelength bands are used for estimating the reflectivity necessary for deriving total ozone (Fleig et al., 1986 a). The TOMS irradiances are sampled in 3° steps ±51° from nadir across the ground track, providing 35 samples every 7.95 seconds (Heath et al., 1975). The extreme cross-scan positions (1 and 35) are used only in the equatorial and very high latitude polar regions to provide

complete spatial coverage. Away from the equator and polar regions, there is enough overlap between the adjacent orbits that only the central scans need to be used (Herman et al., 1993).

This mode of operation is similar for both Meteor-3 TOMS and Nimbus-7 TOMS. With the 104 minute Nimbus-7 orbital period (109 minutes for Meteor-3), the 8 second scan cycle and 1 second retrace, means that successive cross course scan lines are displaced about 50 km along the orbital track. Due to the Earth's rotation, each sun-synchronous orbit of the Nimbus-7 satellite was located approximately 26° of longitude west of the preceding orbit. The resultant scan geometry provided total Earth coverage more than once per day.

The SBUV instrument is a nadir viewing scanning double monochromator designed to measure total ozone and the vertical ozone profile over the pressure (altitude) range 0.7 hPa to 30 hPa (25 km to 50 km). The raw data are obtained by accurate measurement of the atmospheric albedo at 12 discrete wavelengths from 255.5 to 339.8 nm. Wavelengths from 312 to 340 nm are used to calculate total ozone, while wavelengths from 273 to 312 nm are used to infer the ozone profile. Furthermore, it makes periodic scans of wavelengths continuously from 160 nm to 400 nm, so that the UV flux and the Earth radiance measurements are available as a by-product of the SBUV measurement. The Nimbus-7 experiment performed about 1300 measurements between 81°N and 81°S, to infer total ozone and the ozone profile along the orbital track each day in a 200 km × 200 km field of view. The resulting data set was one of the few sources of global vertical ozone profiles.

### 3.3.2 The TOMS subsystem

The TOMS subsystem operates as a step scanner to provide statistical independence of the samples and the same field of view on the ground for each wavelength, as the wavelengths must be scanned sequentially using a single detector to preserve the radiance measurements (Heath et al., 1975). During any given measurement, the motion of the satellite causes the area being viewed to change by about 3%, resulting in small errors.

A schematic view of the TOMS subsystem optics is shown in Figure 3.2. The single Ebert-Fastie monochromator with a fixed grating and an array of exit

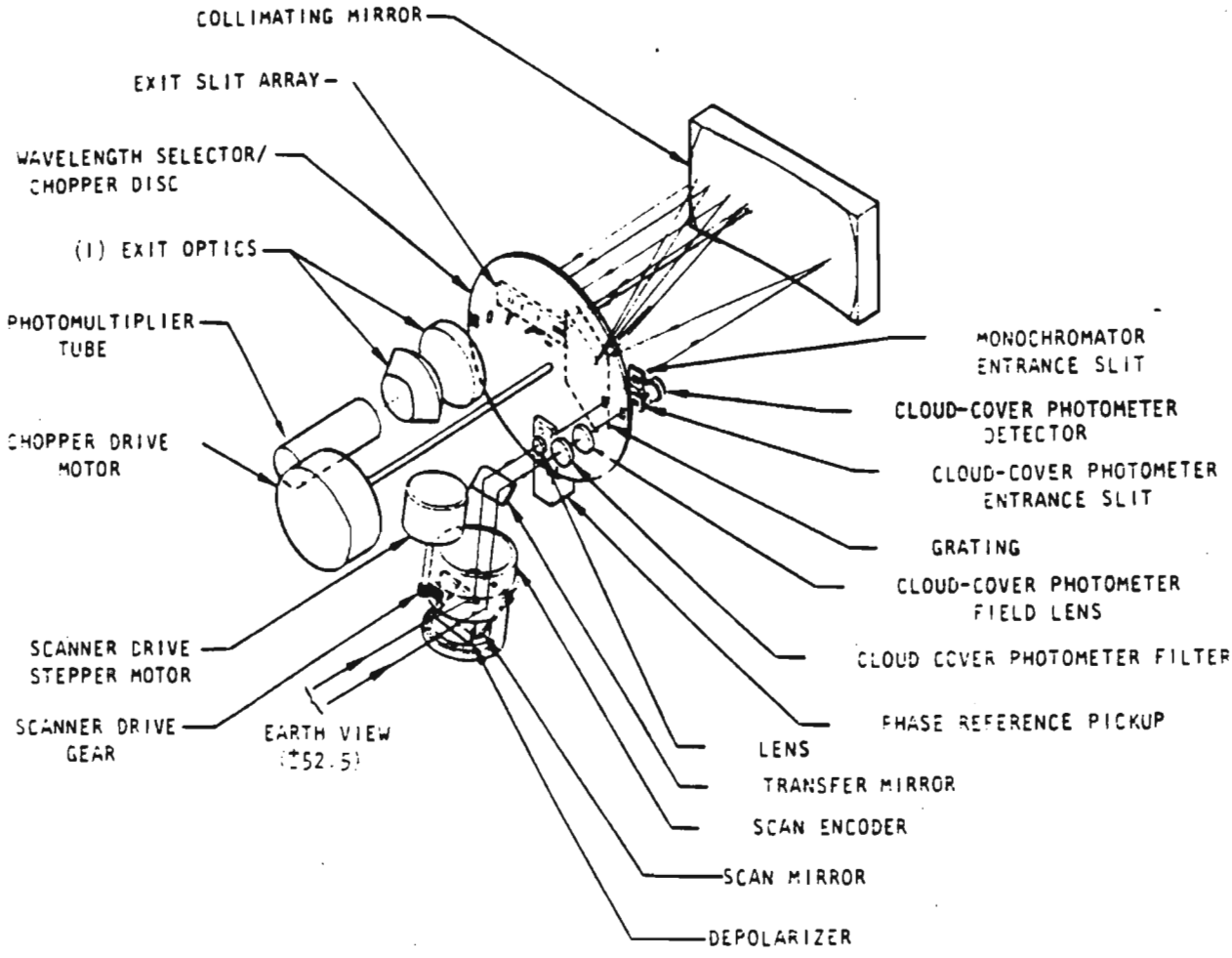


Figure 3.2: The TOMS subsystem optics. From Heath et al. (1975)

slits, is very similar to one half of the SBUV double monochromator discussed below. A rotating wavelength selector is used to gate the dispersed light from the desired exit slit to the detector. The cloud cover photometer is separate from the monochromator, but uses the same chopper disc and sees the same field of view as the monochromator. Incident light passes through a photometer filter having a 50 nm spectral bandpass at 800 nm and a field defining lens. The signal produced by the photomultiplier tube (PMT) is processed by a feedback current digitizing electrometer.

### 3.3.3 The SBUV subsystem

The SBUV subsystem is based on the double monochromator system that was used in the BUV instrument. This instrument was a tandem Ebert-Fastie spec-

trometer, measuring spectral intensities at 12 wavelengths ranging from 255.5 to 339.8 nm with a 1 nm bandpass. The SBUV subsystem is similar in design but features a monochromator which maximizes transmittance, provides accurate wavelength resolution and very high stray light rejection. Furthermore several additional features have been designed into the SBUV instrument to improve performance and to accommodate the continuous scan mode and the wider wavelength range (160 to 400 nm vs. 255 to 340 nm).

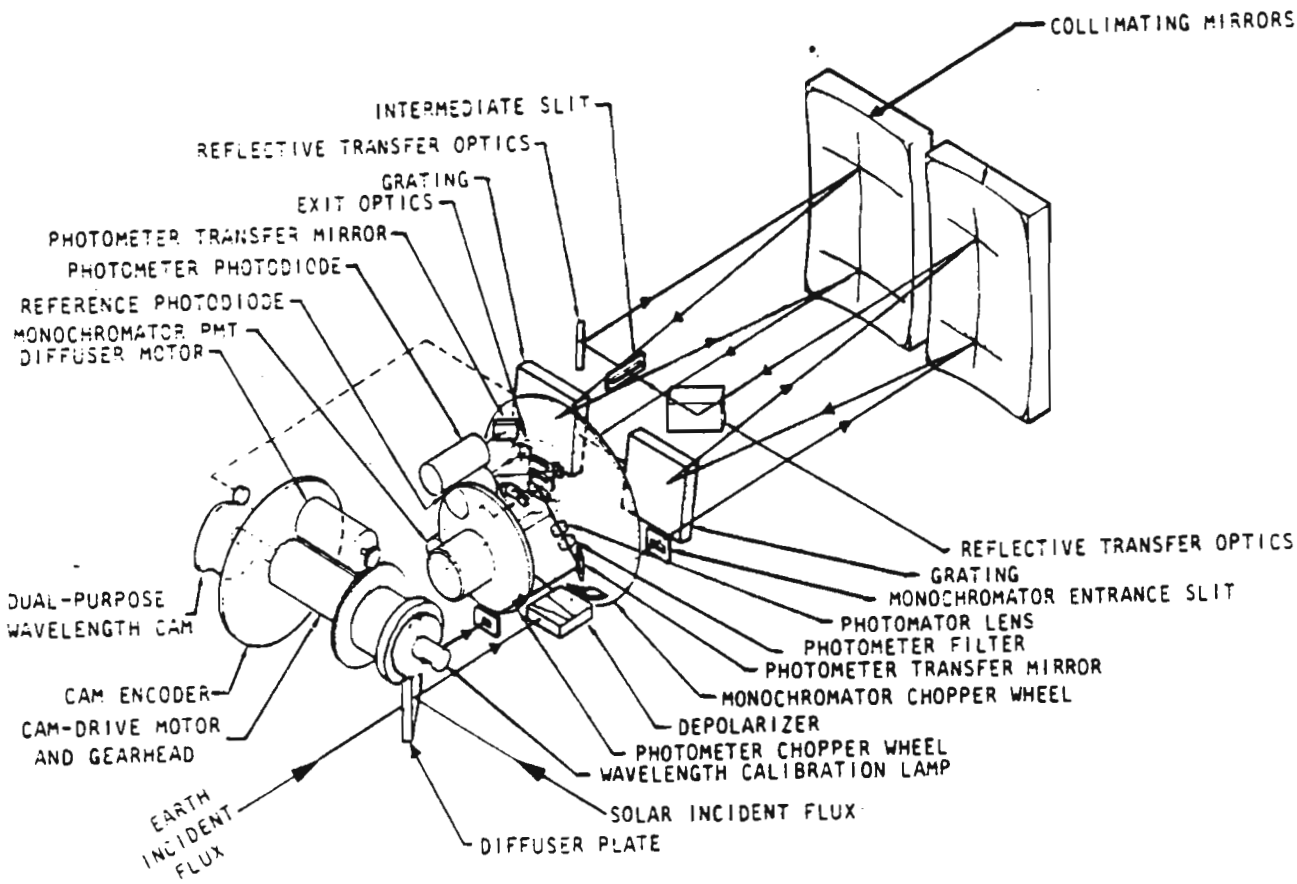


Figure 3.3: The SBUV subsystem optics. From Heath et al. (1975).

Reflected sunlight enters the monochromator optical system through the entrance slit. It is necessary to subtract detector dark current (as well as electronic and photometric offsets in general) from the SBUV/TOMS data since these offsets are much greater than the desired noise level. Furthermore, the space radiation belts tend to cause rapid and large modulation of the dark current signal. This problem is alleviated by chopping the incoming sunlight with a 50 Hz chopper wheel and measuring the background current during the dark time. This dark current is then subtracted from the next measurement. The chopped optical signal from

the monochromator illuminates the cathode of a PMT, the output of which is amplified by a range-switched electrometer. This voltage is then converted by a voltage to frequency converter (VFC) to a digital pulse train whose frequency is proportional to the electrometer output. The electronics subsystem then processes this signal. The 339.8 nm filter photometer, which operates on the high signal portion of the SBUV wavelength range, uses a vacuum photodiode instead of a PMT. The reference photodiode channel signal (339.8 nm), used to monitor the PMT gain, is processed in the same manner as that of the monochromator channel.

The SBUV monochromator is a double Ebert-Fastie type with fixed entrance and exit slit widths equivalent to a 1 nm spectral bandwidth near 300 nm. Each Ebert-Fastie monochromator of the SBUV double monochromator uses a single spherical collimating mirror with a radius of curvature of 500 mm. These mirrors have high surface finishes for low light scatter and are aluminized and overcoated with $MgF_2$ for 80% reflectance at 160 nm.

To obtain the ozone profile from the measured UV spectra, it is necessary to calculate the atmospheric albedo to within 3%, which requires measurements of the radiance to 1% or better. To achieve this level of accuracy, it is necessary that the bandpass, wavelength accuracy, and repeatability of the spectrometer measurement contribute only a small part of this error. At 305 nm, the radiance increases at a rate of two decades per 12 nm for a typical set of data. To obtain an error of less than 1%, the value of $\Delta\lambda$ must be less than 0.026 nm (Heath et al., 1975). If the instrument bandpass is too wide, small absorption features within the measured spectra are smoothed with loss of information.

## 3.4 The TOMS and SBUV data retrieval algorithms

With each improvement to the TOMS data retrieval algorithm, a new version of the data is released. Part of the front end optics of the SBUV/TOMS instrument is a diffuser plate, common to both subsystems, which is directly exposed to space. Outgassing from the satellite, bombardment by space debris and exposure to solar UV radiation causes a steady reduction in the diffuser plate reflectivity (Herman et al., 1990). This results in an over-estimate of the severity of ozone depletion

in the stratosphere. The calibration adjustment made to correct this problem for the Nimbus-7 TOMS/SBUV was the result of an internal method based on the fact that different wavelengths used to measure total column ozone have different sensitivity to calibration error. The long-term drift between total ozone measured at two different wavelength pairs (see below) was used together with this assumption to determine the relative calibration error in each pair (McPeters et al., 1990). By assuming that the sea surface reflectivity is stable in the long term, the less critical absolute calibration of the longer reflectivity wavelengths could be inferred. This approach is known as the Pair Justification Method (PJM) and provides a long-term calibration for TOMS that is accurate to about 1% in total ozone over a period of a decade. This technique was used in the updating of V5 data to V6 data. A brief description of the PJM and the TOMS data reduction algorithm is given in this section.

Details of the irradiance measurements made by TOMS are given by Cebula et al. (1988) while the basic ozone retrieval algorithm is described in detail by Klenk et al. (1982). Assuming that the atmosphere does not scatter, and that the ozone layer is confined to a narrow altitude range in the stratosphere, the measured albedo (A) is given by (Hudson and Kim, 1993):

$$A = R \exp(-\alpha\Omega s - \beta\Delta s) \tag{3.1}$$

where $R$ is the reflectivity of the ground, $\alpha$ is the ozone absorption cross-section, $\Omega$ is the total column ozone, $s$ is the path length determined geometrically, $\beta$ is the Rayleigh scattering coefficient for the atmosphere and $\Delta$ is the column content of the atmosphere. Hence measurements of the ratio of the radiance scattered from the atmosphere and reflected by the Earth's surface to the solar irradiance (backscattered directional albedo) are the primary data for determining total column ozone (Herman et al., 1991). The four shorter wavelengths at which TOMS makes measurements (listed in Table 3.1) are combined into pairs. Ratios of the irradiances for each pair provide parameters known as N-values. The wavelength pairs used to calculate the N-values are listed in Table 3.2 (Herman et al.,1991). The TOMS data retrieval algorithm requires an a priori estimate of the vertical ozone and temperature profile to make a total column ozone measurement. Climatological ozone and temperature profiles for various latitudes are used to initialize a radiative transfer calculation (Dave and Mateer, 1967) to construct a table of precomputed radiances (expressed as theoretical N-values) against which measured backscattered UV radiances are compared (Wellemeyer et al., 1993).

| Pair (nm) | Name | Separation (nm) |
|---|---|---|
| 312.5/331.2 | A | 18.8 |
| 317.5/331.2 | B | 13.8 |
| 331.2/339.8 | C | 8.5 |
| 312.5/317.5 | A′ | 5.0 |
| 317.5/339.8 | B′ | 22.3 |

Table 3.2: **The wavelength pairs used in the TOMS data reduction algorithm.**

The values in the table span the expected range of total column ozone, solar zenith angle (used to calculate the optical slant path length), surface pressure, and latitude dependent ozone profile (Herman et al., 1991). The vertical temperature profile is required since ozone absorption cross-sections are temperature dependent. Measured N-values can then be converted into total column ozone amounts by interpolation between adjacent values in the tables of theoretical N-values. Data quality may be monitored by requiring that calculations from each wavelength pair yield the same total column ozone. There remain sources of error in the TOMS data reduction algorithm and these are discussed in section 3.4.1.

Based on the need to minimize the errors in the reduction algorithm, the A and B′ pairs are selected for derivation of total column ozone for most of the globe (Herman et al., 1991). For the large solar zenith angles encountered in the polar regions, only the C pair penetrates to the lower boundary to include all of the ozone in the atmosphere. The original method for total ozone retrieval from the backscattered radiances of the SBUV was based on a technique proposed by Dave and Mateer (1967). The reduction of an ozone profile from the set of measured albedos is made by using a partial derivative inversion algorithm (Klenk et al., 1983) that incorporates the optimum statistical method of Rodgers (1976). The inversion algorithm also uses climatological profiles and an associated covariance matrix containing layer standard deviations and correlations between layers. Total ozone calculated from TOMS represents a further constraint on the resultant vertical ozone profile.

### 3.4.1 Errors in TOMS data retrieval

In discussing the errors inherent in the TOMS data retrieval algorithm, an attempt has been made to address the following questions:

1. What is the source of the error?
2. When/where does it occur?
3. What are its effects?
4. How serious is it?
5. How can the error be avoided?

1) Profile shape errors

Differences between the assumed climatological ozone and temperature profiles and the true profiles lead to errors in the estimated total column ozone. For example, seasonal biomass burning over equatorial regions causes significant perturbations to tropospheric ozone (Fishman et al., 1990) leading to discrepancies between the shape of the tropospheric climatological ozone profile and the true ozone profile.

The error occurs predominantly at solar zenith angles greater than 70° where long optical path lengths reduce penetration of the radiation to the ground (Herman et al., 1991). Since TOMS measurements are made at local noon, large solar zenith angles correspond to high latitudes during the winter months where the error maximizes.

Long-term changes in the vertical ozone profile could lead to errors in the trends derived from the TOMS data. Over the first 11 years of the Nimbus-7 TOMS lifetime, it appears that global average stratospheric ozone decreased by 2.9 $\pm$ 1.3% (Herman et al., 1991) while tropospheric ozone increased (Fishman et al., 1990). This would result in a steady increase in the discrepancy between the climatological ozone profile and the true ozone profile. The profile shape errors occur as over-estimates of total column ozone trends associated with high altitude changes in the profile, and under-estimates of total column ozone trends associated with the tropospheric profile changes (Wellemeyer et al., 1993).

Using simultaneous data sets obtained from ozonesondes, SBUV and TOMS, Wellemeyer et al. (1993) indicated a standard deviation of approximately 10% in

the retrieved total ozone at solar zenith angles of 86° and higher. They were also able to estimate the errors in the long-term ozone trends derived from TOMS at high latitudes. Their sensitivity calculations indicated that the derived total column ozone was not sensitive to differences in climatological and actual vertical ozone profiles close to the ozone maximum. The study further indicated that long-term trends derived from TOMS data at 63°N would contain a seasonally dependent systematic error leading to an over-estimate in long-term total column ozone decreases by 2 to 4% per decade in the winter months. The error in long-term ozone trends derived from TOMS at the highest solar zenith angles, associated with long-term changes in the ozone profile measured by SBUV, is estimated as a 5% per decade over estimation. If trend studies using the V6 (see later) grid TOMS product are restricted to 60°N to 60°S, the maximum error in derived trends is about 2% per decade.

At high latitudes, profile shape errors can be avoided by making use of measured vertical ozone profiles rather than climatological profiles. However, this approach is severely limited by data availability.

2) Climatological cloud top height errors

The TOMS algorithm is designed to compute total column ozone above the Earth's surface, although only the ozone above dense (optical depth greater than one) cloud tops and mid-tropospheric levels can be measured from space. As a result, tropospheric ozone is treated implicitly in the radiative code used to construct the N-value tables, while the hidden ozone below cloud tops is estimated from climatological cloud top data (550 hPa at high latitudes) and average tropospheric ozone mixing ratios. Hence errors in the cloud altitude estimates result in total ozone retrieval errors.

The error causes an over-estimate in total column ozone in regions of persistent low cloud, usually over oceans, and an under-estimate in regions of persistent high cloud, usually orographic clouds forming over large mountain ranges. During the TRACE-A/SAFARI campaign (August to November 1992) it was found that persistent marine stratocumulus clouds off the southwest coast of Africa caused an over-estimate in TOMS total column ozone by 6 to 7 DU (Thompson et al., 1994).

This problem may be corrected by using cloud cover data to obtain true cloud top heights. Data availability again limits the feasibility of the error correction.

3) Surface reflectivity errors

Studies by Hudson and Kim (1993) showed that over much of tropical and sub-tropical Southern Africa TOMS may under estimate total column ozone by 10 to 20 DU. Possible reasons are (Thompson et al., 1994):

- The ozone retrieval algorithm under-estimates ozone below 500 hPa if the reflectivity of the surface is low.
- Biomass burning in this region may cause large increases in tropospheric ozone (Fishman et al., 1990) and the standard climatological profiles used in the TOMS reduction algorithm do not account for this, as mentioned above.

4) Aerosol loading induced errors

The presence of small amounts of stratospheric aerosol (typically from volcanic eruptions) gives rise to a small wavelength dependent scattering term that is not accounted for when calculating the N-value tables (Herman et al., 1991). This omission leads to a small error in the derived total column ozone under most conditions (Dave, 1978) and may be corrected by using pairs of closely spaced wavelengths when calculating the N-value tables.

Herman et al. (1993) have shown that errors caused by the short-term presence of stratospheric aerosols in the TOMS zonally averaged ozone data are less than 1% before correction, and have no significant effect on ozone trend determination. However, the presence of volcanic aerosols can have other effects on TOMS ozone retrievals as discussed below.

5) Sea glint errors

Sea glint may occur in the irradiance measurements for nadir and near nadir scan positions whenever the angle between the sun and the viewing direction results in the reflection of the sun from the ocean falling within the field of view of the instrument. Increasing zenith angle for a given latitude reduces the occurrence of sea glint, introducing a systematic error in the 380 nm boundary reflectivity determination (Herman et al., 1991). During 1982 to early 1983, the reflectivity determination was affected by stratospheric and tropospheric aerosol loading from the El Chichón volcanic eruption. Cases of sea glint may be identified and removed from the data set as part of the TOMS data reduction algorithm.

6) Synchronization errors

From 1985 onwards there was a small synchronization error between the timing of slits passing in front of the spectrometer entrance aperture and the photon counting electronics (Fleig et al., 1986 a). The location of slits in the chopper wheel relative to the spectrometer entrance aperture are synchronized with the photon counting electronics so that the dark current is automatically subtracted from the total signal for each wavelength as described in section 3.3.3. The instrument and software are designed to reject data where a preset threshold for a 'nonsync' condition is detected. By September 1990 the instrument had stabilized such that 20% of cross track scans had at least one observation affected by 'out-of-sync' errors. Although this problem does not have serious consequences for short term studies of the ozone hole formation, it does have an impact on long-term trends derived from TOMS, since accuracy of 1% or better is required. In the worst case the effect on an individual ozone value can be up to 10%; cases where the error could be larger are rejected in the ozone derivation algorithm. Since only a very small number of data are rejected, nonsync errors represent a minor problem for TOMS data reduction (Herman et al., 1991).

However, the synchronization error results in a related problem referred to as 'toggling'. This is the apparent rapid switching of the TOMS instrument between two fairly well defined synchronization states (Herman et al., 1991). Toggling is the major source of error in the TOMS data after application of the PJM. For the wavelength pairs listed in Table 3.2 the ozone uncertainties arising from toggling are 3 DU for (A, B'), 5 DU for (A, B) and 7 DU for (A, A') (Herman et al., 1991). Although this problem could not be eliminated, it was minimized by selecting the wavelength pairs with approximately the same sensitivity to toggling in the data reduction algorithm.

7) Satellite attitude errors

Errors in the attitude determination of the Nimbus-7 satellite led to small errors in total column ozone retrievals. A 1% disagreement was found between measurements from the extreme off-nadir scan positions, a cross-track bias between positions 1 and 35, or 'far left' and 'far right' (Herman et al., 1991). The largest error corresponded to a satellite roll determination error of 0.1° that translated into an ozone cross-track bias of 2 DU.

8) PSC induced errors

As discussed in point 2, high altitude clouds can cause an under-estimate in

measured total column ozone. In addition to cloud formation over mountain ranges, very high altitude clouds also form over the Antarctic, causing TOMS total column ozone retrievals to be under estimated (Herman et al., 1991).

Polar stratospheric clouds (PSCs) forming near or above the ozone maximum result in measurement errors for solar zenith angles greater than 70°(Torres et al., 1992). Mie scattering model results indicate that NAT type I PSCs (see Chapter 1) produce an approximate 2% under estimation of total column ozone at solar zenith angles greater than 80°. Type II PSCs, which occur less frequently, produce under-estimates of up to 6%. Using ozonesonde data from Palmer station on the Antarctic Peninsula, Krueger et al. (1988) showed that a dense stratospheric cloud at 100 hPa, but assumed by the algorithm to be at 550 hPa, could produce an error in total column ozone of up to 50 DU. This error diminishes at high solar zenith angles where less UV light penetrates to tropospheric levels.

## 3.5 TOMS and SBUV data quality

In any discussion on the use of TOMS data, the aspect of data quality and instrument calibration must be addressed. TOMS data quality is assured by means of comparison with a network of ground-based Dobson spectrophotometers. Discrepancies between the two data sets reveal problems inherent in the TOMS instrument or the data reduction algorithm. Corrections to the data usually result from improved calculations of ozone absorption cross-sections, a better climatology of equatorial ozone, and improved knowledge of the change in instrument characteristics with time eg. the diffuser plate degradation (see section 3.4). Once corrected, new versions of the data are released. The latest is V6 which is discussed in greater detail below.

In comparing monthly means of V4 TOMS data and Halley Dobson data between 1979 and 1982, Bowman (1986) found a 5 to 10% discrepancy between the two data sets. The reasons for differences between these data sets included the fact that the TOMS field of view was of varying size and was averaged over the Dobson field of view. In addition, observation times could have differed and different days were missing from each data set; however monthly means were calculated in an attempt to remove these apparent random errors. An intercomparison of TOMS data and Dobson data in 1984 showed the precision of the TOMS data retrievals

to be better than 2% (Bhartia et al., 1984) while Fleig et al. (1986 a) reported that comparison with the Dobson network indicated no detectable (r.m.s. error less than 1%) degradation in the overall V5 TOMS data quality compared to earlier years.

Comparison of total column ozone measurements from TOMS (V5), SBUV and a network of 36 Dobson stations showed that total ozone measured by the SBUV and TOMS instruments declined relative to the Dobson network by 3% between 1979 and 1988 (Heath, 1988; Reinsel et al., 1988). This calibration drift was most likely caused by under correction of the diffuser plate error as discussed in section 3.4 (Watson, 1988). Although use of the PJM corrected for the diffuser plate degradation, some negative drift in this new V6 TOMS data may still exist relative to the Dobson network due to an increase in tropospheric ozone, which is only partially measured by TOMS as detailed above. McPeters and Komhyr (1991) showed that the V6 data are stable relative to total ozone measured by the world standard Dobson spectrophotometer.

A comparison of the first year of SBUV data with total ozone data from the Dobson Network (Bhartia et al., 1984) demonstrated that the total ozone values from SBUV were lower compared to the Dobson values by 8.3% but had a precision of better than 2%. Fleig et al. (1986 b) made use of 7 years of SBUV data from 7 October 1978 to 7 May 1985 and data from 41 Dobson stations to show that the SBUV measurements had drifted with a rough linear trend of $0.38 \pm 0.13\%$ per year. It was also shown that there was an initial bias between the two data sets due to differences in ozone absorption cross-sections used by the two systems. The drift in the SBUV data was ascribed to three possible sources:

1. Drift in calibration of the SBUV instrument.
2. Drift in the Dobson network.
3. Differences in the effect of real geophysical changes on the two different measurement techniques.

The estimated uncertainty in determining the change in the SBUV instrument calibration at an individual wavelength is 0.1% per year at a 95% confidence level. McPeters et al. (1986) expressed the need for high quality meteorological data above the 7 hPa level in order to obtain accurate vertical ozone profiles from the SBUV data.

## 3.6 The TOMS version 6 data product

The raw data from the TOMS experiment are reduced by NASA to provide gridded TOMS data with a resolution of 1° in latitude by 1.25° in longitude. These TOMS data cells each contain a total column ozone value in DU and cover the entire surface of the Earth. It is this data set that has been used for much of the analysis in this thesis.

The changes made from V5 data to V6 data included an improved calibration, the use of updated standard profiles to characterize the Antarctic ozone hole conditions, corrected errors in satellite attitude determination, re-definition of the B-pair ozone wavelengths (317.5 and 331.3 nm) and recomputation of the pair calibration adjustment factors. The V6 data do not invalidate previous ozone hole studies made using V5 data since ozone losses are much larger than calibration errors (Newman et al., 1991). By combining the effects of all the errors described above, Herman et al., (1991) showed the TOMS one-$\sigma$ error to be 2.4 DU over 11 years. For the global average ozone change the two-$\sigma$ error was shown to be $2.8 \pm 1.3\%$ over 11 years. This level of accuracy is sufficiently high for all studies performed in the following chapters.

Figure 3.4 shows the relative changes at the equator in TOMS (V5), SBUV and SBUV/2 (aboard the NOAA-9 spacecraft), with respect to TOMS (V6).

All time series have been smoothed with a 30 day running average to emphasize their long-term characteristics. If it is assumed that TOMS (V6) total ozone is independent of instrument drift, an assumption probably good to within $\pm 1\%$ (Chandra et al., 1990), Figure 3.4 may be used to estimate the long term changes in the instrument characteristics of TOMS (V5), SBUV and SBUV/2.

## 3.7 Nimbus-7 TOMS and Meteor-3 TOMS comparisons

A modified method for determining the diffuser plate reflectivity has been developed for the Meteor-3 TOMS instrument (Herman et al., 1991). This new approach allows the processing of TOMS albedo data into total column ozone measurements with reasonable accuracy, as it is received, making a realtime prod-

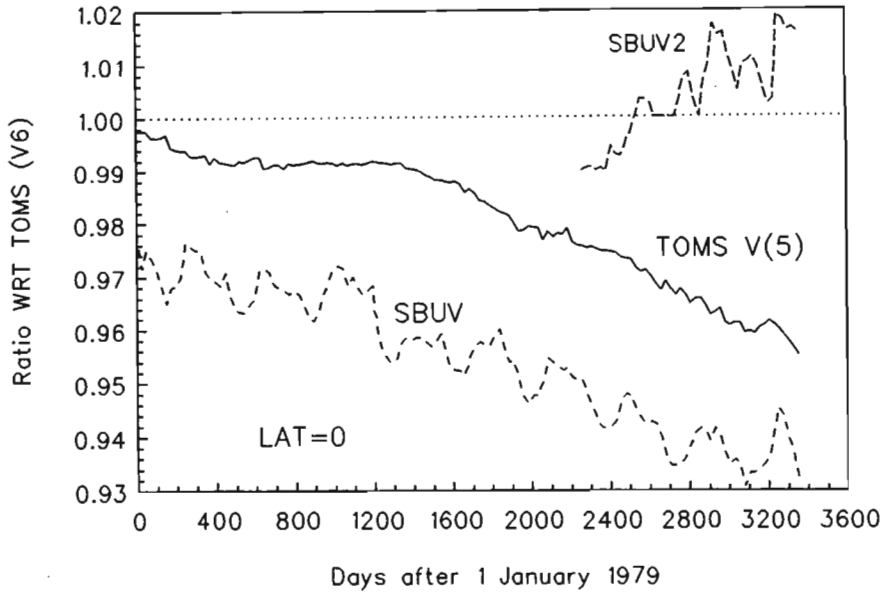


Figure 3.4: Long-term changes in total ozone at the equator for TOMS (V5), SBUV, and SBUV/2 relative to the TOMS (V6) data which have been corrected for long-term instrument change. All the time series are smoothed with a 30 day running mean. From Chandra et al., 1990.

uct available to the user. These data sets were used to direct Antarctic ozone research programmes during the 1993 winter. After an additional year of data have been archived they are reprocessed using the PJM to correct for the small error inherent in extrapolating the diffuser plate reflectivity function.

Simultaneous Nimbus-7 TOMS and Meteor-3 TOMS total column ozone data sets have been obtained for 1993 up until the termination of Nimbus-7 TOMS on 6 May 1993. Figure 3.5 shows a plot of Southern Hemisphere Nimbus-7 TOMS data for 1 January 1993 while Figure 3.6 shows a plot of Meteor-3 TOMS data for the same day. The ozone distribution is shown by means of the colour code indicated on the left of the figure. Lowest ozone levels are shown in hues of blue with values increasing through green and red to the highest values being shown in orange and white. The general ozone morphology in the two images appears to be very similar with low values over the equatorial regions increasing to a small midlatitude zonal maximum before decreasing slightly down to the pole. A difference data set has been generated using

$$\text{Difference} = \text{(Nimbus-7 data)} - \text{(Meteor-3 data)} + 40$$

and is plotted in Figure 3.7. Hues of red, orange and yellow indicate Nimbus-

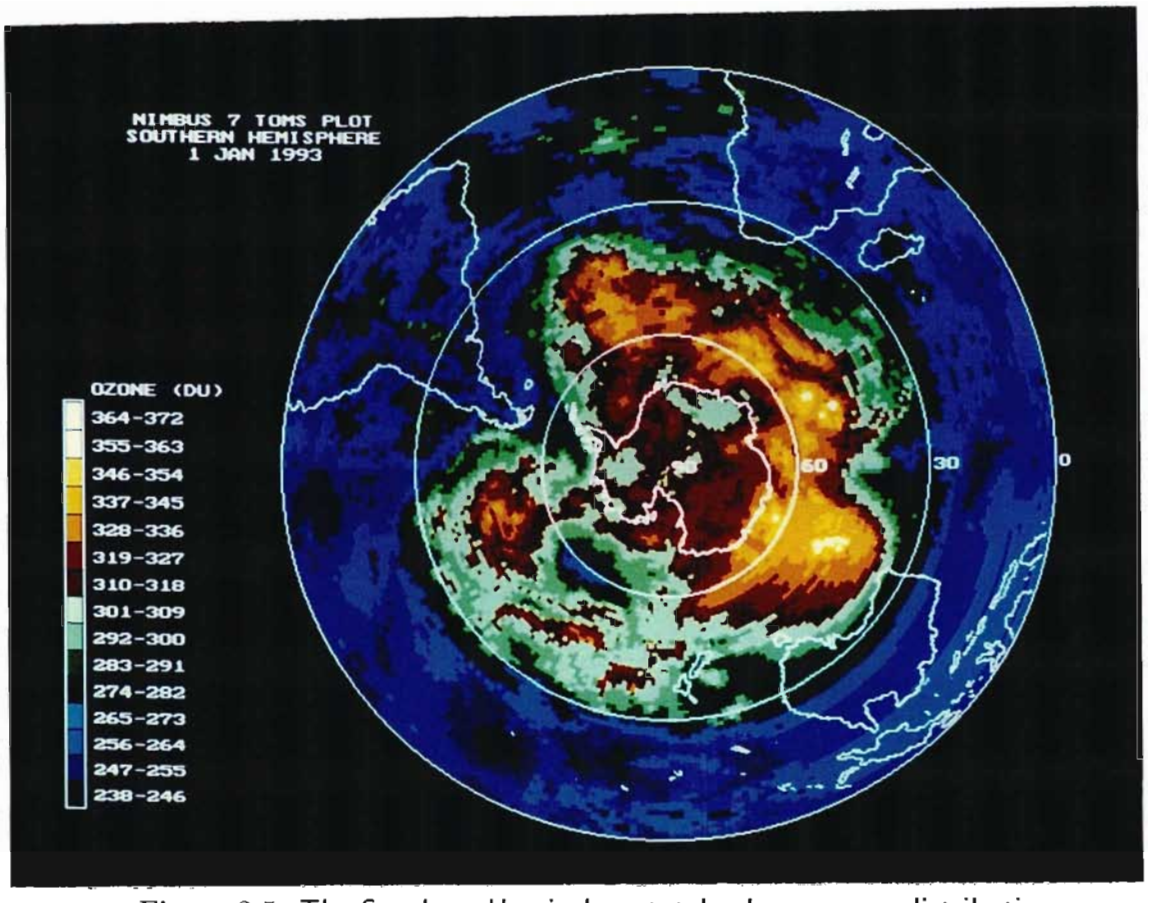


Figure 3.5: The Southern Hemisphere total column ozone distribution as seen by Nimbus-7 TOMS on 1 January 1993.

7 total column ozone in excess of Meteor-3 data while hues of blue and green show regions where Nimbus-7 data are lower than Meteor-3 data. In general, differences are no larger than 15 DU. There appears to be some dependence on orbit path since the orbital track of the Nimbus-7 satellite can be identified in the difference plot (Figure 3.7). Furthermore, there appears to be a region of large positive differences south of Australia. This may arise from temporal differences in the Nimbus-7 and Meteor-3 TOMS measurements. If a large eastward propagating planetary waves disturbance in the total column ozone distribution is sampled at two different times, it is possible that such large differences could be observed. Assuming instrument characteristics to be identical, and knowing the exact time at which measurements were made, it may be possible to make use of differences in the Nimbus-7 TOMS and Meteor-3 TOMS data sets to investigate wave propagation observed in the data and to determine short term variability in global total column ozone. The time difference between the TOMS measurements shown in Figures 3.5 and 3.6 is certainly no larger than 24 hours, while Figure 3.7 shows that variability of up to 34 DU may be expected over this time period.

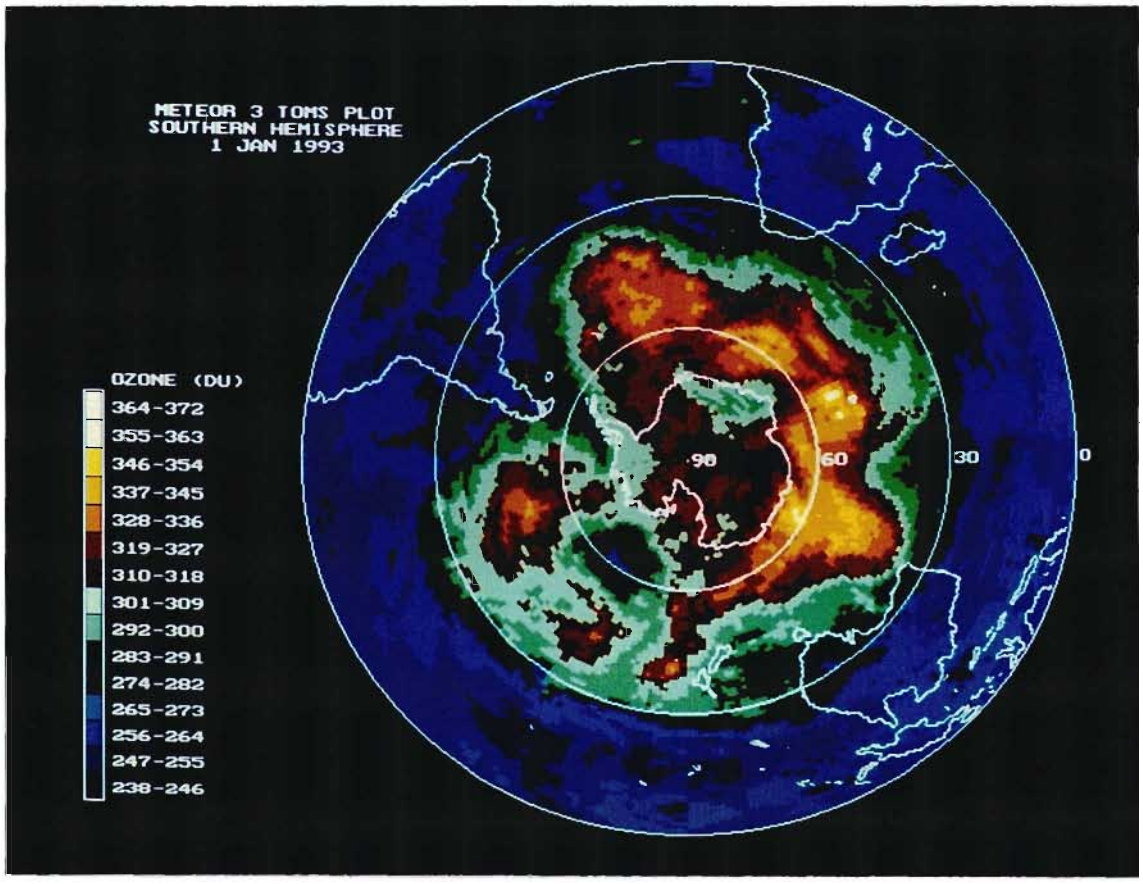


Figure 3.6: The Southern Hemisphere total column ozone distribution as seen by Meteor-3 TOMS on 1 January 1993.

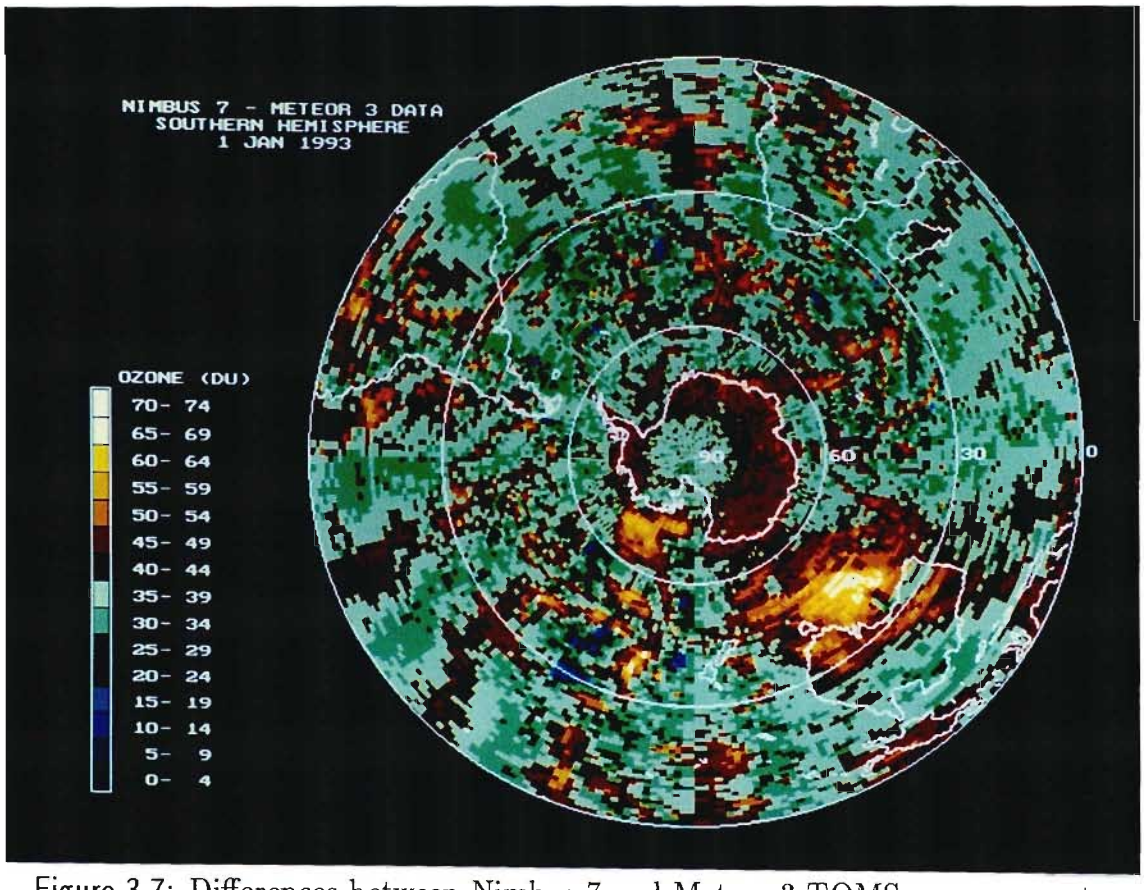


Figure 3.7: Differences between Nimbus-7 and Meteor-3 TOMS measurements on 1 January 1993.

# Chapter 4

# PLANETARY WAVE INFLUENCES ON THE GLOBAL TOTAL COLUMN OZONE DISTRIBUTION

## 4.1 Introduction and Review

The importance of planetary waves in determining the distribution of ozone has long been recognized (Berggren and Labitzke, 1968). However, it has only been with the introduction of two-dimensional models and global satellite data sets, that a more precise understanding of planetary wave influences on ozone has been achieved. Two-dimensional models, with coupled dynamics and chemistry, such as that of Garcia and Solomon (1983), have shown that net transport by steady state planetary waves has a large effect on the mixing ratio of ozone at high latitudes in winter. Figure 4.1 shows the calculated meridional distribution of stratospheric ozone with and without the effect of net chemical transport by planetary waves (Garcia and Solomon, 1983). At high latitudes in the middle and upper stratosphere, this transport significantly enhances ozone mixing ratios. It was shown that at 40 km altitude and 70°N, for example, ozone mixing ratios were almost one third higher when net planetary wave transport was taken into account than when it was neglected. The model excluded the effects of wave transience and dissipation (see section 2.9), making use of a parameterization

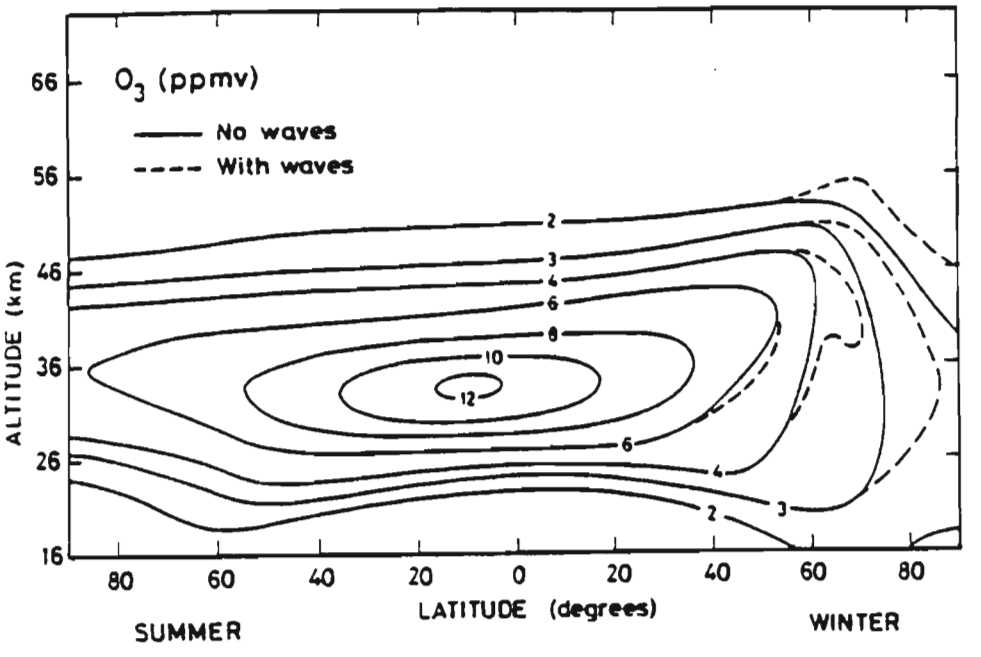


Figure 4.1: Calculated ozone distribution in p.p.m.v. for the Northern Hemisphere winter solstice, with and without chemical eddy transports. From Garcia and Solomon (1983).

scheme for the net transport of chemically active constituents. It was stated that because the effects of wave transience and dissipation are important in the winter stratosphere, the results presented should be viewed as a first approximation.

Such model studies have been supported by observational evidence. Schoeberl et al., (1986) have reported on eastward propagating planetary waves observed in the Southern Hemisphere TOMS ozone distributions. It was shown that during 1981, ozone levels in the circumpolar maximum were modulated by planetary waves whose spectrum varied with season. In the spring, a strong wavenumber 1 modulation was observed with a peak-to-peak amplitude of 100 DU and constant phase. This produced a region of unusually low variance in total ozone near South American longitudes and a maximum, with high variance, near 100°E. A number of eastward propagating planetary waves were recorded in the ozone distribution during the period 1 September 1982 to 30 November 1982 at 70°S, with wave speeds of 10° to 15° of longitude per day. Wirth (1993) showed that these waves are created by a combination of horizontal and vertical advection and are well correlated with waves in the temperature field.

The total column ozone distribution during 1988 showed a zonally asymmetric

pattern with the minimum displaced from the pole towards 90°W (Schoeberl et al., 1989). This shift away from the pole resulted from planetary wave 1 forcing which was large during 1988. The anomalous behaviour of 1988, as a year of reduced Antarctic ozone depletion, will be discussed further below.

Salby and Callaghan (1993) also made use of Nimbus-7 TOMS data, together with simultaneous analyses of the circulation, to examine the origin of large fluctuations in total column ozone in both hemispheres. It was shown that a sizable component of the total ozone variability may be explained by quasi-columnar flow of air in the lower stratosphere. Such flow may be induced by planetary waves. If the layer where total ozone is concentrated is displaced downward (upward) in some region, ozone-rich air will undergo compression (expansion), increasing (decreasing) the ozone number density and hence the column abundance in that region (Salby and Callaghan, 1993).

In addition to TOMS total column ozone data, earlier studies have made use of vertical ozone profile data sets obtained from various satellite experiments.

Chandra and McPeters (1986) made use of ozone measurements from 1970 to 1984 from the Nimbus-4 backscattered ultraviolet (BUV) instrument and from the SBUV on Nimbus-7 (see Chapter 3) to examine the role of planetary waves in determining the spring time ozone distribution above the Antarctic. It was suggested that the contribution to the observed long-term negative trend in zonally averaged ozone data comes mostly from the longitude region between 30°E and 60°W which constitutes the trough of a persistent wavenumber 1 in the Antarctic. Outside this longitude region, the trend in total ozone is much smaller due to strong interannual variability of wave activity (Chandra and McPeters, 1986). These ideas are examined in greater detail in Chapter 5.

Randel (1993 b) also made use of SBUV data to examine normal-mode Rossby waves observed in the stratosphere. In Chapter 2 it was shown that normal-mode Rossby waves always propagate in a westward direction, although, relative to the zonal flow, may appear to propagate eastward. Recall from Chapter 1 (see Figure 1.3) that because of its long lifetime in the lower stratosphere, ozone is advected by propagating waves. However, in the upper stratosphere it is photochemically controlled and its distribution is influenced by temperature perturbations associated with travelling waves (Randel 1993 b). Near-global, coherent westward travelling waves of wavenumber 1 were observed with periods of 5 to 10 days.

The observations also showed good agreement with the predictions of a model of linear wave transport and chemistry.

Planetary wave breaking has been shown to act as a powerful mechanism for the transport of ozone (Leovy et al., 1985) and other long-lived trace gases (Randel et al., 1993) from the tropical to the midlatitude regions. In addition to the TOMS and SBUV instruments on Nimbus-7 there was the Limb Infrared Monitor of the Stratosphere (LIMS) which provided vertical ozone profiles. Using 10 hPa LIMS data, Leovy et al. (1985) showed that between 30 and 3 hPa (the layer in which planetary wave breaking occurs) poleward transport of ozone occurs in narrow tongues drawn out of the tropics and subtropics in association with major and minor warming events. They stated that their results also have implications for interannual and interhemispheric asymmetries in total column ozone. The annual replenishment of the Antarctic ozone hole and the latitudinal increase in the north polar total ozone depends both on the frequency and intensity of midlatitude wave breaking events. The creation of these filamentary structures in turn depends on the number, duration and intensity of stratospheric warmings. In closing, Leovy et al. (1985) suggested that questions about the interannual variability of the Antarctic region must await the development of observational systems capable of mapping the ozone distribution in these regions. In the study presented in this chapter, it is hoped that some light may be shed on the interannual variability of the size of the Antarctic ozone depletion region as determined by midlatitude planetary wave activity. Dunkerton (1991) also made use of LIMS data to analyze wave breaking events and wave propagation in the Northern Hemisphere winter of 1978/79.

In the work detailed below, spatial harmonic analysis is used to determine the power of planetary wave structures observed in the midlatitude total column ozone distribution of both hemispheres. The study presents a long-term climatology of wave behaviour and relates their observed morphology to the interannual variability of the severity of the Antarctic ozone depleted, from 1979 to 1992. The features described below are well above the 1 to 2% error inherent in the TOMS measurements (see Chapter 3).

## 4.2 The fundamental data sets

Daily global total column ozone distributions have been obtained from the version 6 GRID-TOMS data base for each year from 1979 to 1992. Figure 4.2 shows a plot of Southern Hemisphere total column ozone, from this data set, for 4 January 1987. A colour scale is used to show total column ozone where lowest levels are

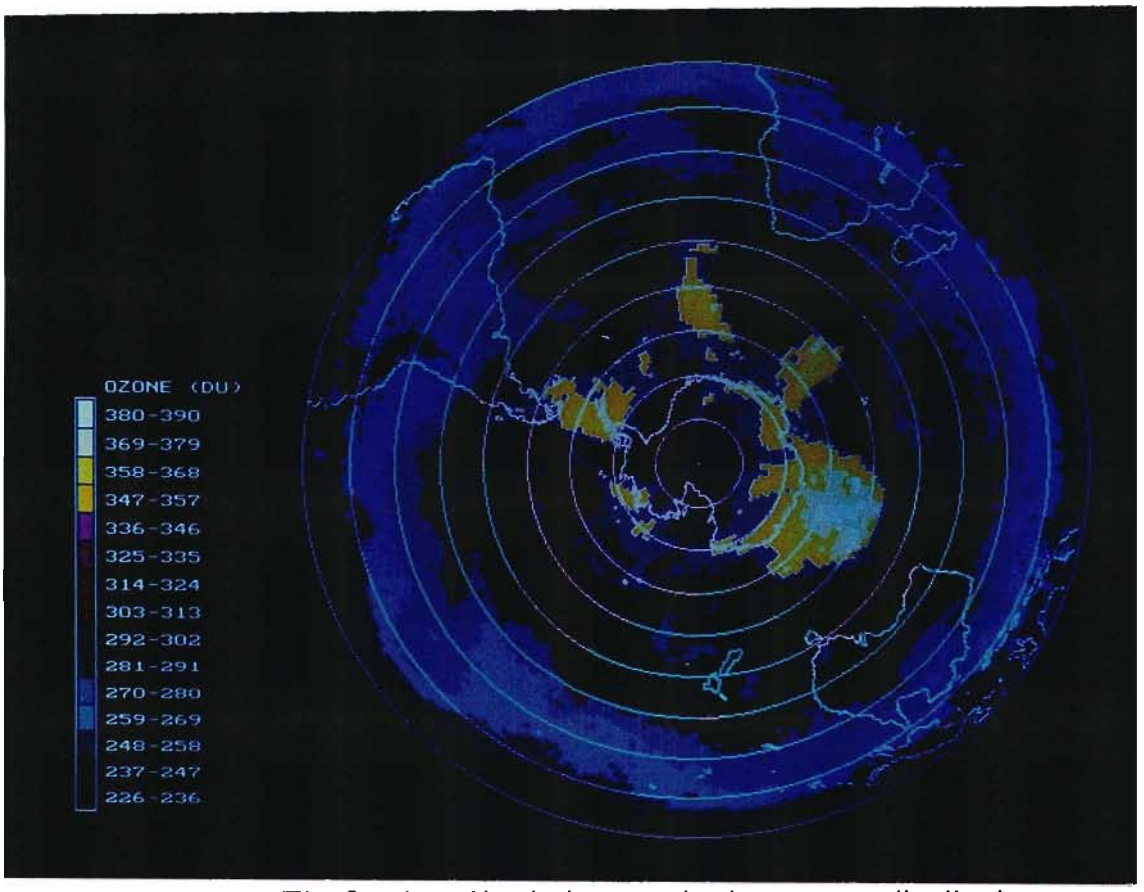


Figure 4.2: The Southern Hemisphere total column ozone distribution as seen by TOMS on 4 January 1987.

shown in hues of blue with values increasing through green and red to the highest ozone levels shown in yellow and white, as indicated by the colour code on the left of the figure. The pentagonal-like morphology of the distribution indicates significant zonal structure, initially thought to result from a planetary wave 5 perturbation to the Southern Hemisphere general circulation.

To examine this structure in greater detail, zonal profiles of total column ozone are averaged from 50°S to 75°S to create a single representative midlatitude profile. This latitude range was selected after a preliminary investigation indicated that

wave powers maximize over this region. Each zonal cell of the GRID-TOMS product spans 1.25° in longitude, resulting in a representative profile of 288 values. There are numerous techniques that may be used to analyze the periodicity in this profile and its temporal evolution. An evaluation of these is given in Appendix A. A preliminary study of planetary waves structures observed in the Northern and Southern Hemisphere total column ozone distribution, from 1979 to 1992, was made using a high resolution Fast Fourier Transform (FFT) algorithm (see Appendix A) to derive daily wave powers. This preliminary study showed that:

- Wave power decreases with zonal wavenumber.
- The sum of the wave powers in the first 6 wave modes provides an accurate representation of the total wave power.
- For wavenumbers 1 and 2, the zonal wave speeds seldom exceed 20° of longitude per day.
- Southern Hemisphere wave 1 power is generally higher than that of the Northern Hemisphere, while Northern Hemisphere wave 2, 3 and 4 power is generally higher than that for the Southern Hemisphere.

These results were used to direct a more detailed, higher quality investigation where a discrete Fourier transform (see Appendix A) was used to calculate the power of the first 6 wave modes in the midlatitude zonal profiles. The decision to calculate wave power in only the first 6 wave modes is supported by Hare and Boville (1965) who reported that large scale planetary waves of wavenumber less than 3 account for most of the wave structure in the stratosphere. Leovy et al. (1985) and Randel (1993 b) have also shown that planetary wave influences on the ozone distribution may be accounted for by examining only the first 6 wave modes. It was initially thought that the use of discrete Fourier transforms in this study would be computationally prohibitive. However, reduction of the spatial transform to the first 6 wave modes increases its speed by a factor of 22. The discrete Fourier transform is then only 5 times slower than the FFT, but provides spectral information of far higher quality (see Appendix A).

Mean midlatitude zonal profiles, obtained from the daily TOMS plots, are reduced to zero mean and passed to the discrete Fourier transform algorithm to obtain the powers for the first 6 wave modes. The midlatitude zonal profile extracted from the data set of Figure 4.2 is shown in Figure 4.3 where the deviation from

the mean, in DU, is plotted against the TOMS cell number. The zonal structure

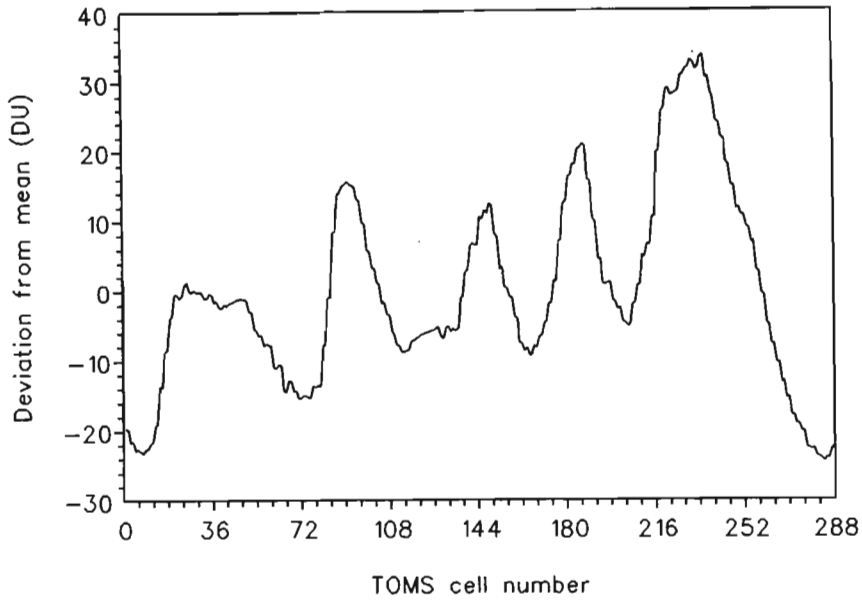


Figure 4.3: The midlatitude representative total column ozone profile obtained by averaging zonal profiles from Figure 4.2 between 50° and 75°S.

is clearly evident in Figure 4.3 where 5 significant local maxima are observed at TOMS cell numbers 27, 91, 149, 187 and 236. The average distance between these local maxima is 52 TOMS cells which when divided into 288 (the number of TOMS cells encircling the globe), suggests a wavenumber 5 or 6 morphology.

The discrete Fourier transform output associated with the data of Figure 4.3 is displayed in Figure 4.4 where the wave power, in $(DU)^2$, is plotted against the frequency in $(TOMS\ cell)^{-1}$. The frequency, when inverted, gives the period, which when divided into 288 gives the wavenumber as shown above. The small arrows in Figure 4.4, indicating the frequencies associated with the first 6 wavenumbers, show that although large power is observed in modes 1 to 4, there is also a peak at wavenumber 6.

By plotting the daily Fourier transform outputs alongside each other, frequency-time plots of planetary wave power may be created for each year. Figures 4.5 a and 4.5 b show examples of such plots using Northern Hemisphere data for 1986 and 1990. In the following analysis (section 4.5) these will be shown to be the years of strongest and weakest Northern Hemisphere wave powers respectively. In Figure 4.5 the frequency domain has been transformed to wavenumber and is plotted against Julian day number. Wave power in $(DU)^2$ is indicated by a

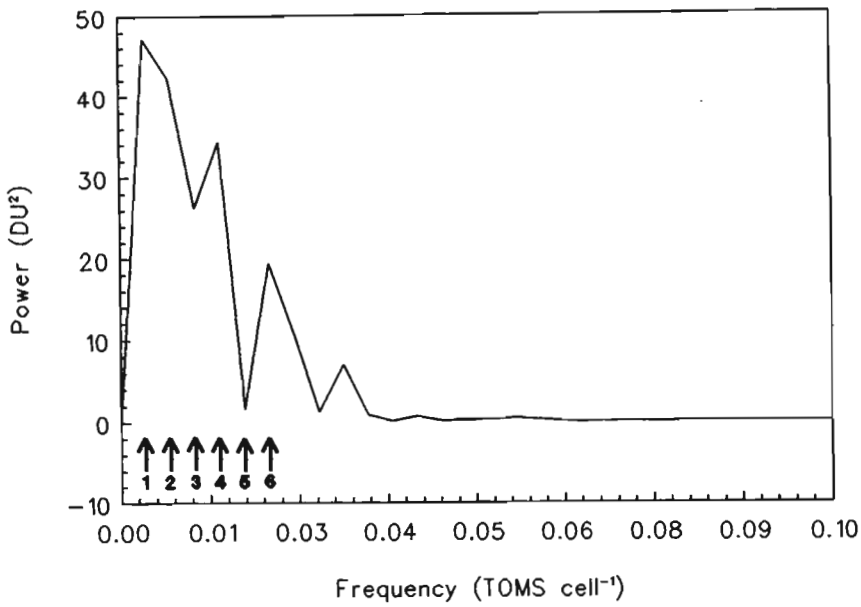


Figure 4.4: The discrete Fourier transform output obtained from the data shown in Figure 4.3. The wavenumbers associated with the frequency plotted on the abscissa are shown using small arrows.

logarithmic colour code shown in the top right corner, the direction of increasing power being from blue through green and red to yellow. Note that although the wave power is concentrated around wavenumbers 1 and 2, there are times where power at higher wavenumbers is observed eg. power at wave 3 on day 31 of 1990 and near day 45 of 1986. In addition, significant wave power is observed only during the Northern Hemisphere winter months, ie. day 300 (27 October) to day 80 (21 March).

Figures 4.6 a and 4.6 b show similar Fourier transform plots for the Southern Hemisphere. Daily wave powers from day 200 to day 335 (discussed later), totalled for wavenumbers 1 to 6 indicate that over the period 1979 to 1992, 1980 (Figure 4.6 a) was the year of minimum Southern Hemisphere planetary wave power while 1988 was the year of maximum wave activity. Compared to the Northern Hemisphere, the wave power maximizes later with respect to the seasons, during the late winter spring and early summer ie. day 200 (19 July) to day 335 (1 December). This period covers the return of solar illumination to the Antarctic stratosphere, the formation of the Antarctic ozone hole, and its dissipation following the breakdown of the circumpolar vortex. It is referred to throughout the following analysis as the *Antarctic Vortex Period* (AVP). In comparison with Northern Hemisphere wave powers (Figures 4.5 a and 4.5 b),

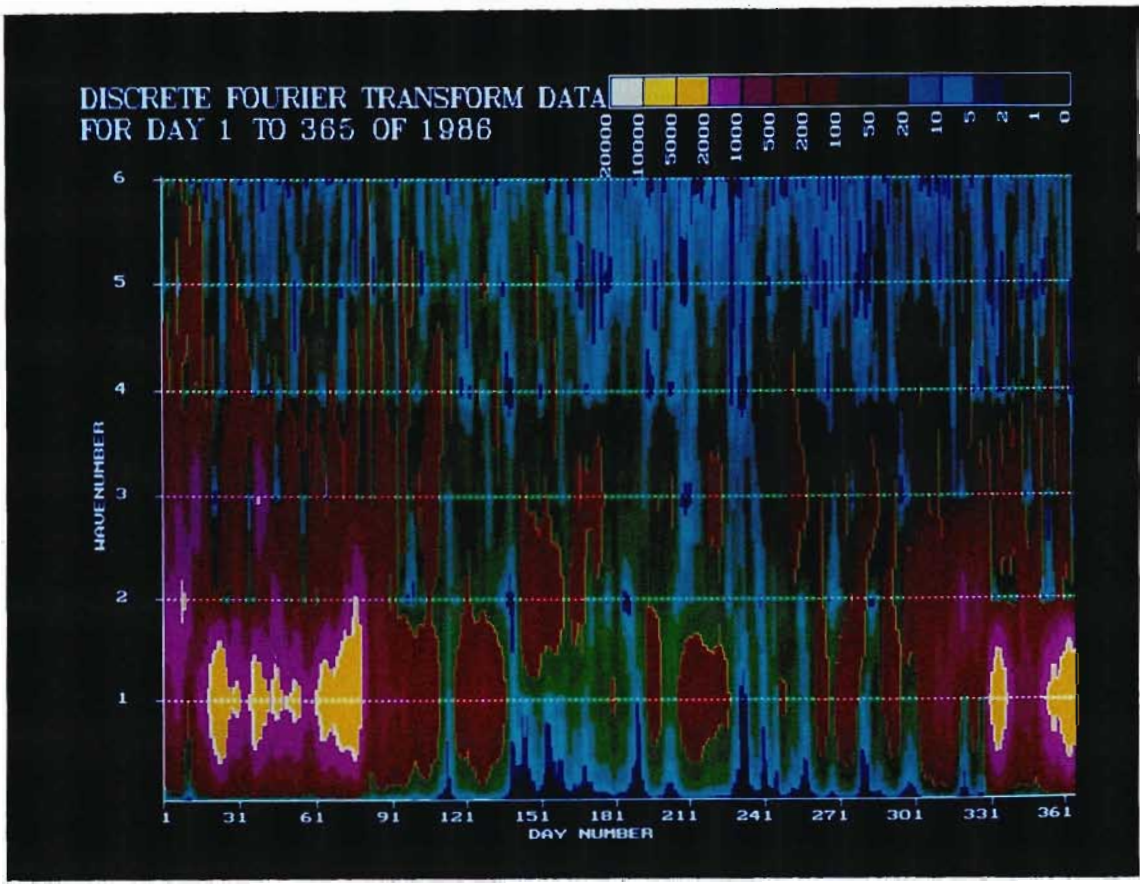


Figure 4.5 a: Northern Hemisphere planetary wave powers for waves 1 to 6 for 1986, derived from TOMS ozone distributions. Powers in $(DU)^2$ are shown by means of the logarithmic colour code shown in the top right of the Figure.

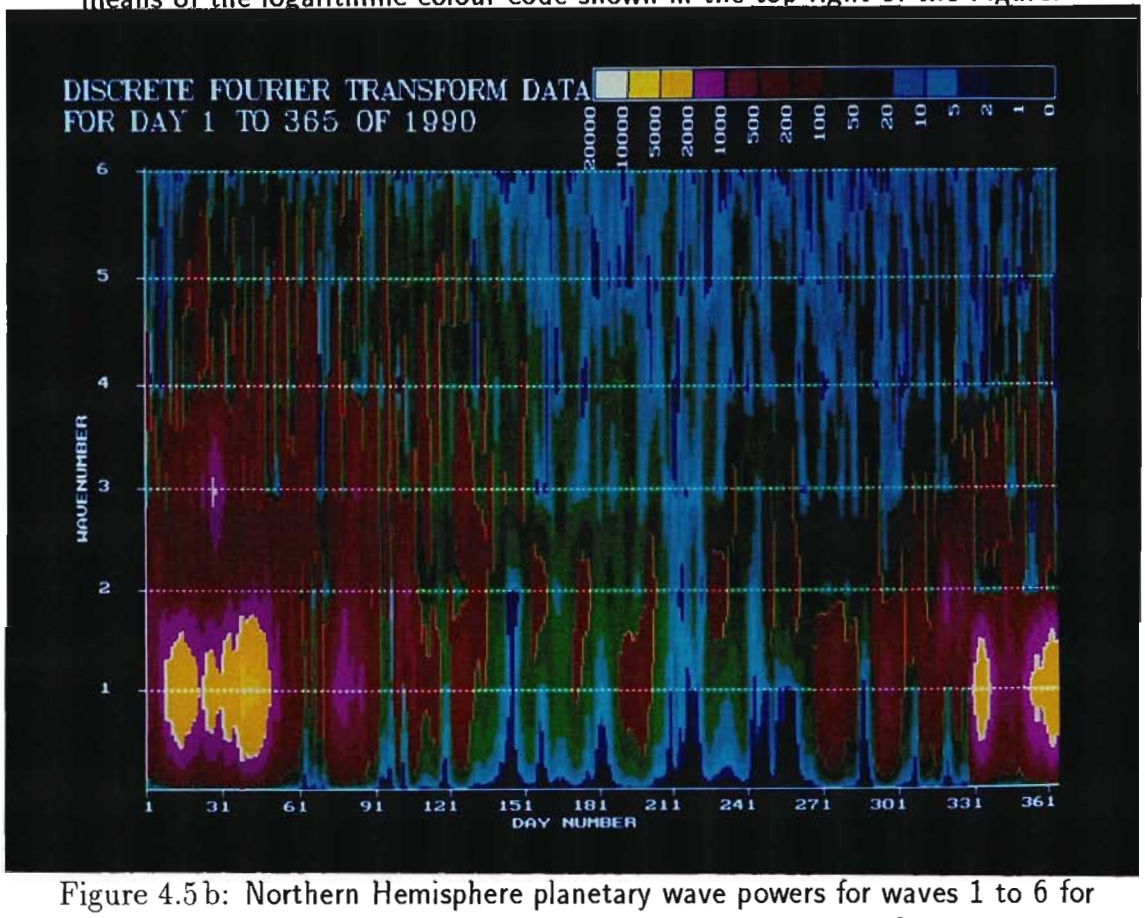


Figure 4.5 b: Northern Hemisphere planetary wave powers for waves 1 to 6 for 1990, derived from TOMS ozone distributions. Powers in $(DU)^2$ are shown by means of the logarithmic colour code shown in the top right of the Figure.

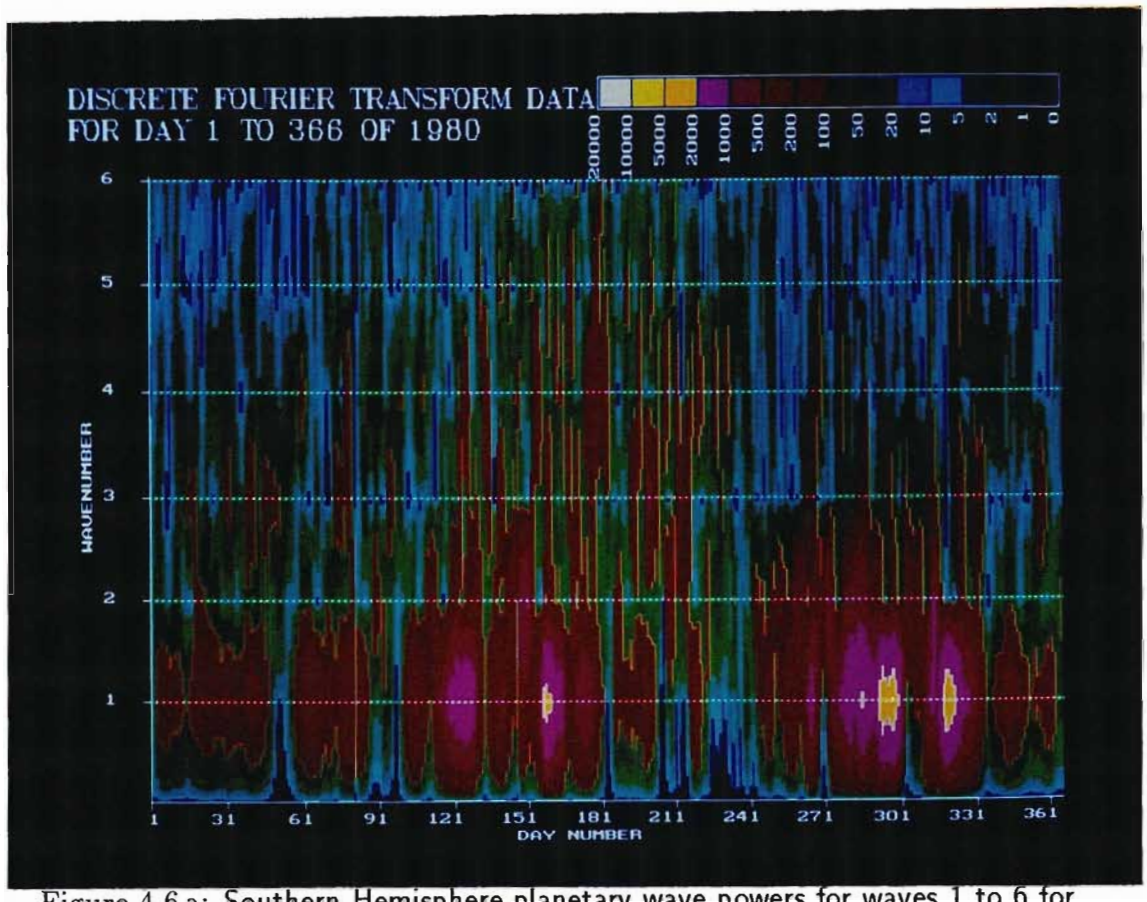


Figure 4.6 a: Southern Hemisphere planetary wave powers for waves 1 to 6 for 1980, derived from TOMS ozone distributions. Powers in $(DU)^2$ are shown by means of the logarithmic colour code shown in the top right of the Figure.

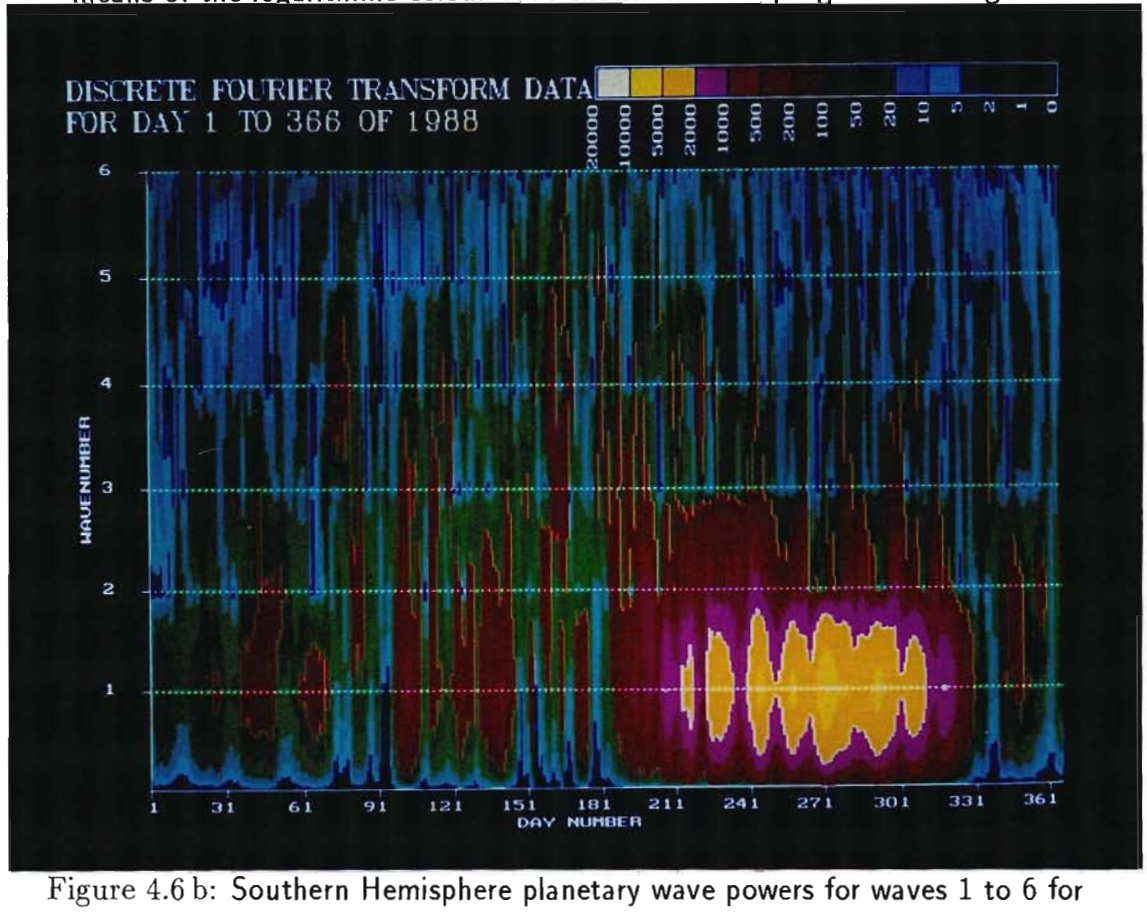


Figure 4.6 b: Southern Hemisphere planetary wave powers for waves 1 to 6 for 1988, derived from TOMS ozone distributions. Powers in $(DU)^2$ are shown by means of the logarithmic colour code shown in the top right of the Figure.

Southern Hemisphere wave powers (Figures 4.6 a and 4.6 b) are more closely constrained to wavenumber 1 and 2.

Plots similar to those shown in Figures 4.5 and 4.6 have been obtained for both hemispheres for each year from 1979 to 1992, providing the fundamental data sets for the analyses presented in sections 4.3 and 4.5.

## 4.3 Southern Hemisphere planetary wave activity and its influence on the severity of the Antarctic ozone hole

### 4.3.1 Southern Hemisphere wave activity: 1979 to 1992

Wave powers for wavenumbers 1 to 6 have been obtained from the discrete Fourier transform outputs for each day from 1979 to 1992. These powers have then been averaged for each calendar day over this period to produce annual mean plots for each wave mode. Figure 4.7 shows the results for the Southern Hemisphere for the four lowest wave modes where the calendar day means in $(DU)^2$, for 14 years, are plotted against day number.

Generally the wave powers decrease with increasing wavenumber. For wavenumbers 1 and 2 the signals peak strongly during the late winter, spring and early summer when the vortex is well formed ie. the Antarctic Vortex Period (the AVP). Wave 3 power shows a local maximum later in the AVP while wave 4 power appears to decrease over this period. In section 2.8.3, an expression was derived for the Rossby critical velocity (see equation 2.136). The equation shows that the critical velocity decreases rapidly with an increase in $k_x$ (the zonal wavenumber). This suggests that the suppression of wave 3 and 4 power during the AVP may be related to the increase in zonal wind speed during this period.

Although there is still uncertainty as to whether or not air is contained within the vortex or processed through it (Hartmann et al., 1989; Tuck, 1989; Randel, 1993 a) interannual differences in planetary wave powers could have marked effects on vortex isolation and chemistry, and hence the severity of the ozone depletion. To prepare the planetary wave power data for comparisons with the annual severity of the Antarctic ozone depletion, wave powers for the first six wave modes have

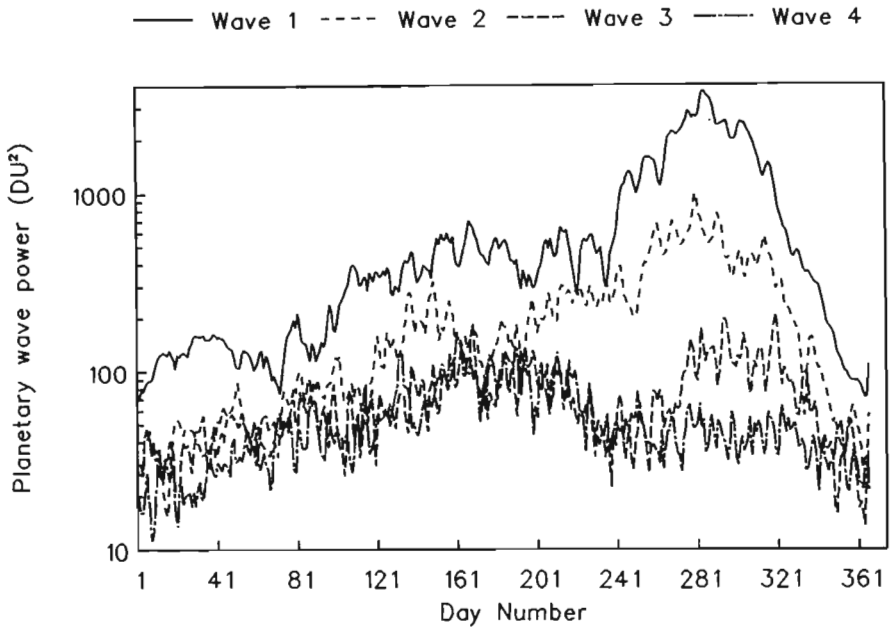


Figure 4.7: Southern Hemisphere planetary wave powers in $(DU)^2$ for wavenumbers 1 to 4 averaged for each calendar day over the years 1979 to 1992. Note the logarithmic scale on the ordinate.

been averaged over the AVP (see above) for each year from 1979 to 1992. Figure 4.8 shows the results for the first four wavenumbers together with the total power of waves 1 to 6.

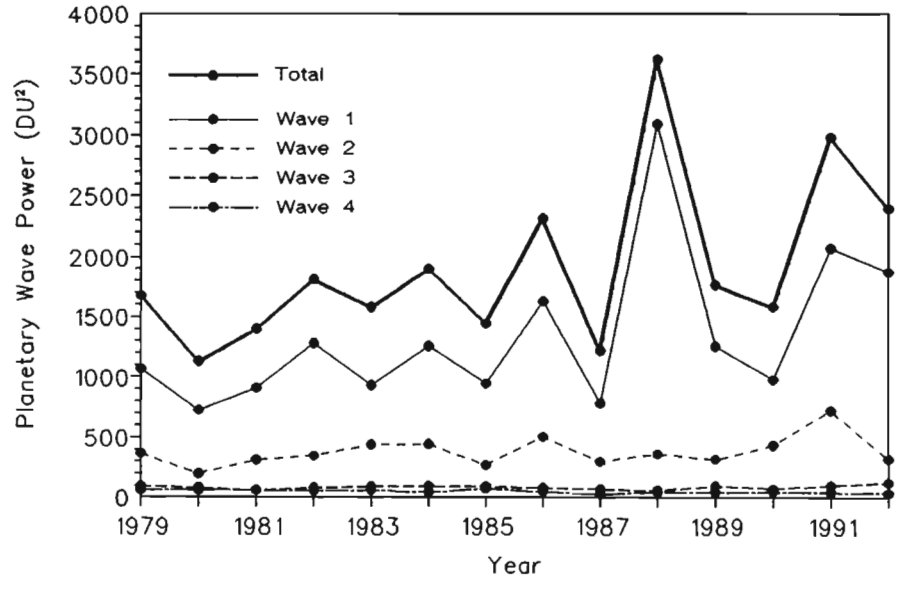


Figure 4.8: Averaged wave powers from day 200 to 335 for wavenumber 1 to 4. The total of the first six wave powers is also shown.

The total wave power signal (bold solid line) shows 1988 to be a year of anomalously high wave power while 1980 and 1987 are the years of weakest wave power. The total wave power for 1988 is more than a factor of 2.9 larger than for 1980 and 1987. The standard deviation of the total power signal is 698 $(DU)^2$ indicating substantial interannual variability in total wave power. The standard deviation for the wave 1 power signal is 643 $(DU)^2$ where the 1988 value (the maximum) is more than a factor of 4.2 greater than the 1980 value (the minimum).

To better illustrate the distribution of the power amongst the different wave modes, Figure 4.9 shows the data used in Figure 4.8 as percentages of the total wave power.

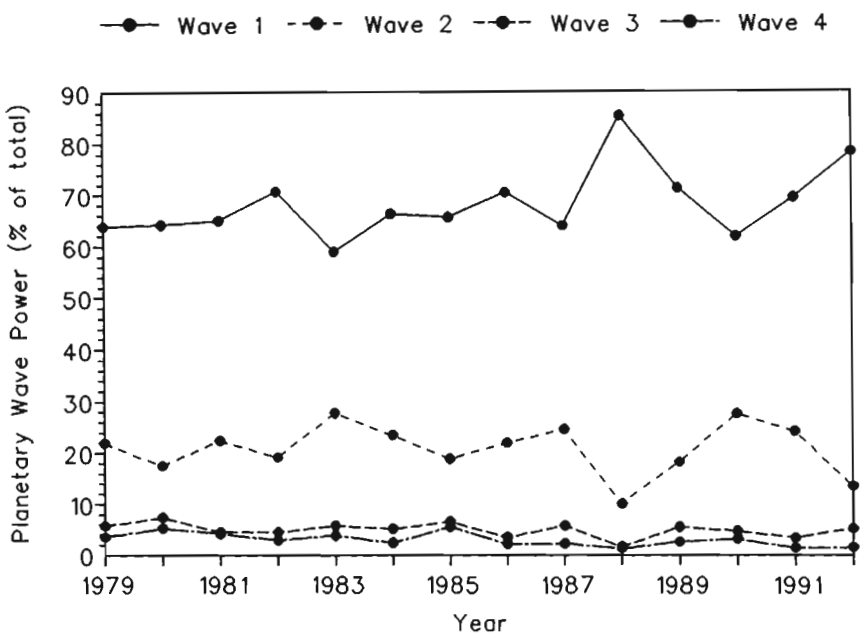


Figure 4.9: Averaged wave powers from day 200 to 335 as a percentage of the total wave power for wavenumber 1 to 4.

Again 1988 appears to differ significantly from the norm where the power is preferentially concentrated at wave 1 and reduced at waves 2 to 4. Furthermore it appears that in certain years of reduced planetary wave activity (eg. 1983 and 1987), the percentage of the total power in wave 1 has been reduced while small increases in the percentage of the total power in waves 2 and 3 are observed.

The interannual variability in the severity of the Antarctic ozone depletion as seen by TOMS is well documented (Gleason et al., 1993; Herman et al., 1993). The total planetary wave power derived in this 1979 to 1992 study has shown 1988 to clearly have the strongest wave power, while 1987 is associated with having one of

the weakest wave powers. It was reported (Schoeberl et al., 1989) that minimum Antarctic total column ozone declined only 15% during September 1988 compared to nearly 50% during September 1987. The maximum total ozone values in 1988 were the largest October monthly mean levels recorded by TOMS since it began operation in 1979. The zonal symmetry in the ozone distribution was severely distorted, with the minimum displaced from the pole towards 90°W. Such a displacement could be caused by a large planetary wave 1 perturbation (Schoeberl et al., 1986). The reduced solar heating of ozone as a result of ozone depletion has been implicated in delaying the Antarctic final warming in 1987 (Kiehl et al., 1988). From these results it appears qualitatively that years characterized as having weak wave powers experience enhanced Antarctic ozone depletion while years identified as having strong wave powers have reduced Antarctic ozone depletion.

### 4.3.2 Comparisons of total planetary wave power and the severity of the Antarctic ozone depletion

To examine the connection between planetary wave power and the depletion of the Antarctic ozone hole in greater detail, a measure of the ozone hole size and depth has been derived for each day from 1979 to 1992. This has been done by first flagging each sunlit TOMS cell of less than 220 DU (Stolarski et al., 1990) ie. the threshold for depletion, from 40°S to 90°S (Krueger et al., 1992). Inside each of these cells, a total mass of ozone may be calculated as follows:

The ideal gas equation is:

$$\begin{aligned} pV &= nRT \\ &= \frac{M}{m}RT \end{aligned}$$

where M = total mass

m = molar mass

$= 3 \times 16 \times 10^{-3}$ kg

The density of ozone at STP is given by:

$$\begin{aligned} \rho &= \frac{M}{V} \\ &= \frac{mp}{RT} \end{aligned}$$

$$= \frac{48 \times 10^{-3} \times 10^{5}}{8.3 \times 273}$$
$$= 2.11 \text{ kg m}^{-3}$$

One DU = $10^{-5}$m of ozone at STP. Therefore, for a column of ozone of cross-sectional area 1 $m^2$, the mass per unit area is given by $2.11 \times 10^{-5}$ D $kg\,m^{-2}$ where D is the total column ozone in DU.

The area of any TOMS cell must be calculated by integrating over the area of the surface of the globe defined by the latitude and longitude borders of the cell as shown in Figure 4.10.

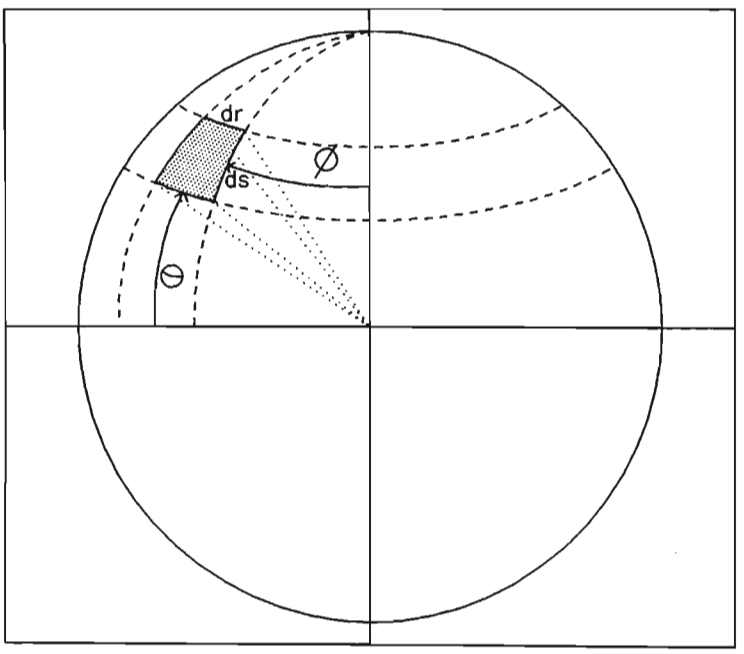


Figure 4.10: Calculation of the area of a TOMS cell.

Then:

$$\begin{aligned}
\text{ds} &= \text{r}\, d\theta \\
\text{dr} &= \cos\theta\, \text{r}\, d\phi \\
\text{Area} &= \text{dr} \times \text{ds} \\
&= \text{r}^2 \cos\theta\, d\theta\, d\phi \\
\text{Area of cell} &= \int_{\phi_1}^{\phi_2} \int_{\theta_1}^{\theta_2} \text{r}^2 \cos\theta\, d\theta\, d\phi \\
&= (\phi_2 - \phi_1)\text{r}^2(\sin\theta_2 - \sin\theta_1)
\end{aligned}$$

where $(\phi_2 - \phi_1)$ = 1 TOMS cell = $2\pi/288$ = 0.021817 and $\theta_2$ and $\theta_1$ are the

poleward and equatorward latitudes of the cell respectively.

The difference in mass calculated using the 220 DU threshold for depletion and the measured total column ozone is taken to be the 'depleted mass' of ozone. These mass differences are summed over all flagged TOMS cells to obtain a daily depleted mass of ozone poleward of 40°S.

Results for 1988 are shown in Figure 4.11 together with the calculated total planetary wave power. The total wave power signal (solid line) shows high

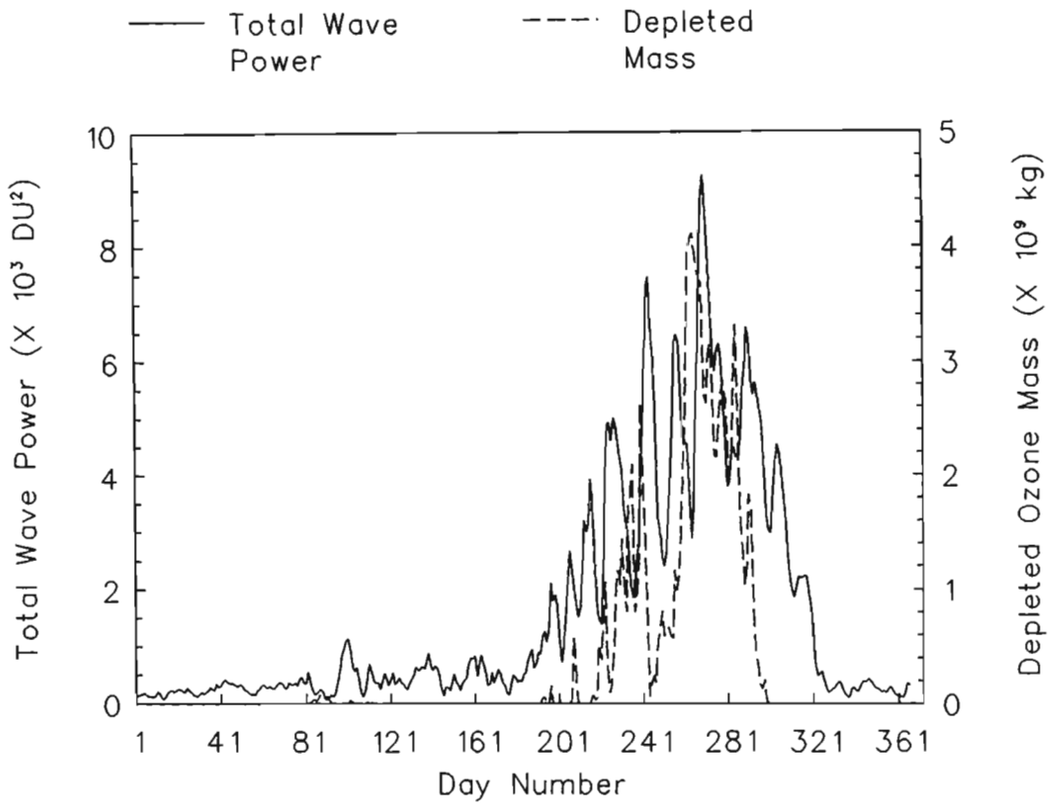


Figure 4.11: Daily total planetary wave power (wave 1 to 6) and depleted ozone mass for 1988

variability during the winter as a result of episodic tropospheric forcing (Randel et al., 1987). Wave powers in the first 6 wave modes have not been analyzed separately for each year, although this could prove instructive as it has been suggested (Randel et al., 1987) that similarity in the dynamical structure of wave 1 and 2 could be indicative of wave packet behaviour. The depleted mass signal (dashed line) consists of two main peaks with almost no depletion just after day 244. The peak before day 244 may result from the ozone hole center being shifted away from the pole towards partially sunlit regions of the Southern Hemisphere.

Recall that the daily depleted mass values were calculated using only sunlit TOMS cells with the result that the depletion will be underestimated before the spring equinox (day 264). However, it may also be argued that since it is the sun that drives the photochemical destruction of ozone, it should not be expected that depletion will occur within non-illuminated TOMS cells unless there is sufficiently strong mixing inside the vortex. A previous analysis (Schoeberl et al., 1989) made use of interpolation across the Antarctic to obtain data poleward of the polar night terminator. This suggests an avenue for further work. The second depleted mass peak, occurring after day 241 (Figure 4.11), indicates the main Antarctic ozone hole. Small peak-for-peak anticorrelations between the total wave power signal (solid line) and the depleted mass signal (dashed line) are observed near day 221, 236 and 264 in Figure 4.11. These may also be indicative of wave 1 forcing of the vortex center away from the pole as discussed in Chapter 1. In general, good peak-for-peak anticorrelations are observed between the depleted mass and total wave power signals through all years, although the calculated correlation coefficients are not statistically significant.

To investigate the interannual differences in the severity of the Antarctic ozone depletion and its dependence on wave activity, the depleted ozone mass signal has also been averaged over the AVP for each year from 1979 to 1992. Figure 4.12 shows these averaged depleted mass values together with stratospheric chlorine loading over the same period.

The chlorine data were obtained from Solomon (1990) up to 1988, while a 3.5% annual rate of increase (Toohey, personal communication) was used to project values to 1992. Chlorine levels are seen to increase approximately linearly from 2.55 p.p.b.v. in 1979 to 4.02 p.p.b.v. in 1993, a rise of 58%. Likewise the AVP averaged ozone depletion increases from practically zero in 1979 to $9.08 \times 10^9$ kg per day in 1992. Although the Antarctic ozone depletion is seen to increase with the rise in stratospheric chlorine there are considerable interannual differences. In particular, the ozone depletion in 1987 is severe while the depletion in 1988 is greatly reduced as discussed above. Assuming that the severity of the Antarctic ozone depletion depends only on stratospheric chlorine loading and planetary wave activity, attempts have been made to normalize the AVP depleted mass values with respect to the chlorine loading.

A first attempt to normalize the integrated depleted mass signal assumes a linear relationship between the stratospheric chlorine loading and the depleted ozone

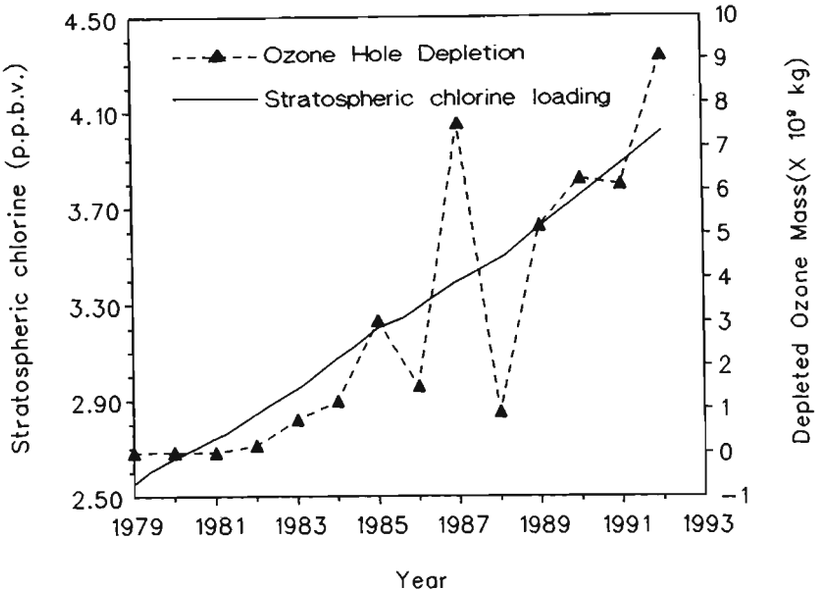


Figure 4.12: Depleted ozone mass averaged over the AVP as a measure of Antarctic ozone depletion and stratospheric chlorine loading. Chlorine data as cited in Solomon (1990) and Toohey (personal communication).

mass, as suggested by Figure 4.12. Deviations from values predicted by a regression analysis are then compared with the AVP averaged planetary wave powers. The linear relationship between the depleted ozone mass signal and the stratospheric chlorine loading is given by:

$$\text{Depleted Mass} = 5.898 \times \text{Chlorine Loading} - 16.16 \qquad (4.1)$$

where the depleted mass is measured in Mt (1 Mt $= 1 \times 10^9$ kg) and the chlorine loading (C.L.) is measured in p.p.b.v.. The regression coefficient ($r^2$ value) for this linear fit was 0.75. Deviations of the measured AVP depleted mass values from the linear prediction are displayed in Figure 4.13 (dashed line) together with the AVP averaged total planetary wave power (solid line).

Note that an inverted scale has been used on the left ordinate to show more clearly the strong negative correlation between normalized depleted mass and the integrated wave power, particularly from 1984 to 1992. The year of 1987 has both the maximum normalized depleted mass and one of the lowest total wave powers while 1988 has the minimum normalized depleted mass and the maximum total wave power. In Figure 4.13 it appears that although wave power was higher in 1991 than in 1986, the Antarctic ozone depletion was worse in 1991. This could

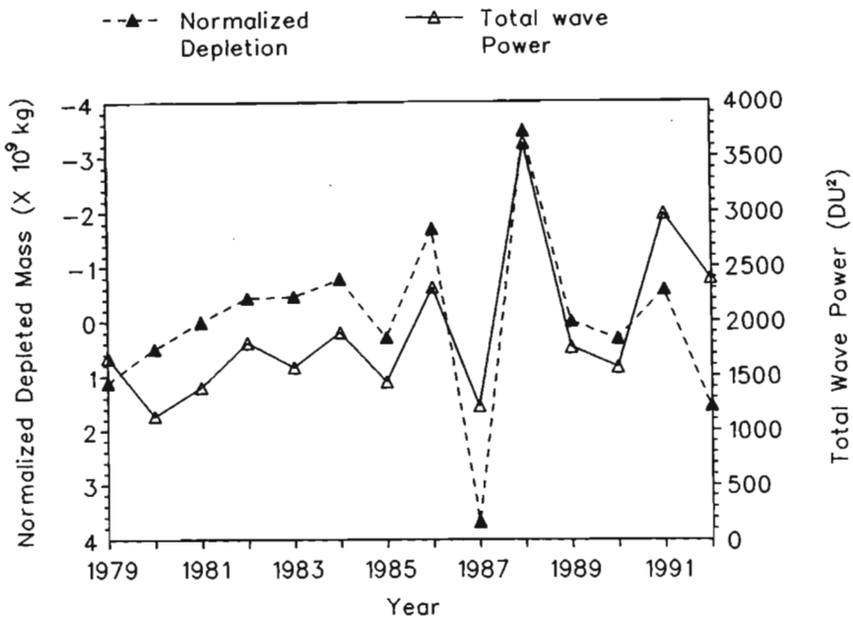


Figure 4.13: Total wave power averaged over the AVP together with the normalized depleted mass signal averaged over the same time period. In this case a linear normalization scheme has been used.

result from a non-linear dependence of ozone depletion on stratospheric chlorine.

To allow for a non-linear dependence of Antarctic ozone depletion on stratospheric chlorine loading, a second attempt at normalization has been made. A second order polynomial was fitted to a plot of depleted ozone mass versus chlorine loading to obtain:

$$\text{Depleted Mass} = 2.590 \times (\text{C.L.})^2 - 11.04 \times (\text{C.L.}) + 11.02 \qquad (4.2)$$

where, as before, the depleted mass is measured in Mt and the chlorine loading (C.L.) is measured in p.p.b.v.. The regression coefficient ($r^2$ value) for this fit was 0.77. Deviations of depleted ozone mass from this second order function are plotted together with averaged total planetary wave power in Figure 4.14. The total wave power averaged over the AVP is plotted on the right ordinate while the normalized depleted mass signal, averaged over the same time period, is plotted on an inverted scale on the left ordinate. A clear negative correlation is seen where years of high (low) wave activity are associated with weak (severe) Antarctic ozone depletion. Regression coefficients show that the linear normalization scheme provides a better anti-correlation between Antarctic ozone depletion and planetary wave power.

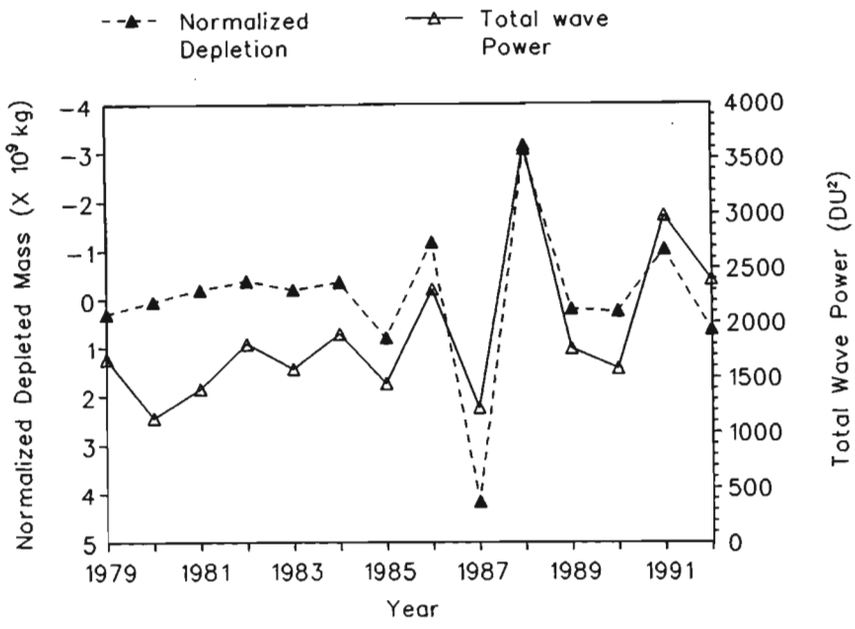


Figure 4.14: Total wave power averaged over the AVP together with the normalized depleted mass signal averaged over the same time period. In this case a second order normalization scheme has been used.

### 4.3.3 Planetary wave influence on Antarctic upper air temperatures

Consideration has also been given to the question of whether planetary wave activity might modulate the upper air temperatures of Antarctic sites. It is expected that during years of strong planetary wave activity, the vortex will not be centered on the pole but will oscillate around the continent in response to wave forcing. For locations at the edge of the continent, it might be expected that at certain times the site will lie within the vortex (experiencing low upper air temperatures) and at other times outside the vortex (experiencing higher upper air temperatures). It is also possible the heat generated by planetary wave breaking at the vortex edge and high altitude heat transport into the vortex (Chapter 1) may contribute to higher stratospheric temperatures of certain Antarctic stations. All of the effects may act as mechanisms for a correlation between years of high planetary wave power and years of high stratospheric temperature for a particular Antarctic site.

Figure 4.15 shows a plot of the AVP averaged total wave power in $(DU)^2$ and the AVP averaged upper air temperatures at 100 hPa obtained at SANAE (70° 18' S, 2° 21' W) for each year from 1980 to 1991 (temperature data were not available

for 1979 and 1992).

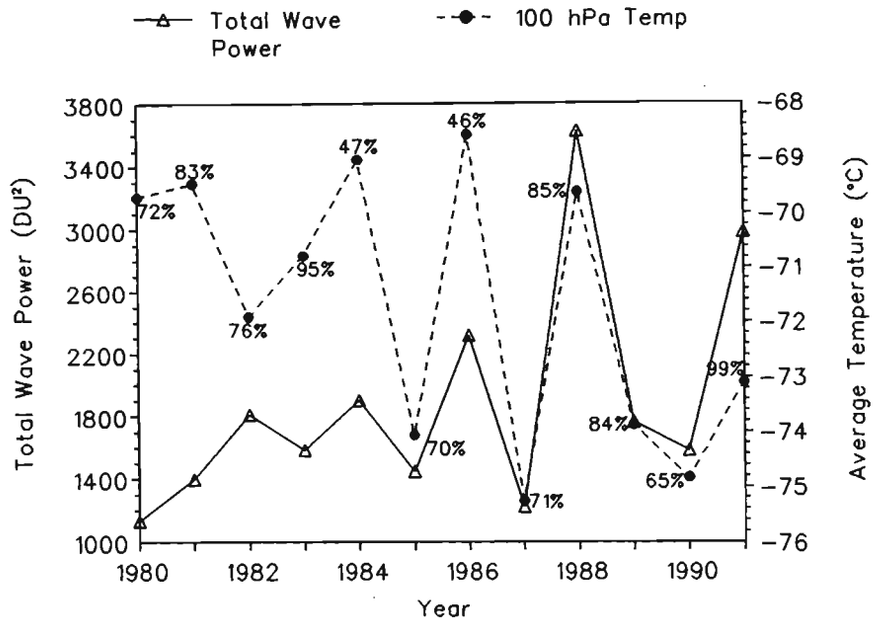


Figure 4.15: Total wave power and 100 hPa temperatures averaged over the AVP from the SANAE Antarctic base.

Temperature readings were not available for every day within the AVP as a result of rawinsonde balloons bursting before reaching 100 hPa. The availability of 100 hPa temperature data for each year is indicated as a percentage next to each temperature data point in Figure 4.15. This figure indicates that, particularly over the period 1985 to 1991, wave power is positively correlated with temperature with a correlation coefficient over this period of r = 0.71. Wave power and temperature correlations may also be expected over shorter periods as the vortex edge moves backward and forward across the measurement site in response to wave forcing. An investigation of this phenomenon is delayed until section 4.5 where a more detailed account of wave amplitudes and phases is given.

### 4.3.4 Discussion

It follows from the foregoing that although the long-term temporal evolution of the Antarctic ozone hole is related to the increase in stratospheric chlorine loading, interannual differences in the severity of the Antarctic ozone depletion are large and may be accounted for by dynamic processes driven by planetary waves. There currently exists disagreement as to whether the Antarctic vortex acts as a 'flowing processor' or as an isolated material entity (Randel, 1993 a).

The long-standing belief is that the vortex isolates the Antarctic stratosphere from the stratosphere of the rest of the Southern Hemisphere, creating a containment vessel inside which chemical depletion of ozone occurs (Chapter 1). In the middle and upper stratosphere, the vortex edge seems to be nearly impermeable, with the result that there is little or no transport of midlatitude air into the vortex. Under this assumption, planetary wave activity may affect the Antarctic vortex in the following ways:

- Wave 1 forcing of the vortex center away from the pole can perturb the $NO_x$ equilibrium, resulting in active chlorine returning to the reservoir species chlorine nitrate (see Chapter 1).
- Planetary wave erosion of the vortex edge tends to decrease the size of the vortex limiting the area over which PSCs may occur and reducing the severity of the annual ozone depletion. PSCs are a prerequisite for the heterogeneous chemistry associated with polar ozone depletion as discussed in Chapter 1.
- Enhanced meridional heat transport by planetary waves creates a warmer vortex, further reducing the probability of PSC formation.

All three of these processes contribute to the negative correlation between planetary wave power and the severity of the Antarctic ozone depletion observed in Figures 4.13 and 4.14.

Significant nonlinearity and feedback exists in this system. A year of reduced meridional temperature gradient results in a reduced meridional pressure gradient which in turn creates a weaker geostrophic wind constituting the circumpolar vortex (see section 2.4). It is well known that the propagation of planetary waves depends on the strength of this stratospheric westerly jet (section 2.8.3; Charney and Drazin, 1961; Schoeberl and Geller, 1976). A weaker geostrophic wind permits greater planetary wave activity, resulting in an increase in the frequency of wave breaking events and a concomitant increase in poleward heat transport at high altitude. Increased wave activity indirectly produces higher polar ozone as discussed above. The increased poleward heat transport, together with enhanced stratospheric heating due to increased ozone concentrations, further reduces the meridional temperature gradient and the strength of the geostrophic wind. Hence

a small initial perturbation to the planetary wave activity may significantly influence the size and depth of the ozone hole through these feedback mechanisms.

Planetary wave activity may also influence the timing of the vortex breakdown. It could be expected that during years of enhanced wave activity, the vortex breakdown will occur early in the season, limiting the period for which unique Antarctic stratospheric conditions permit ozone depletion. This possibility has been investigated by examining the daily depleted mass signals for 1986 and 1988 (years of reduced Antarctic ozone depletion) and 1987 and 1992 (years of enhanced Antarctic ozone depletion). The four signals are shown in Figure 4.16 where the daily depleted ozone mass is plotted against day number from day 280 (7 October) to day 350 (16 December). Note that not only are the depleted

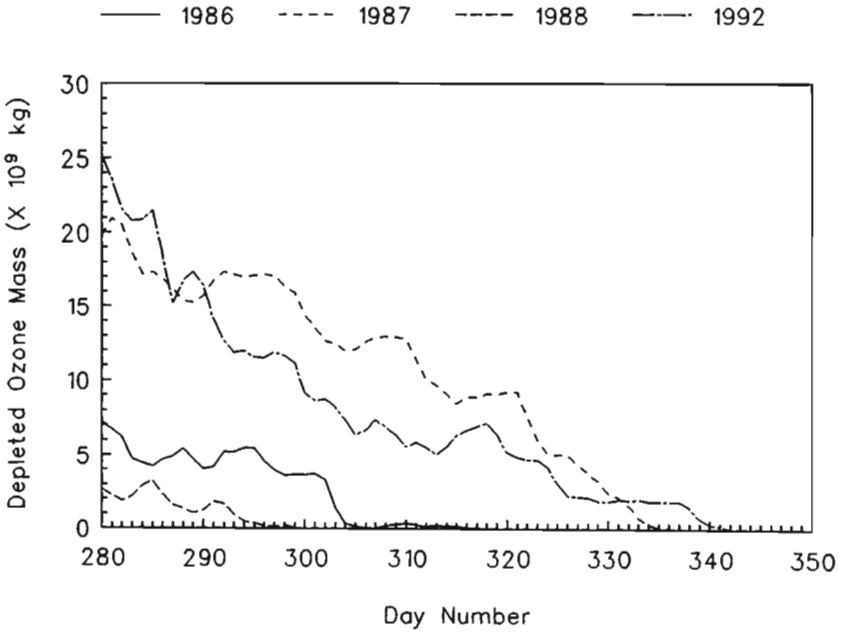


Figure 4.16: Daily depleted mass values for 1986, 1987, 1988 and 1992.

mass signals for 1987 and 1992 significantly higher than those for 1986 and 1988, but they also extend later into the year by approximately one month. This may be indicative of the extended vortex lifetimes during 1987 and 1992 as a result of suppressed planetary wave activity.

The results presented here show that interannual differences in the midlatitude planetary wave activity account for much of the variability in Antarctic ozone depletion. This immediately prompts the question of what causes interannual variability in midlatitude planetary wave activity. A discussion of the influence of the Quasi-biennial Oscillation (QBO) on planetary wave activity is delayed until

the next chapter where the effects of the QBO on the global ozone distribution are examined in greater detail.

In addition to QBO influences, interannual variability in planetary wave forcing mechanisms could account for differences in wave activity. Since topographic forcing does not change from year to year this leaves only interannual variability in land-sea heating contrasts and latent heat release in cumulus cloud systems. Neither of these two possibilities have been pursued in this thesis but suggest avenues for future work.

Leovy et al. (1985) speculated that the integrity of the polar vortex acts as a barrier to the filamentary structures in the ozone distribution which transport ozone from the midlatitude surf zone poleward. Therefore, interannual differences in the strength of the geostrophic wind could result in differences in planetary wave propagation and vortex containment. Interannual differences and long-term trends in vortex strength may be related to anthropogenic increases in stratospheric $CO_2$ loading. Since the wintertime cooling of the Antarctic stratosphere occurs primarily via $CO_2$ radiative cooling (Brasseur and Solomon, 1986), an increase in Antarctic stratospheric $CO_2$ could lead to enhanced cooling (Pitari et al., 1992), greater PSC formation and a steeper meridional temperature and pressure gradient which would create a stronger geostrophic wind. These processes also suggest fields for further research.

## 4.4 Comparison of planetary waves observed in the total column ozone distribution and NMC data for the Southern Hemisphere

A previous study by Manney, Farrara and Mechoso (1991), hereafter referred to as MFM, made use of meteorological data from the U.S. National Meteorological Center (NMC) to examine the behaviour of waves 1 and 2 in the Southern Hemisphere stratosphere during late winter and early spring. Analyses of geopotential height fields from 30°S to 80°S, over the period 1 July to 31 October from 1979 to 1988, were used to derive planetary wave amplitudes at the 10 hPa level. Planetary wave phases, calculated at the latitude of maximum wave amplitude, and wave powers were plotted as functions of latitude and time for each year. These

plots completely characterize the behaviour of waves 1 and 2 in the Southern Hemisphere. The amplitude, in terms of geopotential height, provides an exact measure of planetary wave activity while changes in the phase may be related to zonal propagation of the wave structure.

A different technique has been used in this thesis to perform a similar analysis on the TOMS data, providing analogous plots of total column ozone planetary wave amplitudes and phases. The planetary wave analysis described in section 4.2 has been repeated for each year from 1979 to 1988 for the months of July to October (similar to the AVP described above). In this study however, zonal total column ozone profiles have not been averaged to produce a single representative profile for each day but are rather analyzed individually to provide a measure of planetary wave amplitude for each latitude from 30°S to 80°S for each day. Two-dimensional plots of wave amplitude, as a function of latitude and day number, may then be derived as in MFM. Furthermore, instead of using discrete Fourier transforms a robust least squares fit algorithm has been used which is able to more accurately determine the phase and true amplitude (as opposed to the power) of the planetary wave signal. A detailed description of this algorithm is given in section A.3 of Appendix A. Extracting the wave amplitudes at each latitude from 30°S to 80°S, and the wave phase at the maximum amplitude, provides a thorough description of wave behaviour. The entire analysis has been performed for both waves 1 and 2. However, due to space constraints, detailed results are presented for wave 1 while only phase variations of wave 2 are discussed. Since wave 1 contains most of the power, the exclusion of the presentation of the wave 2 amplitude analysis is not a serious omission.

### 4.4.1 Wavenumber 1 behaviour

Figures 4.17 to 4.26 show the results of the comparison of MFM data and wave parameters derived from TOMS data for 1979 to 1988 respectively. The upper panel of each figure shows the results of the analysis using TOMS data while the lower panel shows the analogous plot obtained directly from MFM. Latitude, from 30°S to 80°S, is plotted against time (in days after 1 July) in both panels. The planetary wave amplitudes observed in the TOMS data are plotted in the form of a contour plot with contour intervals of 30 DU. The bottom left corner of each upper panel contains no data since those latitudes are poleward of the polar

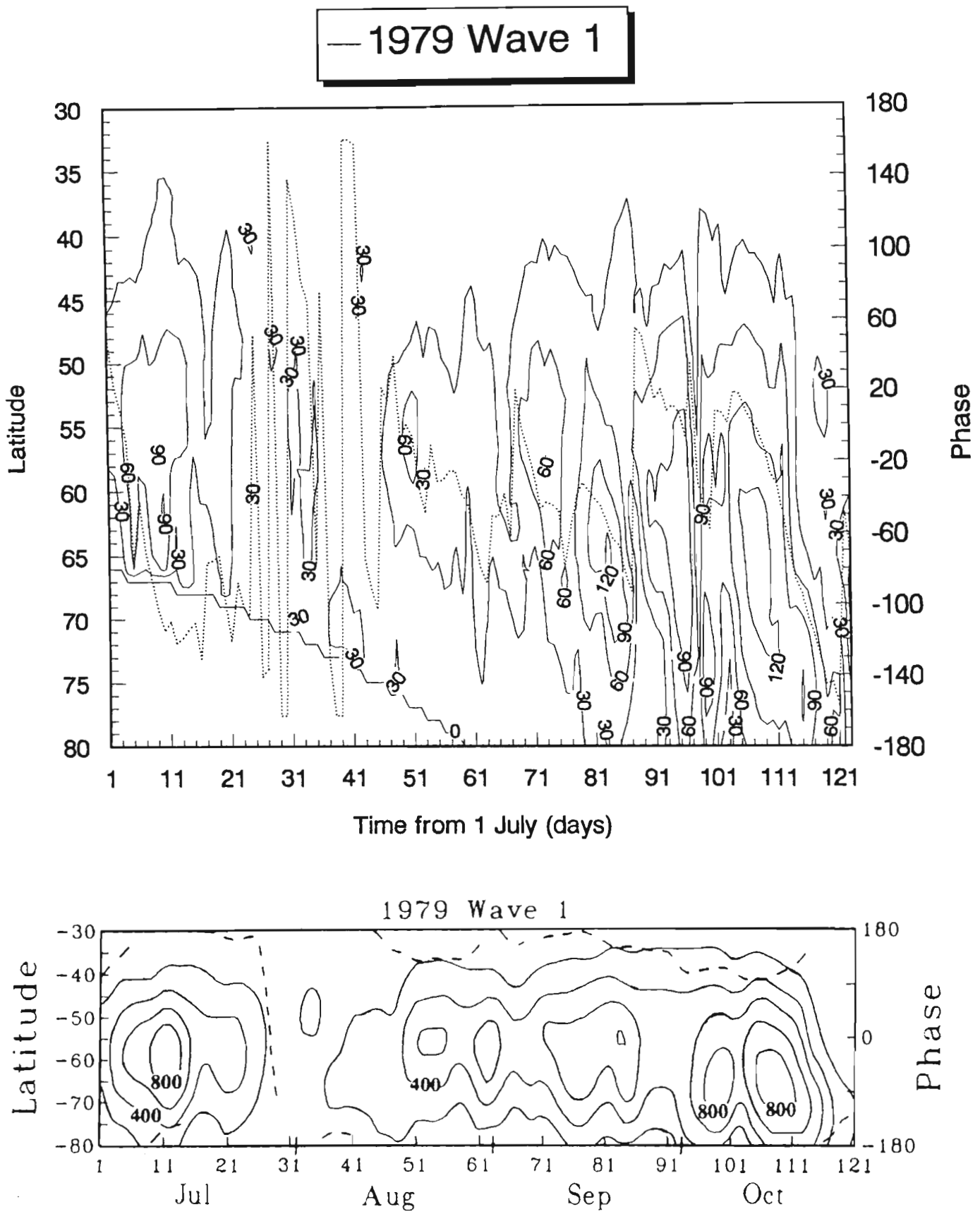


Figure 4.17: Time-latitude contour plots of wave 1 powers for July to October 1979. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

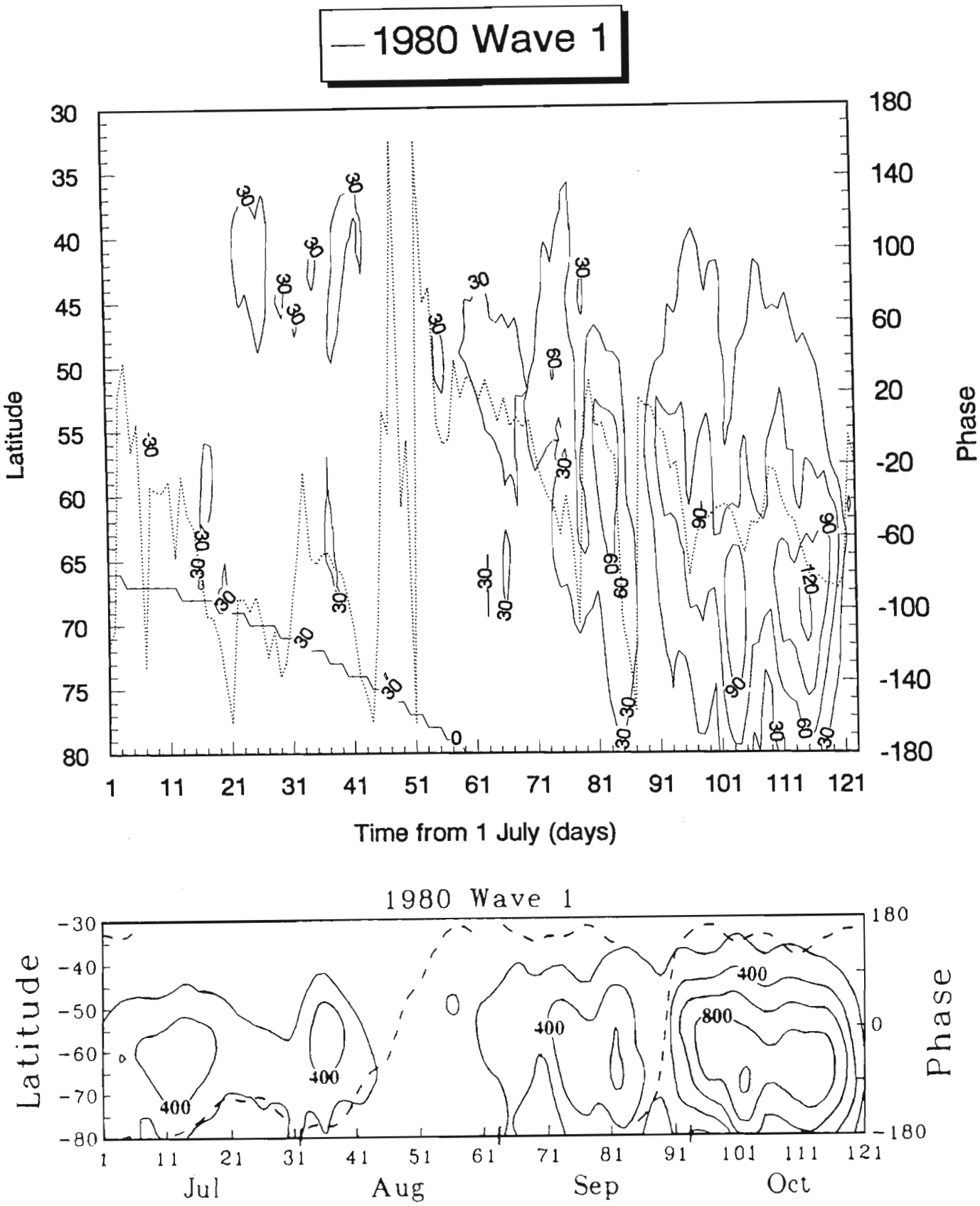


Figure 4.18: Time-latitude contour plots of wave 1 powers for July to October 1980. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

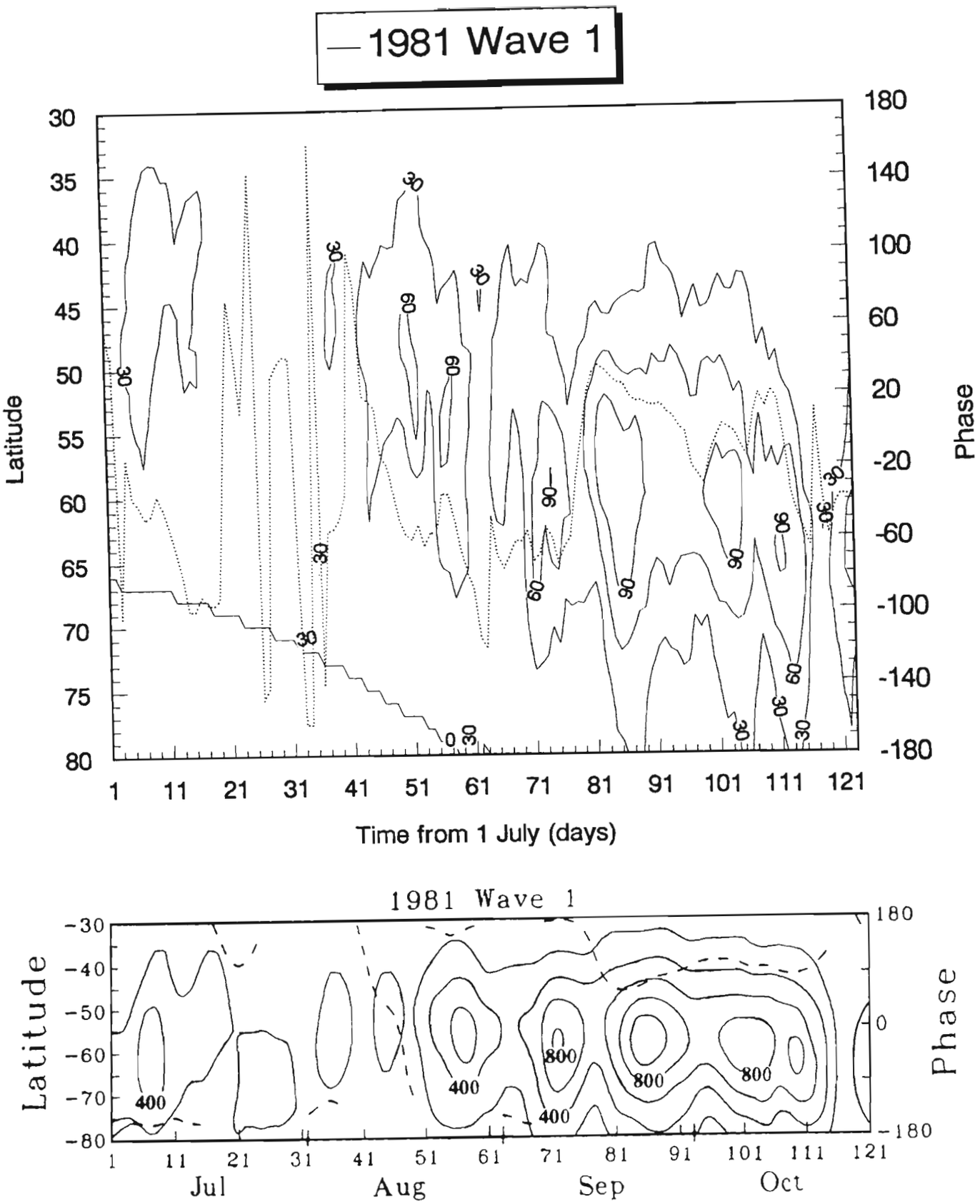


Figure 4.19: Time-latitude contour plots of wave 1 powers for July to October 1981. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

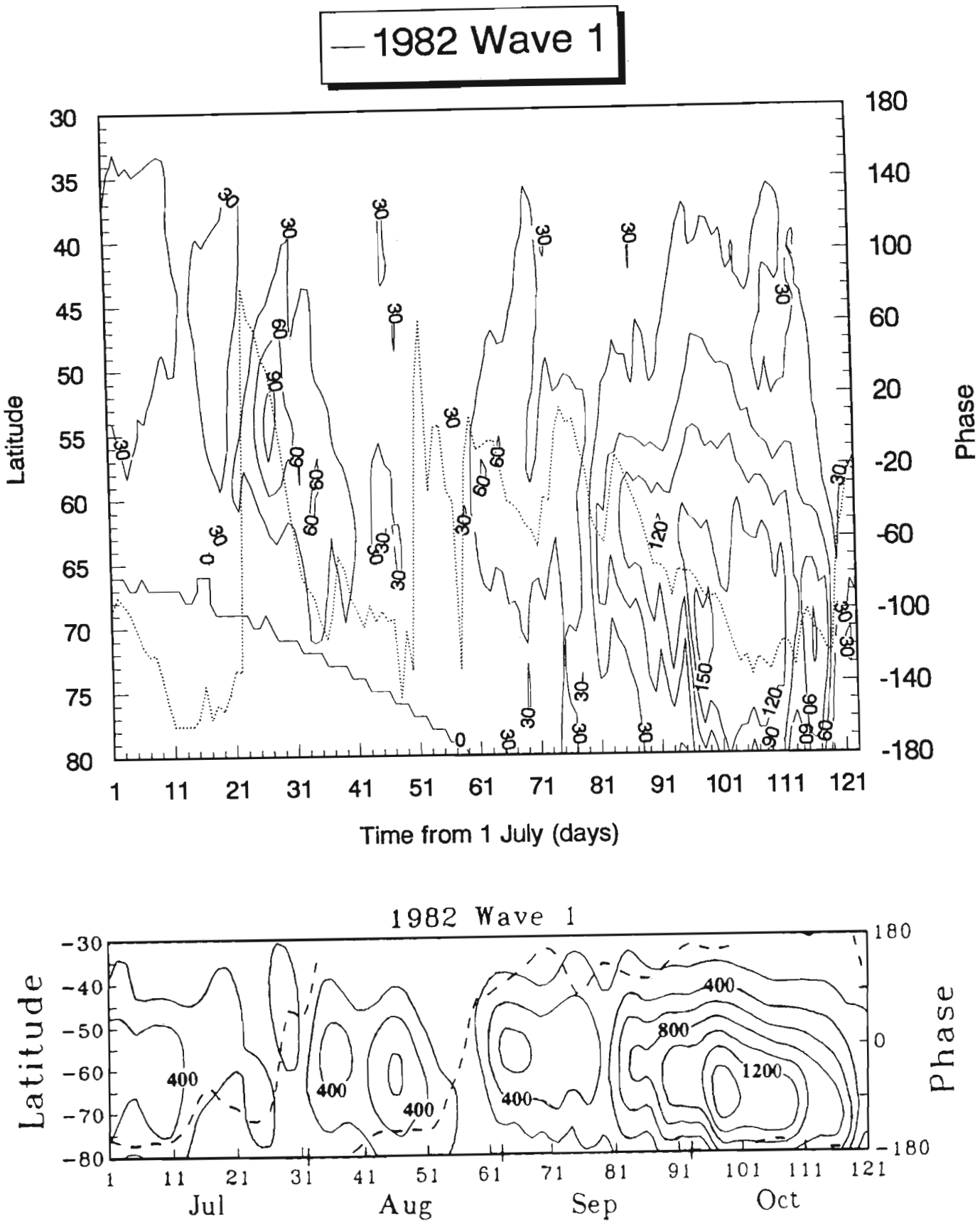


Figure 4.20: Time-latitude contour plots of wave 1 powers for July to October 1982. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

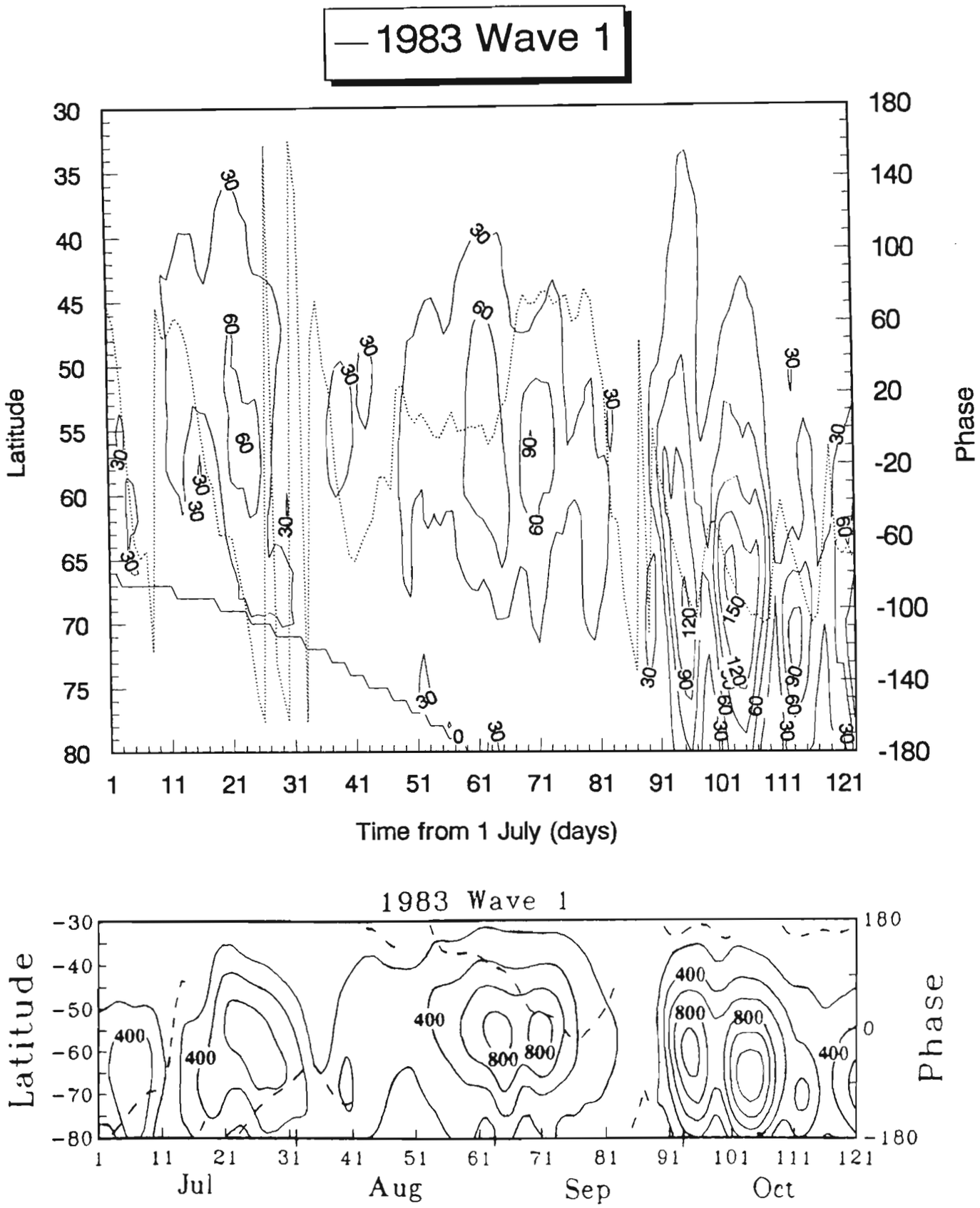


Figure 4.21: Time-latitude contour plots of wave 1 powers for July to October 1983. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

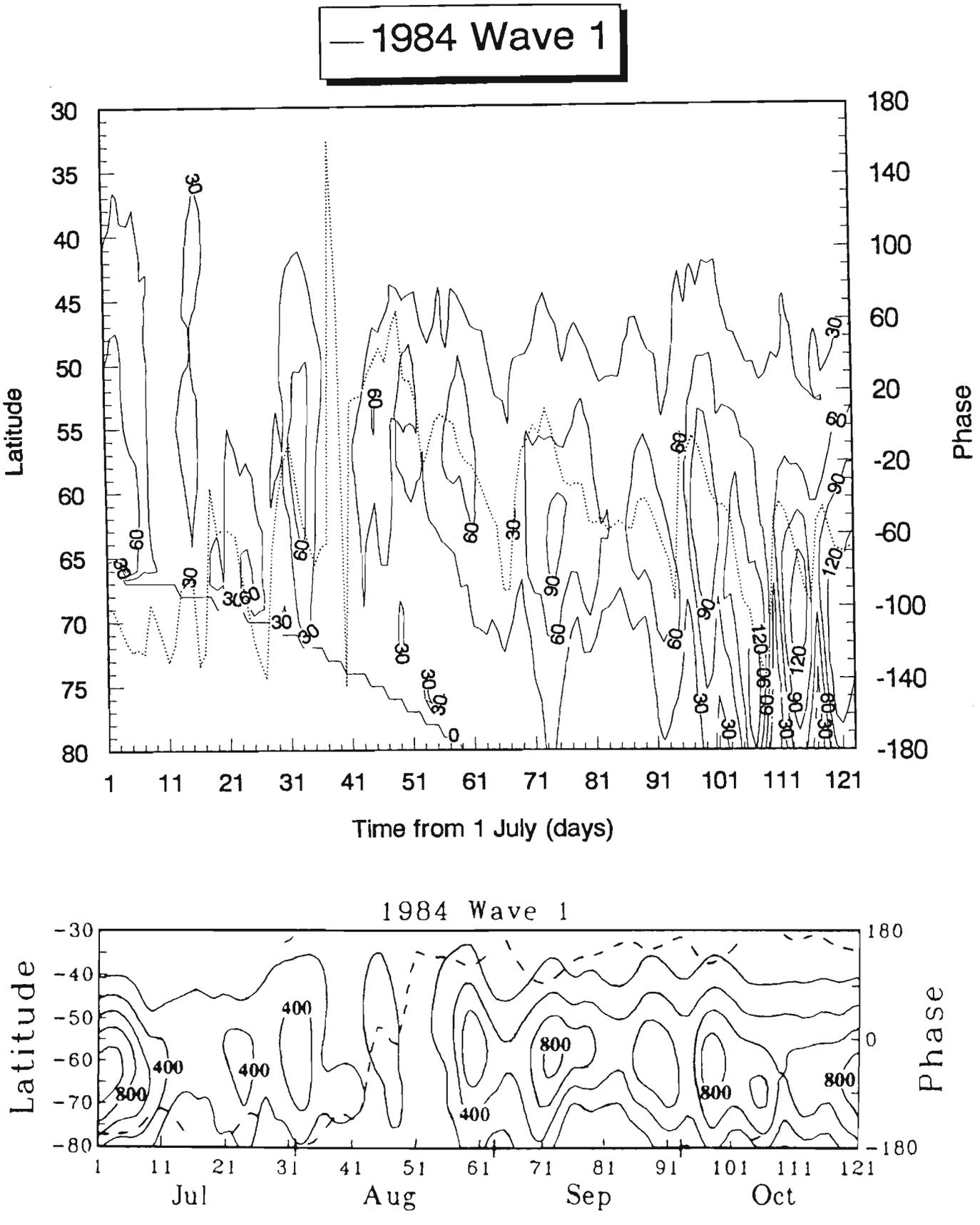


Figure 4.22: Time-latitude contour plots of wave 1 powers for July to October 1984. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

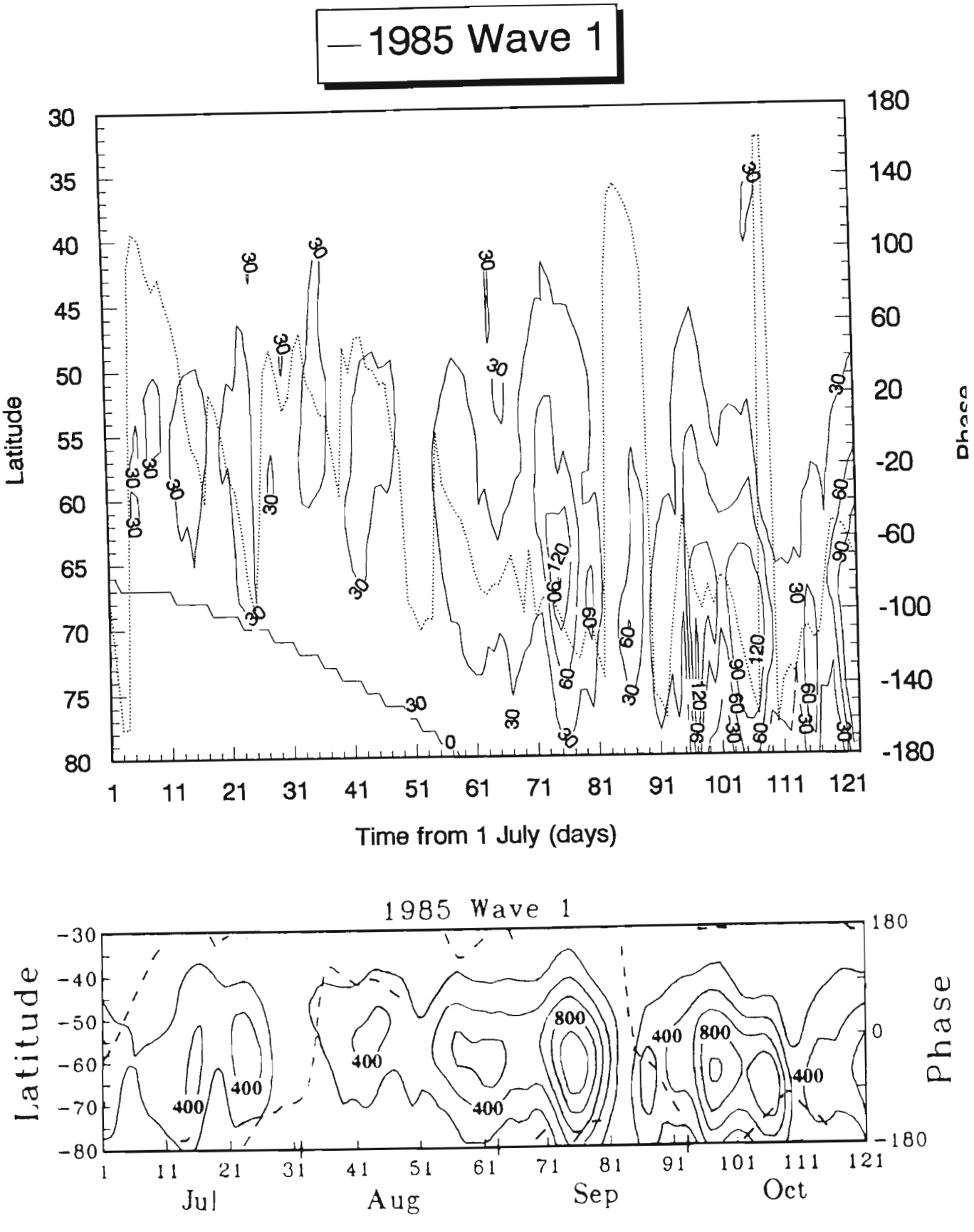


Figure 4.23: Time-latitude contour plots of wave 1 powers for July to October 1985. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

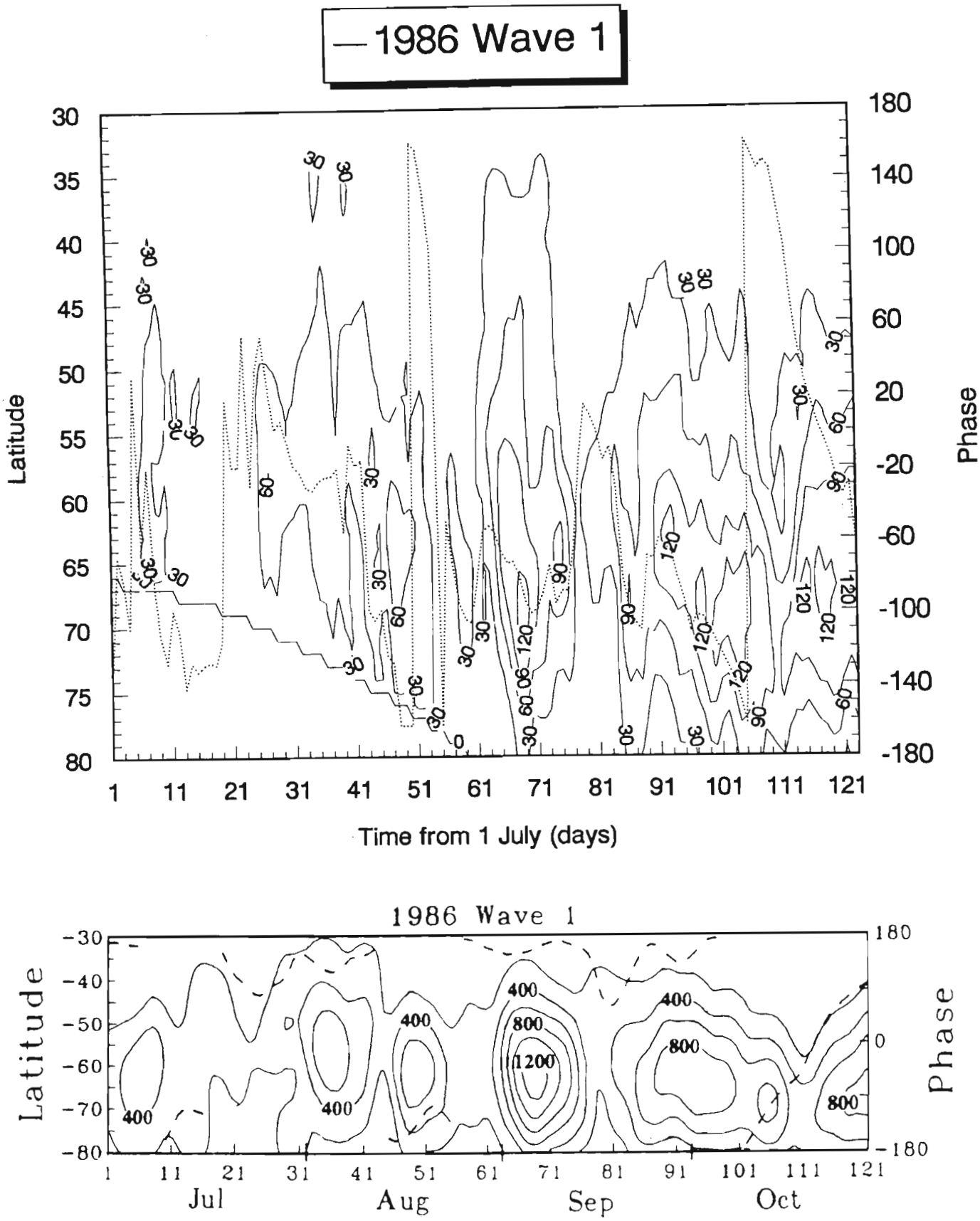


Figure 4.24: Time-latitude contour plots of wave 1 powers for July to October 1986. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

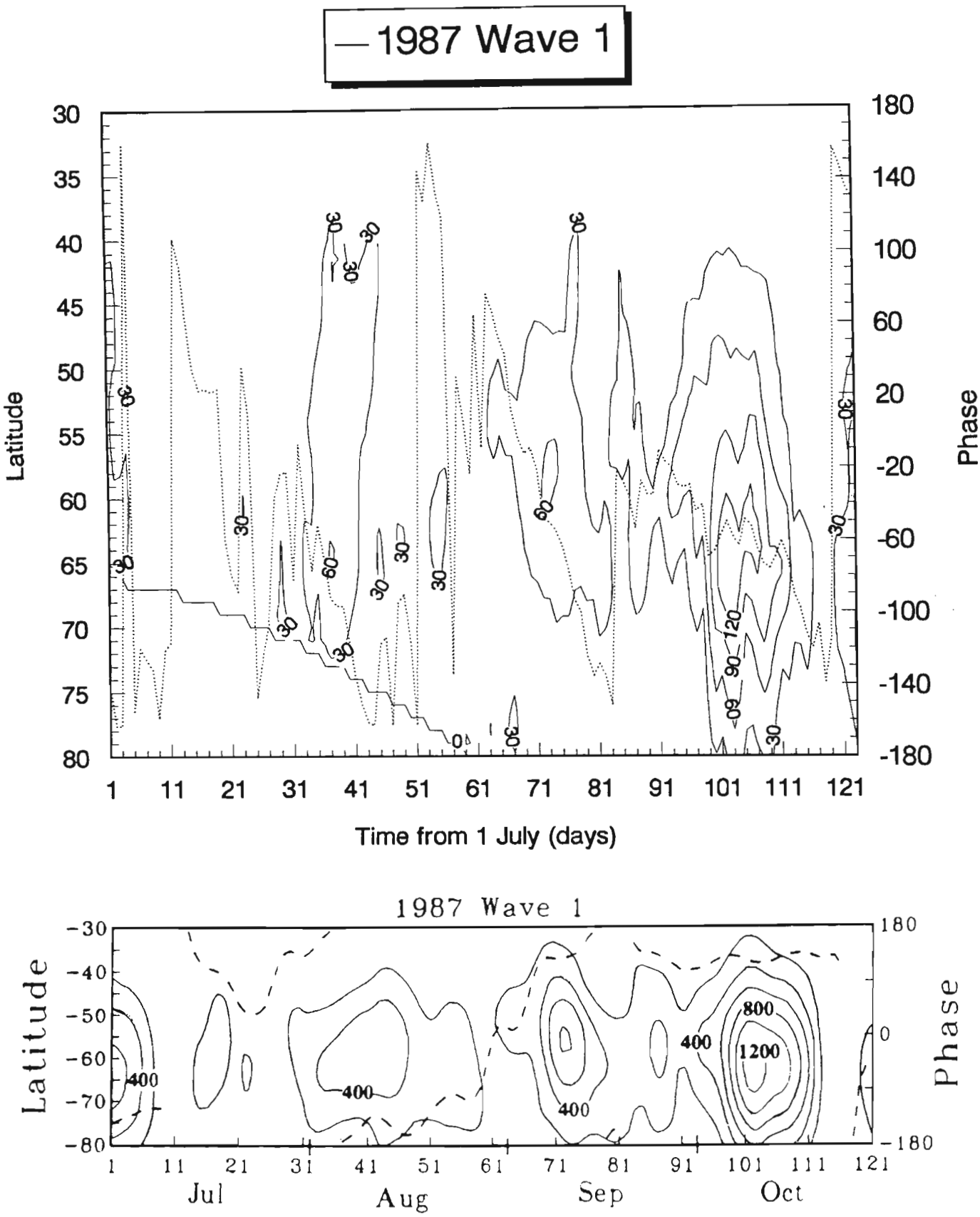


Figure 4.25: Time-latitude contour plots of wave 1 powers for July to October 1987. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

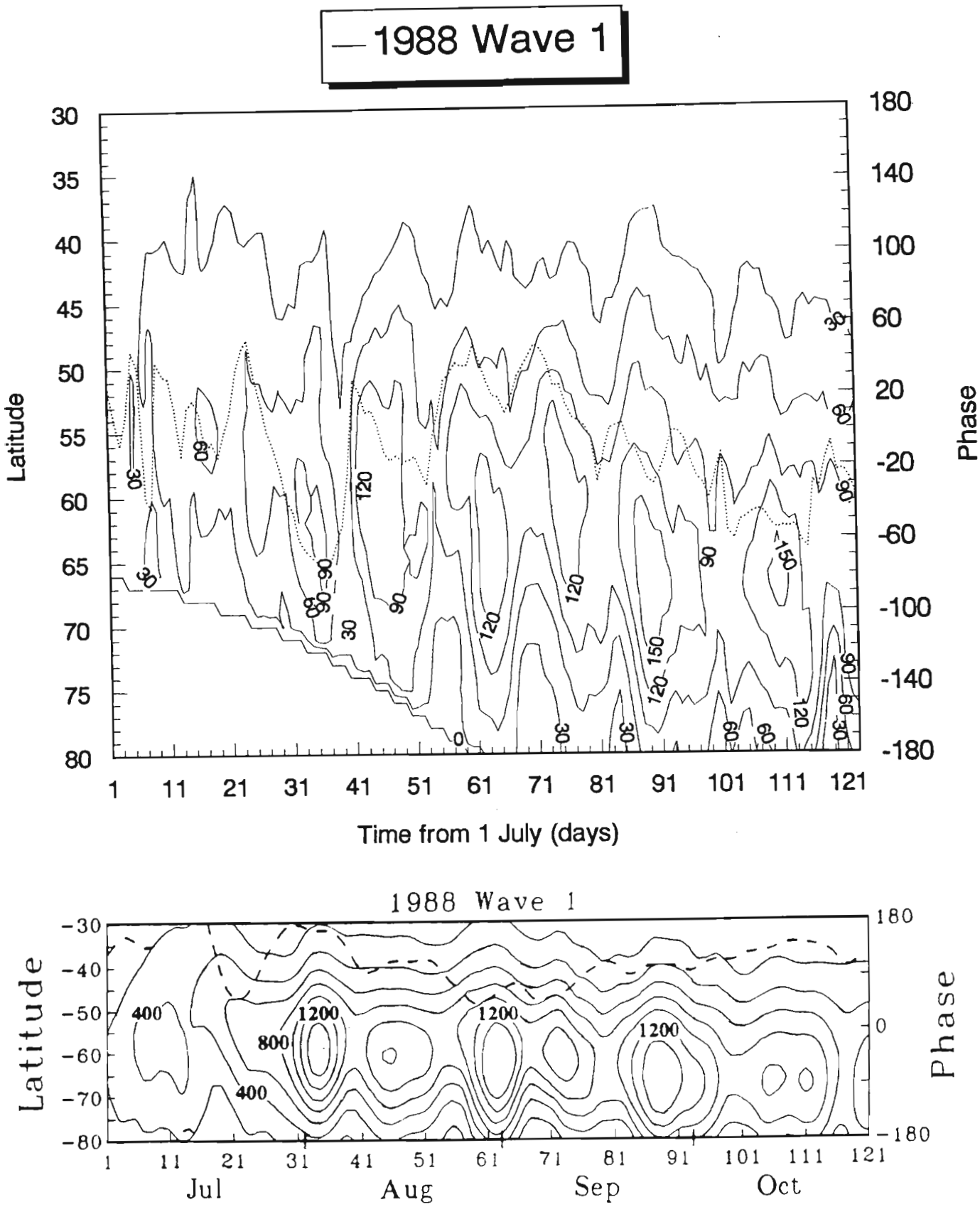


Figure 4.26: Time-latitude contour plots of wave 1 powers for July to October 1988. The upper panel shows wave amplitudes derived from TOMS data where the phase is plotted as a thin dotted line. The lower panel shows the geopotential height amplitude at 10 hPa over the same period where the phase is plotted as a dashed line, from Manney et al., (1991).

night terminator at that time of the year and hence no TOMS measurements are possible. For the data plots derived using TOMS data, the wave phase at the latitude of maximum amplitude (dotted line) is plotted from 180° to -180° on the right ordinate.

A similar contour plot technique is used to display the planetary wave amplitudes observed in the geopotential height data of MFM. In each lower panel, the contour interval is 200 geopotential meters. The phase is plotted as a dashed line. MFM separated geopotential height fields into cosine and sine components of zonal harmonics 0 to 6 with harmonic 0 representing the zonal mean. Time series of the zonally asymmetric fields (wavenumbers 1 to 6) were analyzed by calculating power and phase spectra using a standard FFT routine at each latitude. This technique of phase analysis has the opposite sign convention to the technique used in the TOMS data analysis. Hence for the TOMS data study, decreasing (increasing) phase indicates eastward (westward) movement whereas in the case of MFM, decreasing phase is interpreted as westward movement.

**Analysis of wave amplitudes**

The general morphology in each of the upper and lower panels is similar with good peak-for-peak correlations in latitude and time. Part of this correlation may be understood in terms of the increased height of the atmospheric column resulting in increased total column ozone. High correlations between ozone and geopotential height fields have long been recognized (Reed, 1950). At 10 hPa the geopotential height is typically 30 km (Houghton, 1989). For a wave of amplitude 1200 m (approximate maximum wave amplitude in the lower panel of Figure 4.26) this corresponds to a 4% increase in column height. The height for the same total column ozone wave in the associated TOMS data set is approximately 120 DU, which is 38% of the mean of 315 DU. This disproportionately increased ozone column arises for two reasons:

- Planetary wave amplitudes increase with altitude through the lower stratosphere as do ozone concentrations. Hence the planetary wave induced compression of the air column is greater in the stratosphere where ozone mixing ratios maximize. This causes a significantly stronger anomaly in the total column ozone (Salby and Callaghan, 1993).

- Vertical motion associated with the wave may move ozone rich air from the middle stratosphere into the lower stratosphere.

Two dimensional spatial correlation would provide a quantitative measure of the agreement between these two data sets but has not been performed in this study as the MFM data were available only in graphical format. In particular 1987 is shown to have very small planetary wave 1 amplitudes as seen in both panels of Figure 4.25. Wave powers remain low until the second week in October when a strong wave event of 1200 m amplitude occurs. This is mirrored by a 120 DU amplitude wave event in the TOMS ozone distribution. An earlier, smaller event of 800 m amplitude (60 DU) occurs near day 71. For 1988 (Figure 4.26) the wave amplitude morphology is vastly different. Six large wave events are visible in both the TOMS and MFM wave amplitude distributions. The characteristics of these events are listed in Table 4.1 where the maximum wave amplitude (in DU for TOMS and geopotential meters (gpm) for MFM), latitudinal position of the maximum, and day number of the maximum are given for both the TOMS data derived plots and the plots obtained from MFM. Note that for the MFM data, the results listed in

| TOMS derived data | | | MFM data | | |
|---|---|---|---|---|---|
| Amp. (DU) | Position (°S) | Day no. | Amp. (gpm) | Position (°S) | Day no. |
| 99 | 66 | 36 | >1400 | 58 | 34 |
| 121 | 59 | 43 | >1200 | 60 | 46 |
| 140 | 63 | 63 | >1400 | 62 | 63 |
| 147 | 62 | 76 | >1200 | 60 | 72 |
| 176 | 68 | 88 | >1400 | 66 | 89 |
| 159 | 67 | 111 | >1000 | 67 | 109 |

Table 4.1: The maximum wave amplitude, the latitudinal position and the day number of the 6 wave events observed in Figure 4.26. Characteristics for the TOMS derived data are obtained from the upper panel while MFM data are obtained from the lower panel.

Table 4.1 may not be accurate since they were visually extracted from the MFM plot. Both the temporal and spatial characteristics of the wave events are shown to compare well in spite of the inaccuracy of the data reduction from the MFM plot and the fact that a completely different algorithm has been used to derive the wave characteristics in the TOMS data. The correlation between wave amplitudes appears to be weak but may again result from poor data reduction from the MFM

plot. The wave amplitude derived from the TOMS data for the first wave event (99 DU) is significantly lower than would be expected when compared with the remaining data.

From Table 4.1 and both panels of Figure 4.26, it can be seen that the last two wave events are centered significantly more poleward than previous events. It will be shown below that this poleward motion is closely associated with changes in the phase of the wave and the spatial coverage of the Antarctic ozone hole.

### Analysis of wave phases

Due to differences in the sign convention for wave phases of MFM and wave phases observed for TOMS wave structures, direct comparison of these data is difficult. However, interpretation of the phase data in these two data sets should lead to similar conclusions. A discussion of the phases for the total column ozone planetary waves is given below, while an interpretation of the two data sets is given in section 4.4.3.

The daily phase data obtained from the TOMS data analysis are shown in Figure 4.27 (small stretched crosses) for each year from 1979 to 1988. Also shown on each plot is the daily latitudinal position of the maximum wave amplitude (solid line). The latitude of maximum amplitude (between 30°S and 80°S) is plotted on the left ordinate in each panel while the time (in days after 1 July) is plotted on the abscissa and the phase (from 180° to -180°) is plotted against the right ordinate. In general, the phases show little variability indicating that the wave 1 modes seldom propagate zonally. However, occasional pronounced negative gradients are observed in the phases, indicative of eastward propagation of the wave. Examples of eastward propagation are seen near the end of the data series in 1979 and 1986, and between day number 51 and 81 in 1987. An algorithm has been developed to identify the most prominent wave propagation events in the 10 year data record. A 15 day data window, inside which linear regression coefficients are calculated, is passed over the 123 day data series for each year. When $r^2$ values exceed 0.95, the regression coefficients are retained. In this way the 8 best (as determined by high $r^2$ values) wave 1 propagation events have been identified. Furthermore, the derived regression coefficients allow the calculation of wave speeds in terms of degrees of longitude per day. To convert these speeds to $m\,s^{-1}$, the average latitude of the maximum wave amplitude over the 15 day

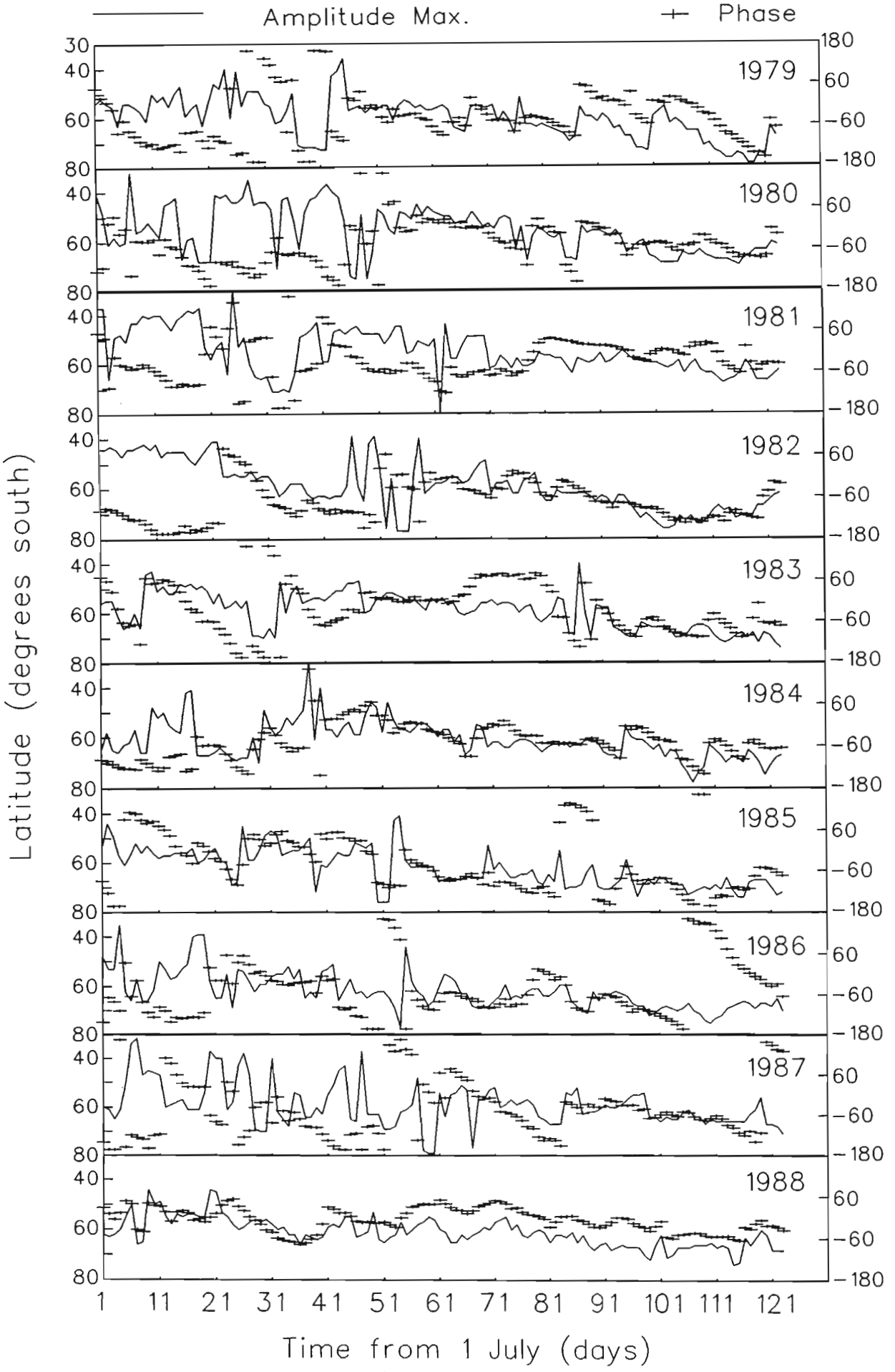


Figure 4.27: Wave 1 meridional positions and phases, July to October, 1979 to 1988. The latitude of the wave maximum is plotted against the left ordinate (solid line) while the wave phase is plotted against the right ordinate (stretched crosses).

period has also been recorded. The results are summarized in Table 4.2 which lists the year, the time period of the 15 day data window, the wave speed in degrees of longitude per day, the $r^2$ regression coefficient, the average latitude of the wave maximum and the wave speed in $m\,s^{-1}$.

| Year | Period (days after 1 July) | Wave Speed (degrees/day) | $r^2$ | Average Lat. of maximum | Speed ($m\,s^{-1}$) |
|---|---|---|---|---|---|
| 1979 | 107 to 121 | 12.15 | 0.9904 | 77 | 3.512391 |
| 1981 | 82 to 96 | 2.43 | 0.9594 | 60 | 1.560135 |
| 1982 | 23 to 37 | 15.35 | 0.9577 | 55 | 11.308509 |
| 1983 | 12 to 26 | 16.20 | 0.9829 | 49 | 13.655330 |
| 1986 | 91 to 105 | 6.72 | 0.9564 | 68 | 3.234973 |
| 1986 | 107 to 121 | 15.53 | 0.9659 | 68 | 7.475595 |
| 1987 | 67 to 81 | 10.64 | 0.9776 | 57 | 7.444493 |
| 1988 | 22 to 36 | 10.17 | 0.9565 | 57 | 7.119458 |

Table 4.2: Characteristics of 8 wave 1 propagation events observed in TOMS total column ozone distributions.

All events in this table involve eastward propagation. Other than the anomalous event of 1981 and 1986, wave speeds generally range between 10 and 15 degrees of longitude per day. This result has also been reported by Schoeberl et al. (1986). It was mentioned previously that the meridional position of the wave maximum tends to drift poleward from July to October of each year. Since the mapping of wave speed from degrees of longitude to $m\,s^{-1}$ varies with the cosine of the latitude, it might be expected that wave speeds in $m\,s^{-1}$ will be higher earlier in the period. The wave speed for the event late in 1979 is low as a result of the high latitudinal position of the wave maximum. Similarly wave speeds for events early in 1982 and 1983 are above the mean due to the low average latitude of the wave maximum. The two consecutive wave propagation events of 1986, occurring at the same average latitude, have very different speeds. Including the anomalously slow wave propagation event of 1981 it appears that wave speeds vary between $1.5\,m\,s^{-1}$ and $14\,m\,s^{-1}$.

Investigation of the latitudinal position of the wave maxima (solid lines in Figure 4.27) confirms that, in general, wave positions move poleward over the period July to October. This is related to a decrease in the area of the vortex. Table 4.3 lists the latitude, height and intensity of the polar night jet obtained from the

1979 to 1990 monthly mean NMC zonal wind fields (Randel, 1992). Over the

| Month | Latitude (°S) | Height (km) | Intensity ($m\,s^{-1}$) |
|---|---|---|---|
| July | 44 | >50 | ~90 |
| August | 58 | 37 | ~80 |
| September | 61 | 34 | ~70 |
| October | 65 | 28 | ~40 |

Table 4.3: **Monthly mean characteristics of the polar night jet obtained from NMC data (Randel, 1992).**

period of July to October the jet core decreases in altitude and moves poleward, indicative of the reduction in the area of the vortex. The jet maximum is centered at the latitude of maximum meridional temperature gradient (Chapter 2). With the onset of spring, the position of the maximum in the meridional temperature gradient is expected to move poleward as the Southern Hemisphere atmosphere warms. The effect of the decrease in vortex area on the total column ozone distribution is demonstrated in Figure 4.28.

Zonal mean ozone levels are shown by means of the colour code in the top right of the figure. Latitude is plotted on the ordinate while day number through the year is plotted on the abscissa. The decrease in the meridional extent of the vortex through the late winter, spring and summer results in a poleward drift in the midlatitude ridge of maximum total ozone values.

Note also in Figure 4.27 that at times there appears to be a correlation between the wave phase and its latitudinal position ie. eastward propagating waves move poleward. Since wave erosion of the vortex edge is known to decrease the area of the vortex, it is possible these propagating waves drift poleward as a result of this process. However, this is purely speculative and further investigation is required to confirm this as a plausible mechanism.

### 4.4.2 Wavenumber 2 behaviour

Analyses similar to that for wave 1 have been performed for wave 2. For conciseness, a full presentation of correlations between TOMS and MFM plots, as for wave 1, is not included. However, the correlations were again shown to be high. The results from the 10 TOMS wave amplitude plots have been summarized in

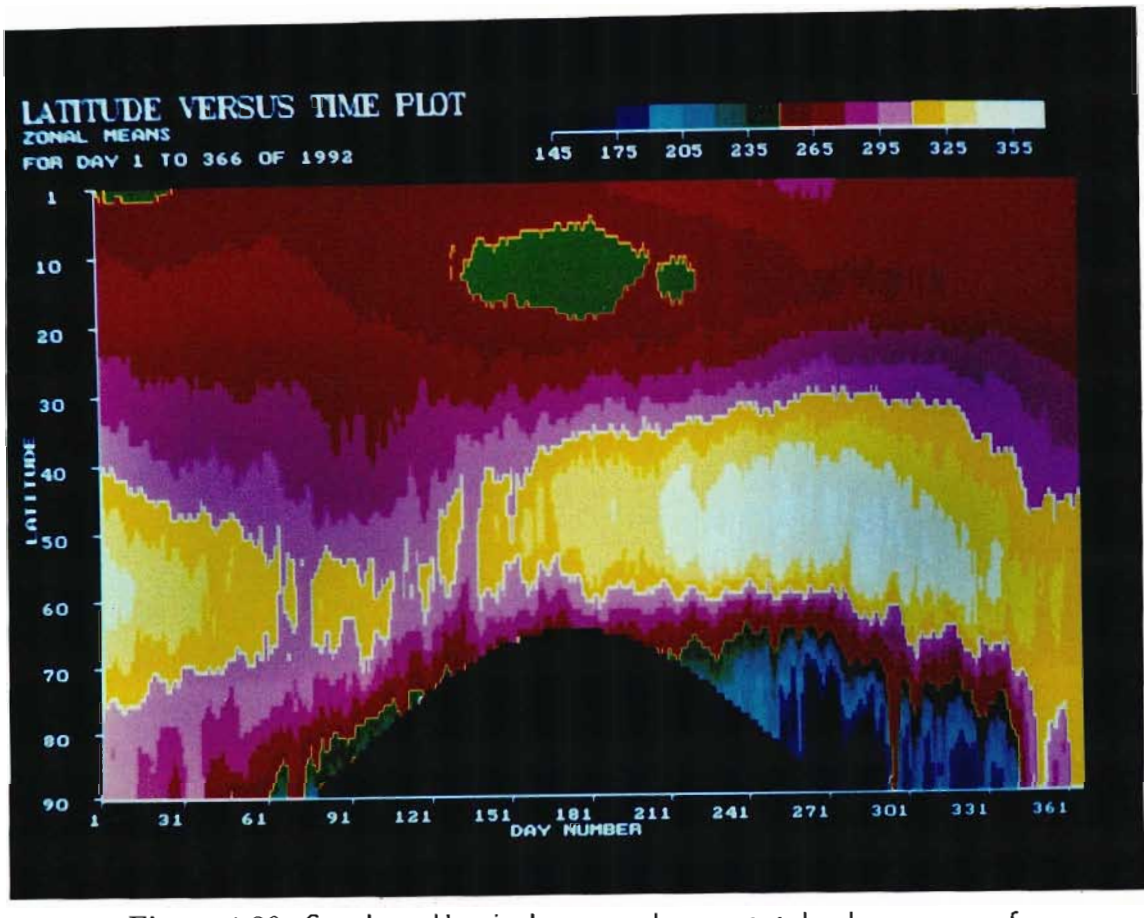


Figure 4.28: Southern Hemisphere zonal mean total column ozone for each day of 1992.

Figure 4.29 where the latitudinal position of the maximum amplitude of wave 2 is plotted together with the phase of the wave (as for wave 1 in Figure 4.27). Wave phases obtained from the analysis of the TOMS data cannot be compared directly with those obtained from MFM due to different phase conventions, although a comparison of the interpretation of the results will be given in the discussion (section 4.4.3).

In Figure 4.29 the phase signals vary between 0° and -180° as a result of the zonal mode 2 structure of the wave. In comparison with the wave 1 phases shown in Figure 4.27, there is considerably greater eastward propagation in the wave 2 mode. The algorithm used for wave 1 phase analysis has been applied to the wave 2 phase data to identify the 8 most prominent wave propagation events, as determined by the $r^2$ regression coefficient. In this case the regression analysis window was reduced to 12 days since many of the wave events circumnavigate half the globe over this period. The characteristics of the wave propagation events

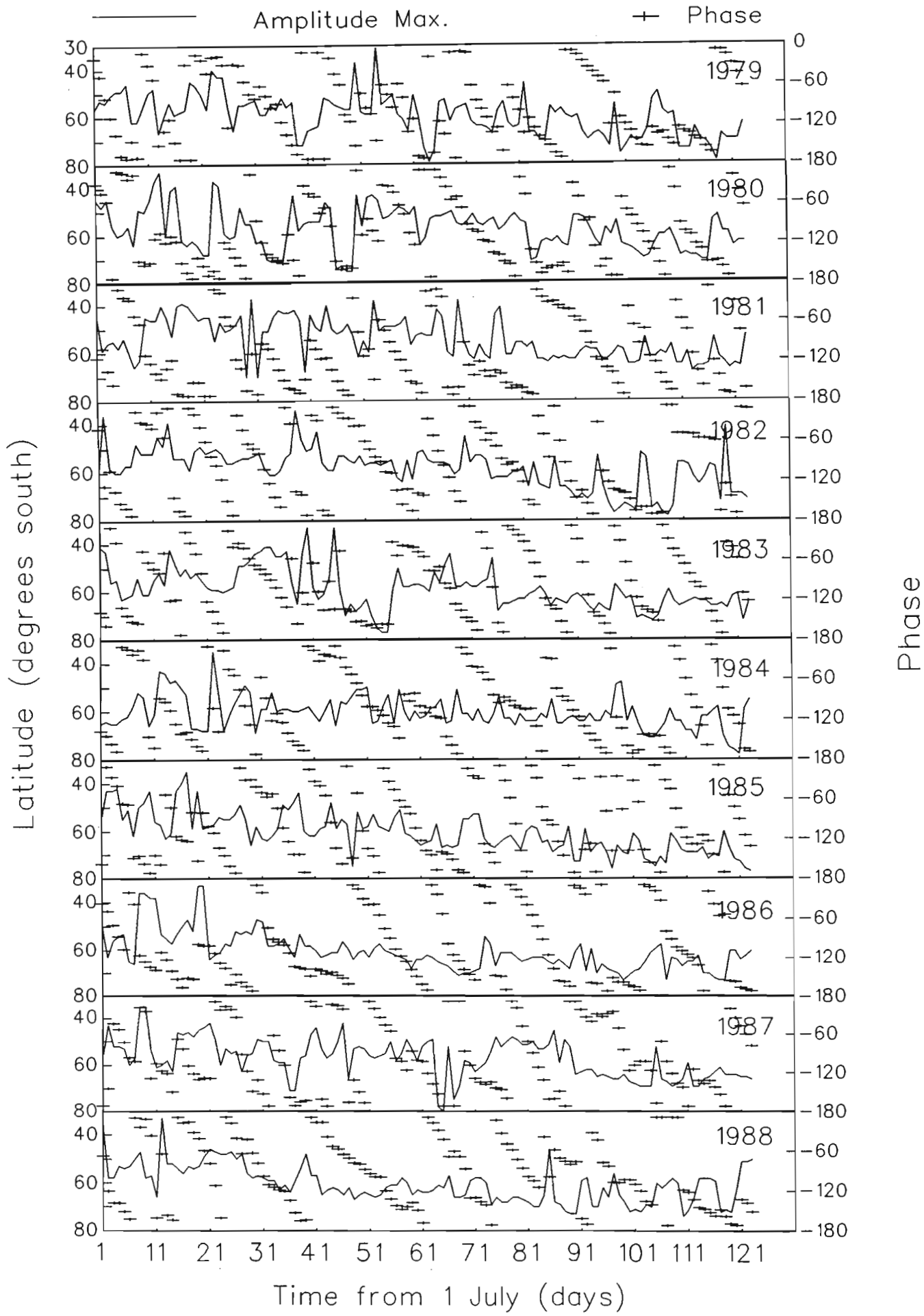


Figure 4.29: Wave 2 meridional positions and phases, July to October, 1979 to 1988. The latitude of the wave maximum is plotted against the left ordinate (solid line) while the wave phase is plotted against the right ordinate (stretched crosses).

are listed in Table 4.4.

| Year | Period (days after 1 July) | Wave Speed (degrees/day) | $r^2$ | Average Lat. of maximum | Speed ($ms^{-1}$) |
|---|---|---|---|---|---|
| 1980 | 65 to 76 | 11.38 | 0.9793 | 56 | 8.175013 |
| 1981 | 89 to 100 | 12.29 | 0.9850 | 59 | 8.131176 |
| 1983 | 25 to 36 | 9.62 | 0.9778 | 44 | 8.887717 |
| 1983 | 78 to 89 | 15.74 | 0.9783 | 58 | 10.716056 |
| 1983 | 107 to 118 | 14.02 | 0.9945 | 64 | 7.896573 |
| 1986 | 51 to 62 | 12.75 | 0.9888 | 62 | 7.687581 |
| 1986 | 78 to 89 | 15.92 | 0.9805 | 63 | 9.282669 |
| 1988 | 43 to 54 | 6.28 | 0.9736 | 67 | 3.151988 |

Table 4.4: Characteristics of the 8 best wave 2 propagation events observed in TOMS total column ozone distributions.

As for wave 1, all events are cases of eastward propagation. Including the low wave speed event in 1988, wave speeds vary between 3 and 11 $ms^{-1}$, similar to those of wave 1. It appears that although wave 2 speeds are similar to those for wave 1, wave propagation events are more common for wave 2. As for wave 1, investigation of the meridional position of the maximum wave amplitude (solid lines in Figure 4.29) indicates poleward drift over July to October of each year.

### 4.4.3 Discussion

Both the work of MFM and the analysis of wave structures observed in the TOMS data show that wave 1 is quasi-stationary with occasional eastward wave propagation events, while wave 2 propagation events are considerably more frequent. Using NMC data similar to that of MFM, Randel (1987) showed that waves 1 and 2 contribute the dominant transient geopotential height variance above 100 hPa. It was also shown that during the winter of 1983 wave 1 transient fluctuations at 10 hPa (averaged over 50°S to 60°S) are quasi-stationary or slowly westward moving while wave 2 moves regularly eastward at 24° to 36° of longitude per day (10 to 15 day period). There is still uncertainty as to why wave 2 shows eastward propagation while wave 1 remains quasi-stationary. It has been suggested (Randel, 1987; Randel et al., 1987) that wave 1 and 2 in the stratosphere result from upward propagation of wave 1 and 2 in the troposphere. However, this

mechanism may not explain the observed eastward propagation of wave 2 in the stratosphere.

MFM showed 1987 to be a year of extremely low wave 1 activity over the period 1979 to 1988. Analysis of TOMS data confirmed that this was the year of second lowest total wave power (exceeded only marginally by 1980) over the period 1979 to 1992. It was also shown in section 4.3 that 1987 was the year of most severe Antarctic ozone depletion when normalized to the stratospheric chlorine loading (see Figures 4.13 and 4.14). MFM also showed that 1988 was the only year when the wave 1 amplitude exceeded 1400 m in the 10 hPa geopotential height field, reaching an amplitude in excess of 1800 m. Analysis of TOMS data from 1979 to 1992 confirmed 1988 to be the year of highest total wave power and in addition showed this to be the year of smallest Antarctic ozone depletion. These results confirm that planetary wave activity has marked influences on the Antarctic ozone depletion.

These results suggest a number of possibilities for future work. MFM reported that while wave 1 is usually quasi-stationary, a number of instances where wave 1 moves eastward with wave 2 are observed, lasting from 4 to 10 days. Similar behaviour was observed in the wave 1 signals derived from the TOMS data as evident in Figure 4.27, although correlations between wave 1 and 2 propagation events were not considered. Chandra and McPeters (1986) showed that the wavenumber 1 pattern observed in October mean total column ozone from 1979 to 1982 was very stable in the south polar region except for a slight shift of the wave minima towards the east.

Furthermore, since variations in wave amplitude were ignored when analyzing the propagation events, future work should include the signal of maximum wave power in order to investigate how amplitudes vary with eastward propagation and poleward drift.

## 4.5 Comparisons between daily wave power and SANAE 100 hPa temperatures

It was shown in section 4.3.3 that there is a long-term correlation between SANAE 100 hPa temperatures and total wave power averaged over the AVP. Previous

analysis (Bodeker and Scourfield, 1993) has shown that day-to-day ozone and temperature correlations are highest at the 100 hPa level during the months of July to October. Day-to-day comparisons between SANAE 100 hPa temperatures and wave amplitudes are now examined over the period July to October. Again, due to space constraints, results are not presented for all years. The years of 1987 and 1988, experiencing particularly low and high wave powers respectively, have been used as examples.

The previous study has shown that the wave 1 amplitude is significantly higher than the wave 2 amplitude over the period July to October and therefore provides the predominant forcing of the Antarctic vortex. For this reason, only wave 1 amplitudes are considered in this study. Daily maximum wave 1 powers (between 30°S and 80°S) and wave phases, for 1987 and 1988, were extracted from the data sets used in the comparison with MFM data. The phase provides a measure of the longitudinal position of the wave maximum with respect to SANAE. Values close to 90° imply that the wave maximum is close to SANAE which is near the Greenwich meridian. It might be expected that under these circumstances, for strong wave amplitudes, the vortex will be distorted with its center pushed away from the pole and if the vortex edge is forced poleward of SANAE (latitude ~70°S) higher total column ozone and upper air temperatures should be measured. It has been suggested (Bowman, 1986) that large meridional gradients in total column ozone (such as those observed across the vortex edge) can produce local changes of more than 100 DU in less than a week in the presence of large amplitude planetary waves. To this end, daily TOMS data and upper air temperatures have been extracted for SANAE from July to October of 1987 and 1988. SANAE ozone data are not available for most of July as its location is poleward of the polar night terminator.

### 4.5.1 Wave 1 forcing over the AVP of 1987

Figure 4.30 shows a plot of daily maximum wave 1 amplitude, wave 1 phase and SANAE total column ozone for July to October 1987 while Figure 4.31 shows the same wave attributes and SANAE 100 hPa temperatures over the same period.

In both Figures 4.30 and 4.31 the maximum wave amplitude (solid line) is plotted against the left ordinate while the wave 1 phase (open circles) is plotted against the right ordinate. The time in days after 1 July is shown on the abscissa. In

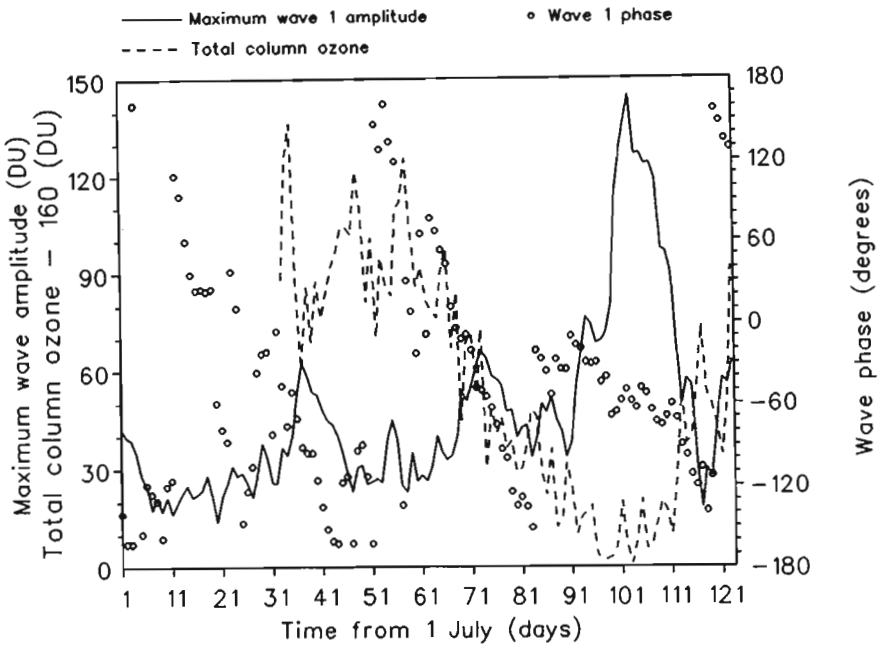


Figure 4.30: Daily maximum wave 1 amplitude, wave 1 phase and SANAE total column ozone for July to October 1987.

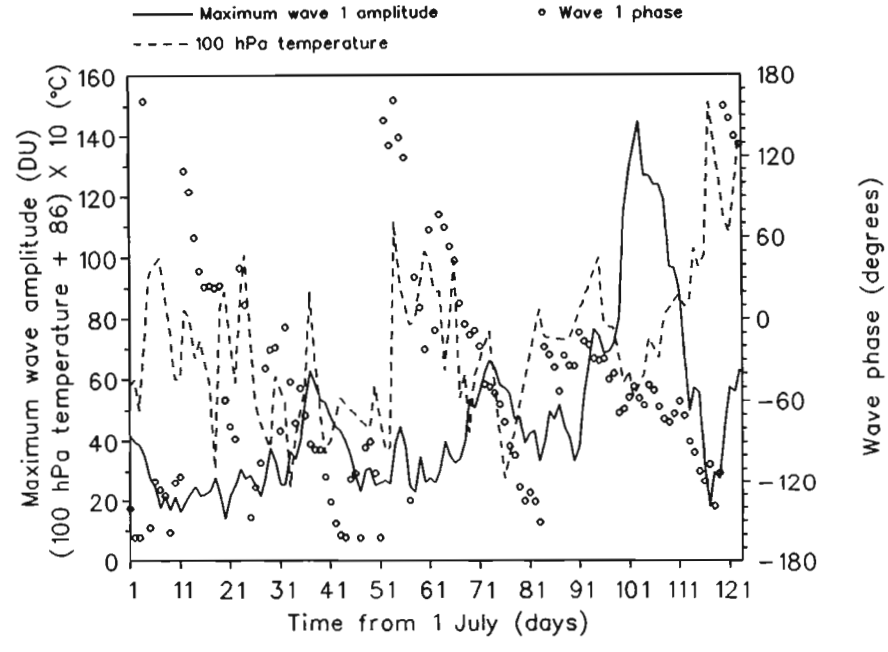


Figure 4.31: Daily maximum wave 1 amplitude, wave 1 phase and SANAE 100 hPa temperatures for July to October 1987.

Figure 4.30 the total column ozone, scaled by subtracting 160 DU, is also plotted (dashed line) against the left ordinate. Likewise, in Figure 4.31, the temperature is scaled by adding 86° C and multiplying by 10 before plotting it (dashed line) against the left ordinate.

Total column ozone values are greater than 220 DU (60 DU on left ordinate of Figure 4.30), the threshold for ozone depletion, up until day 69 (7 September) after which they decrease, remaining low through the rest of September and most of October. These elevated values through August may result either from little or no local chlorine photolysis and subsequent ozone depletion or from the fact that wave phases are between 90° to 180° during the later half of the month. Such phase values indicate that the center of the vortex may be being pushed away from South America (and away from SANAE) towards Australia and New Zealand. However, wave amplitudes are low over this period and elevated ozone levels are more likely the result of minimal chlorine catalyzed ozone destruction.

There are three events visible in Figures 4.30 and 4.31 which may demonstrate the wave forcing behaviour described above. Care must be taken when interpreting these data as it is possible that temporal changes in ozone are a direct consequence of local chemistry and not dynamics. The concomitant change in temperature is then just a consequence of reduced lower stratospheric heating by ozone. This is further complicated by feedback processes where rapid PSC formation associated with rapidly declining temperatures result in episodic ozone depletion. However, such rapid temperature changes are more likely the result of dynamics. Keeping these caveats mind, the three wave perturbation events are discussed below.

1) Amplitude enhancement: Day 32 to 41

Over these 10 days in early August, the wave amplitude increases rapidly from about 30 DU to 60 DU and the wave phase decreases from -6° on day 32 to -135° on day 42 (Figures 4.30 and 4.31). This event may force the vortex center away from the pole towards SANAE. The expected associated suppression in total column ozone is seen in Figure 4.30 (dashed line). However, lack of data prior to the event makes it impossible to determine if this is significantly below the background values. Although the 100 hPa temperatures appear to be suppressed through the first half of August (day 32 to 47 of Figure 4.31) a small increase in temperature is shown during the wave event. No immediate explanation can be found for this.

2) Wave propagation: Day 63 to 81

In Table 4.2 of section 4.4.1, it was shown that there was a strong wave 1 propagation event from day 67 to 81, moving eastward at 10.64° longitude per day. This event, clearly visible in both Figures 4.30 and 4.31 actually extends from day 63 to day 81. Over this period the phase changes from 76° to -128° showing that

the wave maximum moves from north-west of SANAE eastward until it is almost 180° longitude away from SANAE. Over this period it should be expected that ozone levels and temperatures above SANAE will decrease following the reduced proximity to the wave maximum and due to the vortex center being pushed closer to SANAE. The decrease in ozone and temperature during this wave propagation event is evident in Figures 4.30 and 4.31 respectively. Shortly after the event, temperatures rise again as the phase moves back to around -30° while ozone levels remain low, most likely as a result of chemical ozone depletion.

3) Large amplitude enhancement: Day 91 to 121

During October there is a very strong wave 1 event where the wave amplitude increases by more than 100 DU over 15 days as shown in Figures 4.30 and 4.31. The phase varies only slightly during this period, drifting from -40° to approximately -120°. In this configuration the wave maximum is south of Australia and New Zealand, on the opposite side of the Antarctic continent to SANAE. The expected low values of total column ozone and a small decrease in temperature are visible in Figures 4.30 and 4.31 respectively. The total column ozone is further suppressed during this period by local destruction through the chlorine cycles discussed in Chapter 1.

### 4.5.2 Wave 1 forcing over the AVP of 1988

Plots similar to those using 1987 data are shown in Figures 4.32 and 4.33 using data for 1988. Figure 4.32 shows the plot of daily maximum wave 1 amplitude, wave 1 phase and SANAE total column ozone for July to October 1988 while Figure 4.33 shows the same wave attributes together with SANAE 100 hPa temperatures. Note that during 1988, SANAE total column ozone levels at no time fell below the 220 DU threshold for depletion. Note also that the maximum wave amplitudes are significantly higher in 1988 than in 1987 (cf. Figures 4.30 and 4.32). Between day 61 and 121 there appear to be peak-for-peak correlations between wave phase and total column ozone ie. increases in phase (wave maximum moving closer to SANAE) correspond to increases in SANAE total column ozone. Individual forcing events due to wave propagation are not as common in 1988 as in 1987 due to the stationary nature of wave 1 during 1988. There is also a peak-for-peak correspondence between the temperature signal and the wave phase from July to October 1988 (see Figure 4.33).

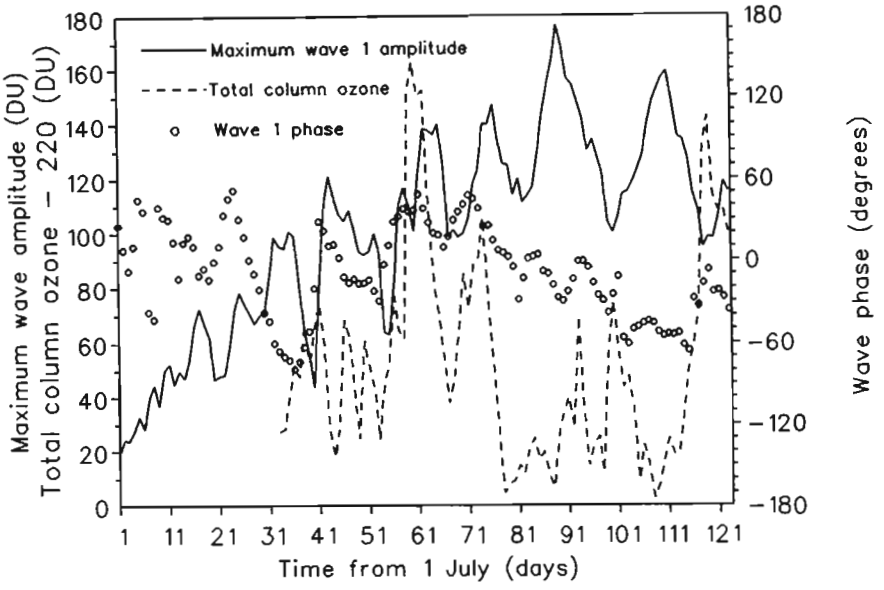


Figure 4.32: Daily maximum wave 1 amplitude, wave 1 phase and SANAE total column ozone for July to October 1988.

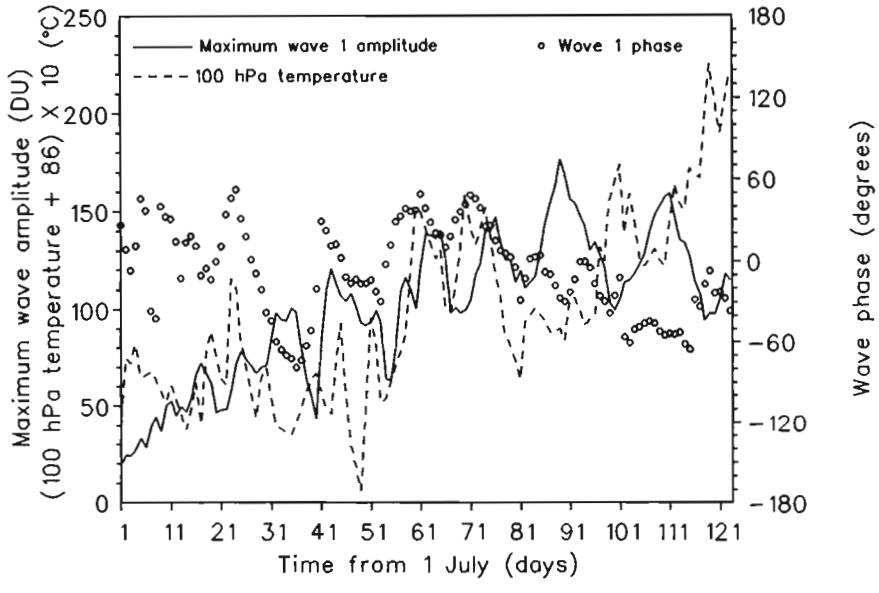


Figure 4.33: Daily maximum wave 1 amplitude, wave 1 phase and SANAE 100 hPa temperatures for July to October 1988.

This study has shown that when examining total column ozone and upper air temperatures from an Antarctic site, it is important to consider both the amplitude and phase of the wave 1 forcing.

## 4.6 Northern Hemisphere planetary wave morphology and its influence on the severity of Arctic ozone depletion

### 4.6.1 Northern Hemisphere wave activity: 1979 to 1992

A similar analysis to that described in section 4.3 for the Southern Hemisphere total column ozone distribution is now presented for Northern Hemisphere data to provide a long-term climatology of wave activity. Although no significant ozone hole forms over the Arctic, analysis similar to that performed for the Antarctic is repeated to determine if the expected enhanced planetary wave activity in the Northern Hemisphere is instrumental in preventing the formation of a 'Southern Hemisphere type' ozone hole.

Wave powers for the first 6 wave modes have been extracted for each day from 1979 to 1992 from the Northern Hemisphere TOMS data. To investigate their mean annual variation, averages for each calendar day over this period have been calculated. The first four wavenumber annual mean curves are plotted against day number in Figure 4.34.

To restrict the study to wintertime behaviour, it was decided to select the period from the end of October (day 300) to the spring equinox (21 March – day 80) as the Northern Hemisphere equivalent of the Southern Hemisphere AVP. Wave powers averaged over this period provide a measure of annual winter wave activity and may be used to assess the affect of planetary wave activity on a Northern Hemisphere ozone hole formation. It can be seen in Figure 4.34 that wave powers maximize over this period and, as for the Southern Hemisphere, wave power decreases rapidly with increasing wavenumber. Furthermore, in comparison with the analogous plot for the Southern Hemisphere (Figure 4.7), it appears that powers for wave 1 and 2 are generally lower and powers for wave 3 and 4 are higher in the Northern Hemisphere. Note in Figure 4.7 that wave powers in the Southern Hemisphere maximize over a small time period while wave powers in the Northern Hemisphere (see Figure 4.34) remain close to the maximum value for a longer period of the year.

These qualitative observations are confirmed in Figure 4.35 where histogram distributions of the wave power signals of Figures 4.7 and 4.34 are compared. The

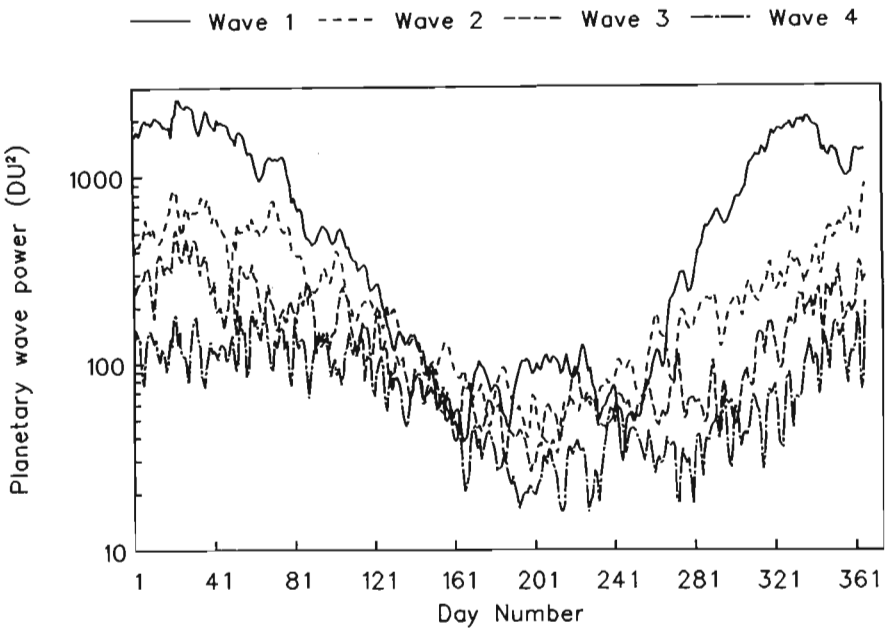


Figure 4.34: Northern Hemisphere planetary wave powers in $(DU)^2$ for wavenumber 1 to 4 averaged for each calendar day over the years 1979 to 1992.

calendar day averaged wave powers for waves 1 to 4 were binned into 100 $(DU)^2$ intervals. The number of data points in each bin is plotted on the ordinate in each panel of Figure 4.35 while the center power for each bin is plotted on the abscissa. Other than for wave 1, Northern Hemisphere wave powers tend to be higher than those in the Southern Hemisphere and show a greater spread in the histogram distributions of Figure 4.35. Furthermore, although the maximum wave 1 powers are higher in the Southern Hemisphere than in the Northern Hemisphere, they occur on very few days. In summary, Figure 4.35 shows Northern Hemisphere total ozone wave powers to be higher than those in the Southern Hemisphere, particularly for waves 3 and 4.

Daily Northern Hemisphere wave powers for the first 4 wave modes as well as the total wave power of the first 6 wave modes have been averaged over day 300 to day 365 and day 1 to day 80 for each year to provide an annual measure of the power in waves 1 to 4 and the total wave power. The results are plotted in Figure 4.36. The annual total wave power (thick solid line) shows considerable variability with 1986 showing maximum total wave power and 1990 showing minimum wave power. The year of 1985 appears unusual in that wave 1 power is lower than the mean while the power in wave 2 is considerably higher than the mean. Comparison of 1985 and 1990 data shows that although wave 1 power was lower

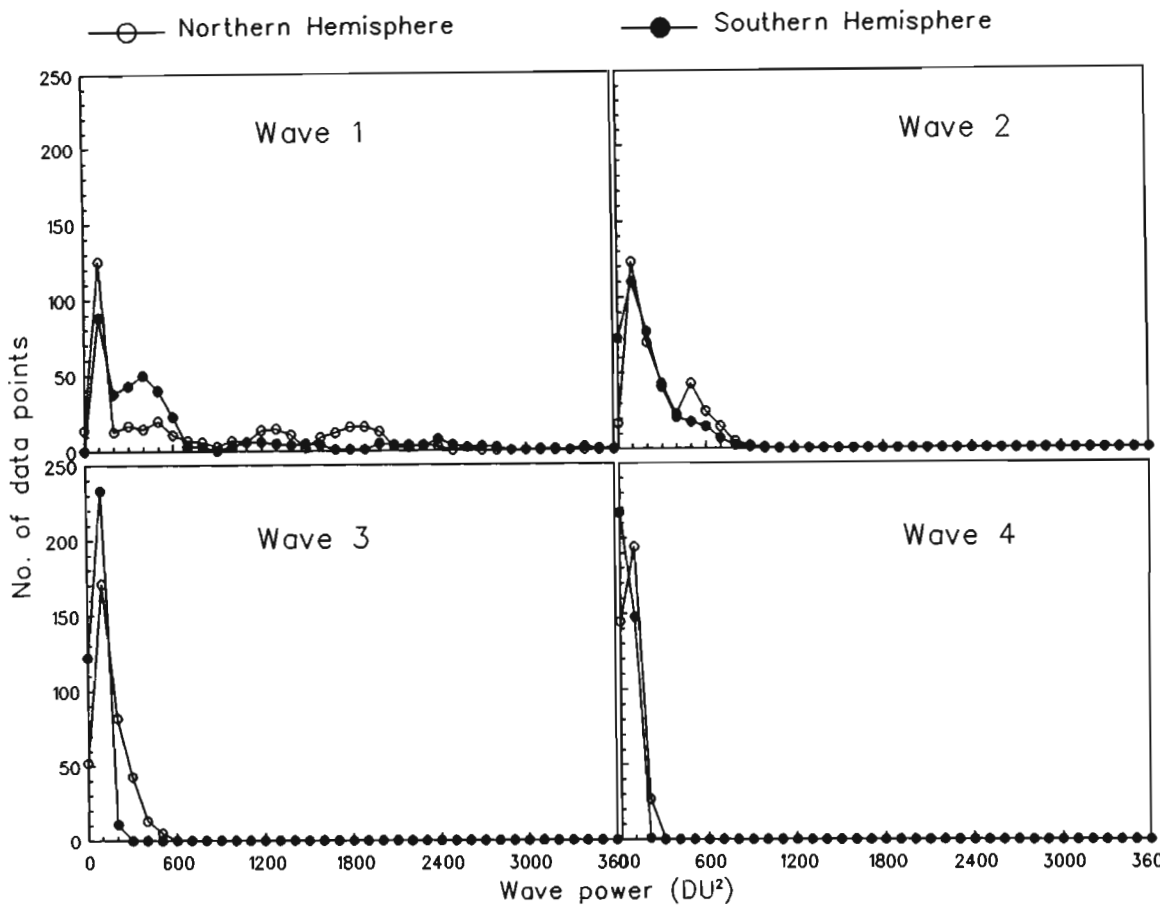


Figure 4.35: Histogram distributions of calendar day averaged wave power signals for the first 4 wave modes for the Northern and Southern Hemispheres.

in 1985, enhancements in wave 2 power result in the total power being higher than that for 1990. This indicates that care must be taken when selecting which wave mode to use as a parameter for planetary wave activity. In the following study, the total wave power data have been used.

### 4.6.2 Comparison of total wave power and Arctic ozone depletion

As for the Southern Hemisphere, a daily depleted ozone mass signal has been obtained where the same 220 DU threshold for depletion has been used. An example is shown in Figure 4.37 where the 1988 daily total wave power (modes 1 to 6) is plotted as a solid line and the 1988 daily depleted mass is plotted as a dashed line.

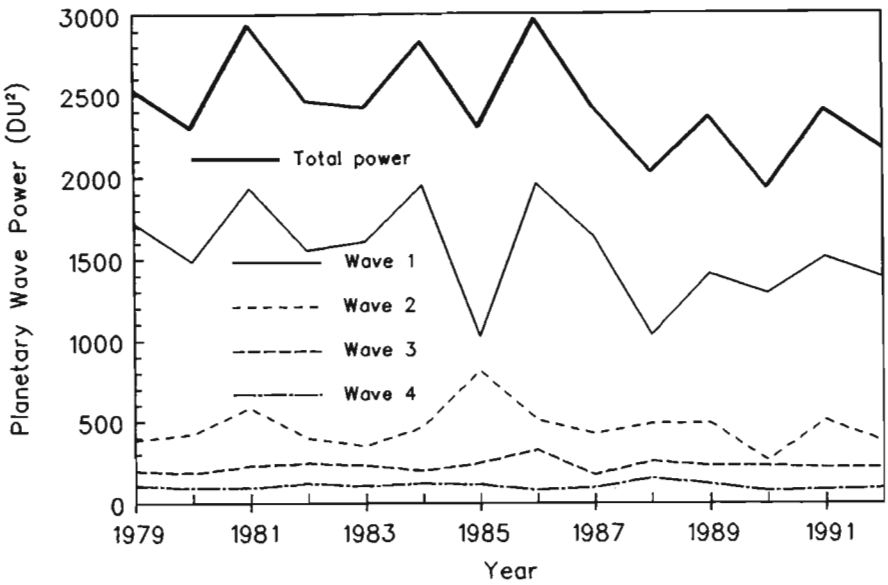


Figure 4.36: **Northern Hemisphere wave powers averaged over day 300 to day 80 for wavenumber 1 to 4. The total of the first six wave powers is also shown.**

The daily total wave power maximizes during the winter and early spring (day 41 to 81) while small isolated episodes of ozone depletion are observed in late fall and early winter. The episodic nature of the depleted mass signal masks any day-to-day correlations with the total wave power. The fact that the two signals maximize at different times of the year may weaken annual correlations of ozone depletion and wave activity, although temporal integration (day 300 to day 80) may alleviate this problem. The behaviour observed in Figure 4.37 differs from that of the Southern Hemisphere (cf. Figure 4.11) where the wave power signal and depleted mass signal maximize at the same time of the year and the depleted mass is less episodic in nature. Not only are depleted mass values more variable than in the Southern Hemisphere, but the daily Northern Hemisphere values are typically 20 times smaller than those observed in the Southern Hemisphere.

To prepare the data for comparisons of total planetary wave power and Arctic ozone depletion, both the depleted ozone mass data and the total wave power data have been averaged over the winter (after day 300 and before day 80) for each year from 1979 to 1992. Note that this may introduce some ambiguity since it is the second half of one winter that is being combined with the first half of the following winter. However, since the same time period is used for both signals, it is hoped that this will avoid inconsistencies. Figure 4.38 shows the

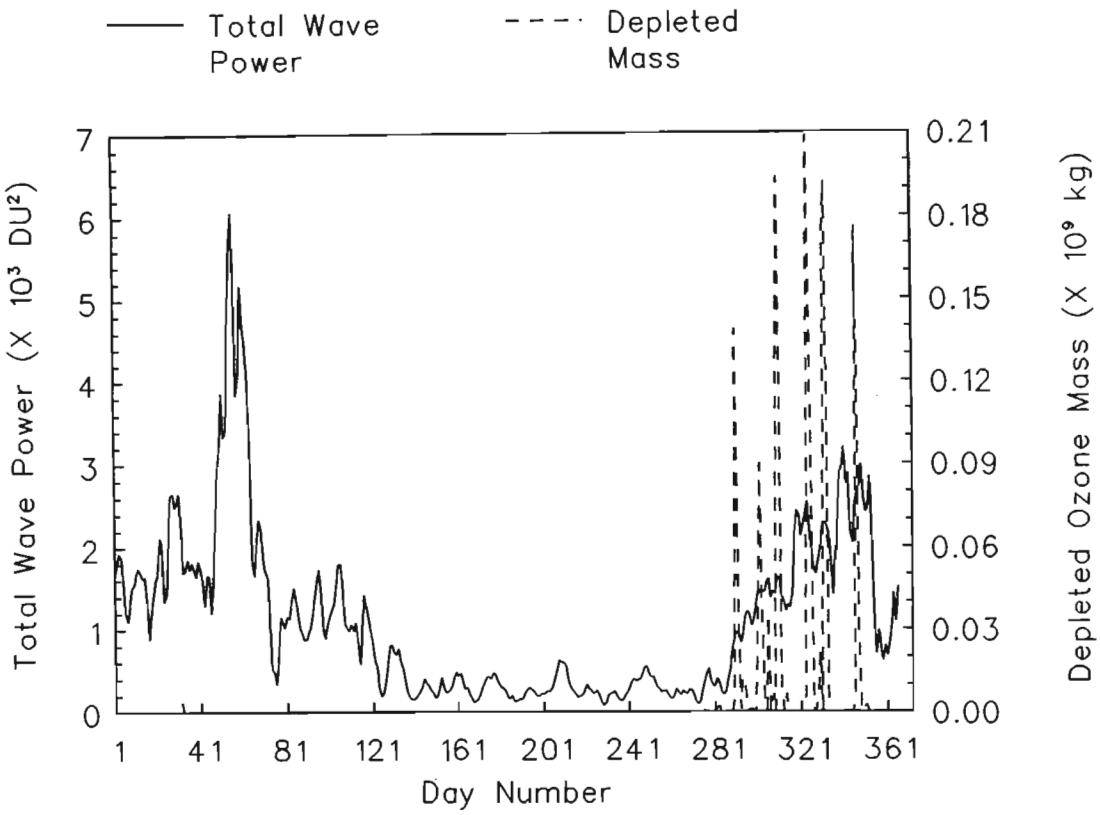


Figure 4.37: **Northern Hemisphere total planetary wave power (wave 1 to 6) and depleted ozone mass for 1988.**

averaged depleted mass values together with the stratospheric chlorine loading signal shown earlier in Figure 4.12.

An increase in Northern Hemisphere ozone depletion appears to follow the increase in stratospheric chlorine loading although the correspondence is not as well defined as for the Southern Hemisphere (cf. Figure 4.12). Note also that the winter averaged depleted mass values are of the order of 200 times smaller than Southern Hemisphere values. This results from Northern Hemisphere depletions being smaller in magnitude and more episodic in nature, as shown previously (Figure 4.37). Interannual differences are significant where it appears that 1985 and 1992 show anomalously high ozone depletion. Note also the quasi-biennial structure in the depleted mass signal over the period 1984 to 1991.

A linear normalization scheme has been used to remove the chlorine dependence from the depleted mass signal. The linear relationship between the depleted ozone mass and the stratospheric chlorine loading is given by

$$\text{Depleted Mass} = 23.33 \times \text{Chlorine Loading} - 61.86 \qquad (4.3)$$

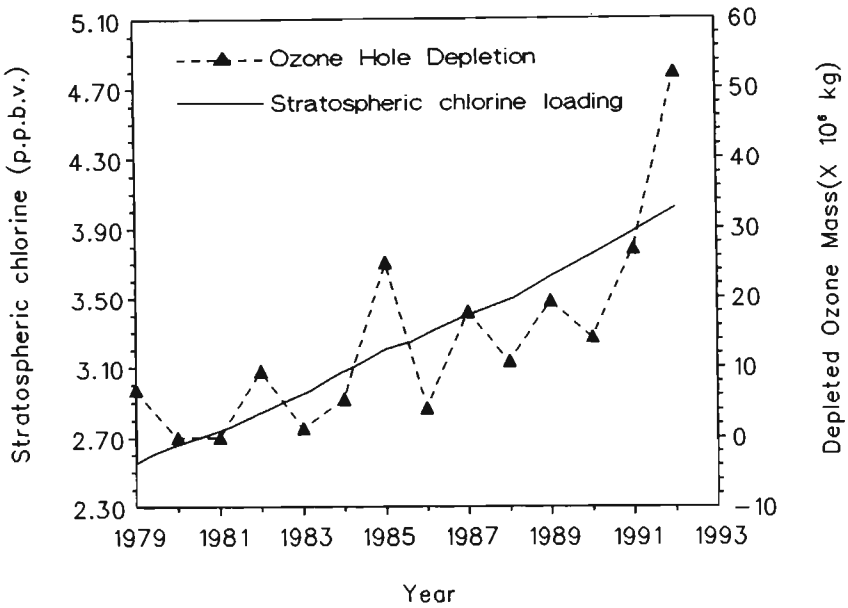


Figure 4.38: Northern Hemisphere winter depleted ozone mass and stratospheric chlorine loading. Chlorine data as cited in Solomon (1990) and Toohey (personal communication).

where the chlorine loading is measured in p.p.b.v. and the depleted mass is measured in kt (1 kt = $1 \times 10^6$ kg)

Depleted mass values predicted from this linear relationship (r = 0.77) are subtracted from the original data series to highlight anomalous behaviour. The results are displayed in Figure 4.39 together with the wintertime averaged total wave power.

Normalized depleted mass is plotted against the left ordinate using an inverted scale, while the integrated wave power is plotted against the right ordinate. An approximate peak-for-peak anticorrelation is observed between the two curves for short periods although the relationship is not as clear as for the Southern Hemisphere. Since the correlation coefficient between the two curves is so low (r = 0.31) it appears that interannual differences in Arctic ozone depletion do not depend on interannual differences in Northern Hemisphere wave activity.

### 4.6.3 Discussion

Previous studies (van Loon and Jenne, 1972) have reported that the amplitudes of forced stationary waves are much smaller in the Southern Hemisphere than

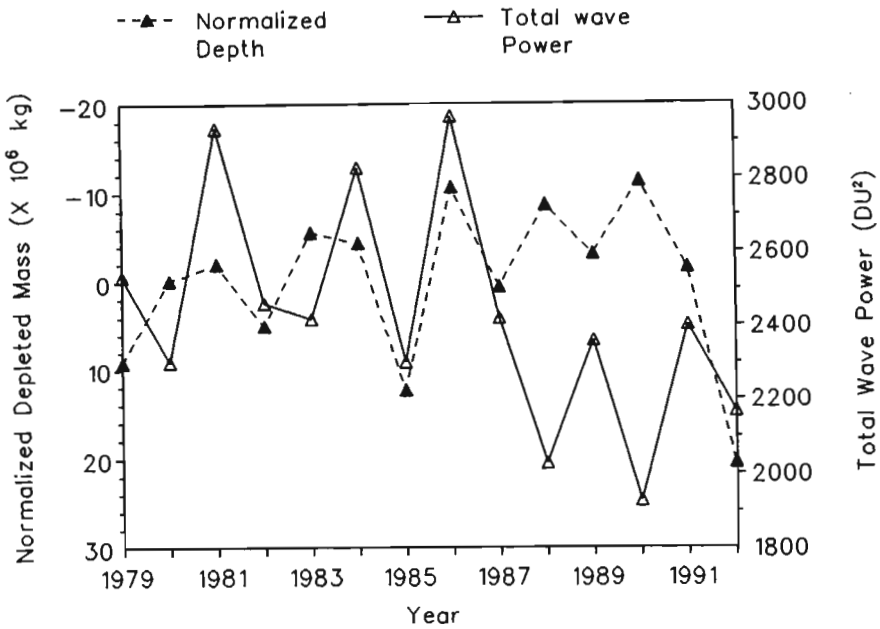


Figure 4.39: Total wave power averaged over the Northern Hemisphere winter together with the normalized depleted mass signal averaged over the same period. A linear normalization scheme has been used.

in the Northern Hemisphere. Large expanses of ocean and lack of topographic features poleward of 50°S and north of the Antarctic continent result in reduced Southern Hemisphere wave activity. Although wave powers for wavenumbers 2 to 4 were observed to be lower in the Southern Hemisphere (Figure 4.35), wave 1 power was lower in the Northern Hemisphere. This discrepancy may result from ozone wave amplitudes not being linearly related to geopotential height wave amplitudes.

Arctic ozone depletion does appear to be increasing with the rise in stratospheric chlorine loading (Figure 4.38). Significant negative trends in total column ozone in the high latitudes of the Northern Hemisphere have been reported in previous studies. Stolarski et al. (1991) reported trends of 0.4 to 0.5%/year between 40°N and 60°N but stated that the statistical uncertainties were larger than in the Southern Hemisphere due to the greater degree of meteorological variability. It was also suggested that this change could be due to effects of polar processed air being exported from the Arctic. This will be discussed further in the next section where a more detail account of global ozone depletion will be given.

Some of the reasons for the large differences in ozone depletion between the Arctic and Antarctic have been discussed in Chapter 1. Previous work (Waters et al., 1993) has shown that lower stratospheric temperatures in the Southern Hemi-

sphere fell below 190 K for approximately 4 months in 1992, but only occasionally so in the Northern Hemisphere 1991-92 winter. Both of these observations (reduced ozone depletion and higher stratospheric temperatures) can be linked to greater planetary wave activity resulting from greater orographic and thermal forcing by the more varied Northern Hemisphere land distribution (Schoeberl and Hartmann, 1991).

The weak dependence of the Arctic ozone depletion on planetary wave activity may result from the fact that the depleted mass signal contribution comes from only a few days during the early winter (see Figure 4.37) while planetary wave power maximizes later in the winter. The weaker Arctic ozone depletion results in total column ozone levels seldom falling below the 220 DU threshold. The mechanisms for the observed depletion include homogeneous gas phase reactions and heterogeneous reactions taking place on sulphuric acid aerosols (Chapter 5). As a result, it may be necessary to find a more appropriate depletion threshold for the Arctic that is relevant to its unique chemistry.

# Chapter 5

# THE INFLUENCE OF PLANETARY WAVES ON THE STATISTICAL PREDICTABILITY OF TOTAL COLUMN OZONE

## 5.1 Introduction

Planetary scale atmospheric disturbances, together with mesoscale weather systems are responsible for much of the day-to-day variability in total column ozone above a location (Reed, 1950). Mechanisms for planetary wave perturbations to total column ozone have been discussed in Chapters 1 and 2. In addition to causing local total ozone perturbations, planetary waves may also influence the ozone distribution non-locally. For example, it has been shown that the total ozone Quasi-biennial Oscillation (QBO) at middle and high latitudes results from the equatorial wind QBO modulation of planetary wave activity and the associated ozone transport (Shiotani, 1992). Furthermore, in the previous chapter it was shown that the severity of the annual Antarctic ozone depletion is significantly influenced by interannual variability in midlatitude planetary wave activity. It might be expected therefore, that during years of anomalously low (high) planetary wave activity, statistical models will over (under) predict total column ozone

at high southern latitudes. This may also be true for high northern latitudes although the midlatitude wave forcing does not significantly affect Arctic ozone depletion (Chapter 4).

The 14.5 year Nimbus-7 TOMS data set is an invaluable resource for the analysis of global ozone trends. It is therefore not surprising that many statistical models of total column ozone have made use of these data (Stolarski et al., 1991; Niu et al., 1992; Hood and McCormack, 1992; Gleason et al., 1993). Although statistical models have been used extensively in the analysis of global ozone trends, they are generally restricted to influences such as long-term linear trends, seasonal (annual) variations, the QBO, and the solar cycle variation. Since it is known that planetary wave activity influences the interannual variability in the global total ozone distribution, especially at high latitudes, the inclusion of a planetary wave activity parameter may improve the accuracy of such models. Clearly, until it becomes possible to predict interannual variability in midlatitude planetary wave activity (possibly never), the inclusion of a planetary wave activity index will not improve the predictive capabilities of statistical models. It will however improve long-term trend estimates derived from these models if the interannual planetary wave induced variability could be removed.

To investigate this hypothesis, differences between measured total column ozone and values derived from a statistical model are compared with the planetary wave activity series calculated in the previous chapter. A brief review of the salient features of previous statistical analyses of TOMS data is given in section 5.2 while a discussion of the results of these studies is given in section 5.5 where they are compared with the behaviour of the model discussed in this chapter.

## 5.2 Review

It appears that the primary aim of most statistical models of long-term ozone variation is to accurately determine linear trend coefficients. Influences of the seasonal cycle, QBO, El Niño Southern Oscillation (ENSO) and solar cycle are often treated as interference signals which must be removed before the linear trend is calculated. The relative amplitudes, phases and periods of these 'interference signals' are seldom compared. Most studies of global total column ozone variation make use of a multiple linear regression model of the form (Hood and McCormack,

1992):

$$O_3(t) = \mu(i) + \beta_T t + \beta_{QBO} X_{QBO}(t - L) + \beta_{SUN} X_{SUN}(t) + \varepsilon(t) \qquad (5.1)$$

where t is the time in months, $\mu(i)$ is a seasonal term equal to the long-term mean for the $i$th month of the year ($i$ = 1, 2, ..., 12); $X_{QBO}(t)$ is a time series representing the tropical QBO; $L$ is the optimum phase lag between $X_{QBO}(t)$ and $O_3(t)$; $X_{SUN}$ is a time series representing the solar variability where no phase lag is assumed; $\varepsilon(t)$ is a residual error term; and $\beta_T$, $\beta_{QBO}$ and $\beta_{SUN}$ are coefficients determined by least squares regression. The Singapore (1.4°N, 104.0°E) 30 hPa zonal winds are usually used to represent the QBO ($X_{QBO}(t)$) while the solar variability ($X_{SUN}(t)$) is tracked using the monthly mean 10.7 cm radio flux (F10.7). The models usually use monthly mean data (Stolarski et al. (1991) used weekly means), spatially averaging in the zonal direction, and over a few degrees of latitude. The sacrifice of longitudinal resolution in favour of seasonal variability is usually justified as the change in long-term trends with season dominates longitudinal variations (Stolarski et al., 1991; Hood and McCormack, 1992; Herman et al., 1993). Once models similar to equation 5.1 have been applied to zonal mean total column ozone and meridional profiles of $\beta_T$, $\beta_{QBO}$ and $\beta_{SUN}$ have been obtained, atmospheric dynamical and chemical mechanisms must be found to account for the meridional variability of the $\beta$ profiles.

Niu et al. (1992) used a more complex equation to derive trends as a function of latitude, longitude and season. Monthly average total column ozone (denoted $y_t$, t = 1, 2, ..., T), calculated in a grid of 10° latitude by 10° longitude, were modelled statistically using (Niu et al., 1992):

$$y_t = \sum_{i=1}^{12} \mu_i I_i(t) + \sum_{j=1}^{12} \beta_j I_j(t) R(t) + \gamma x_t + N_t \qquad (5.2)$$

$$N_t = \phi N_{t-1} + \varepsilon_t \qquad (5.3)$$

where t = 1, 2, ···, T and

$\mu_i$ total column ozone in month $i$, $i = 1, 2, \cdots, 12$.

$I_i(t)$ indicator series for the $i$th month of the year, 1 if month $t$ corresponds to month $i$ of the year, and 0 otherwise.

$\beta_j$ trend in month $j$ of the year, $j = 1, 2, \cdots, 12$.

$R(t)$ linear ramp function, $R(t) = t/12$.

$x_t$ solar 10.7 cm flux series, used as a proxy for the ultraviolet solar irradiance; the value of $\gamma$ is the associated coefficient.

$N(t)$ residual noise series, which is modelled as a first order autoregressive series since there is a month to month correlation in total column ozone values.

$\varepsilon_t$ an uncorrelated series with zero mean, but the variance of $\varepsilon_t$ depends on month; ie., $\sigma_t^2 = \sigma_{t-12}^2$.

This model was applied to TOMS data between 70°S and 70°N from November 1978 to May 1990. After observing significant regional differences in long-term linear trends, Niu et al. (1992) suggested that trends be calculated at a higher spatial resolution of 1.25° longitude by 1° latitude. This suggestion has been incorporated into the model discussed in this chapter.

Since most models deal with seasonal variability, difficulties are encountered at high latitudes during winter months where TOMS cannot make measurements (Chapter 3). To avoid this problem, the data are usually examined between 65°N and 65°S, ignoring the polar regions. This is a serious deficiency in such models since it is at high latitudes, where significant depletion of ozone occurs, that a measure of long-term linear trends is indispensable.

Hood and McCormack (1992) applied equation 5.1 to 13.2 years of TOMS data over this 65°N to 65°S latitude range to calculate latitudinal and seasonal variations in the regression coefficients of the linear trend, the QBO and the solar cycle. Stolarski et al. (1991) also examined TOMS data over this latitude range, from November 1978 to May 1990, to derive long-term trends as a function of latitude and season. The model used was similar to equation 5.1 although second order autoregressive noise was assumed. The global time series showed that total column ozone as measured by TOMS had declined $0.26 \pm 0.14\%$/year over this period. Using high latitude data outside of the polar night, it was also shown that negative trends increase with latitude, as discussed further in section 5.5. Herman et al. (1993) calculated long-term trends in TOMS data as a function of latitude and month over the period January 1979 to December 1991. Daily ozone

data were zonally averaged within 2° latitude bands from 90°N to 90°S.

As mentioned above, to reduce computation time, many models average over a few degrees of latitude. It will be shown below that some features of the meridional variation in model coefficients would be significantly smoothed under such conditions (see for example the meridional variation in the QBO amplitude – Figure 5.19).

## 5.3 Description of the statistical model

### 5.3.1 Model features

The model developed in this chapter has not followed the standard models described in equations 5.1 and 5.2 above. The features of this model are:

- Seasonal variability in model coefficients was neglected completely in favour of high spatial resolution. Whereas, for example, Niu et al. (1992) used a spatial resolution of 36 × 18, the highest possible resolution of 288 × 180 (single TOMS cells) has been used in this model.
- Monthly average total column ozone data were used (as for previous models) from 1979 to 1992 to provide 168 month data series.
- No proxy data were used to track QBO and solar cycle variability. These were identified in the data signals by virtue of their periods only. This is a more flexible approach as it makes no *a priori* assumptions of the forcing mechanisms.
- Interpolation over the polar caps was used to infer data poleward of the polar night terminator in daily TOMS ozone distributions. This procedure, used in previous analysis (Schoeberl et al., 1989), permits derivation of global model coefficients.

The 14 year total column ozone record for each of the 51840 TOMS cells covering the globe has been modelled statistically by fitting the following equation to the 168 (14 years × 12 months) monthly means:

$$
\begin{aligned}
\Omega \;=\; & K_1 \times M + && \text{Term 1}\\
& K_2 + && \text{Term 2}\\
& K_3 \times \sin((M+K_4) \times 2\pi/K_5) + && \text{Term 3}\\
& K_6 \times \sin((M+K_7) \times 2\pi/K_8) + && \text{Term 4}\\
& K_9 \times \sin((M+K_{10}) \times 2\pi/K_{11}) + && \text{Term 5}\\
& K_{12} \times \sin((M+K_{13}) \times 2\pi/K_{14}) && \text{Term 6}
\end{aligned}
$$

where $\Omega$ is the model monthly average total column ozone, M is the month number (Jan 1979 = 1, Dec 1992 = 168) and the K values are the 14 model coefficients. Terms 1 and 2 are the linear trend and offset respectively where $K_1$ and $K_2$ are determined using regression analysis. Terms 3 to 6 characterize the annual variation, the QBO, the solar cycle variation and a semi-annual (6 month period) variation respectively. In each of these terms the $K_{3,6,9,12}$ values are the amplitudes, the $K_{4,7,10,13}$ values are the phases and the $K_{5,8,11,14}$ values are the periods.

An advantage of this form of model is that no additional input data series are required eg. 50 hPa Singapore wind fields to remove the QBO dependence or F10.7 to remove the solar cycle dependence. In this model these signals are labelled only by the agreement of their periods with expected periods. A disadvantage of this version of the model is that quantitative error estimates may be obtained only for the linear trend determination. Some measure of error for the sinusoidal variations may be obtained by comparing the observed periods with theoretical periods as detailed below.

### 5.3.2 Examples of the application of the model

Five locations near the Greenwich meridian at 59.5°N, 29.5°N, 0.5°S, 29.5°S and 59.5°S are used as examples to show the way in which this function has been fitted to the monthly mean data series for each TOMS cell. Figures 5.1 a to 5.1 e show the monthly average total column ozone for each of the 5 locations for the 168 months from January 1979 to December 1992. The long-term linear trends are also shown on each plot using dashed lines. It is clear that the variability in monthly mean signals increases with latitude and Northern Hemisphere locations experience greater annual variability than conjugate locations in the Southern Hemisphere (cf. Figures 5.1 a and 5.1 e). The $K_1$ (linear trend) and $K_2$ (offset) model coefficients obtained from the straight line fits are listed

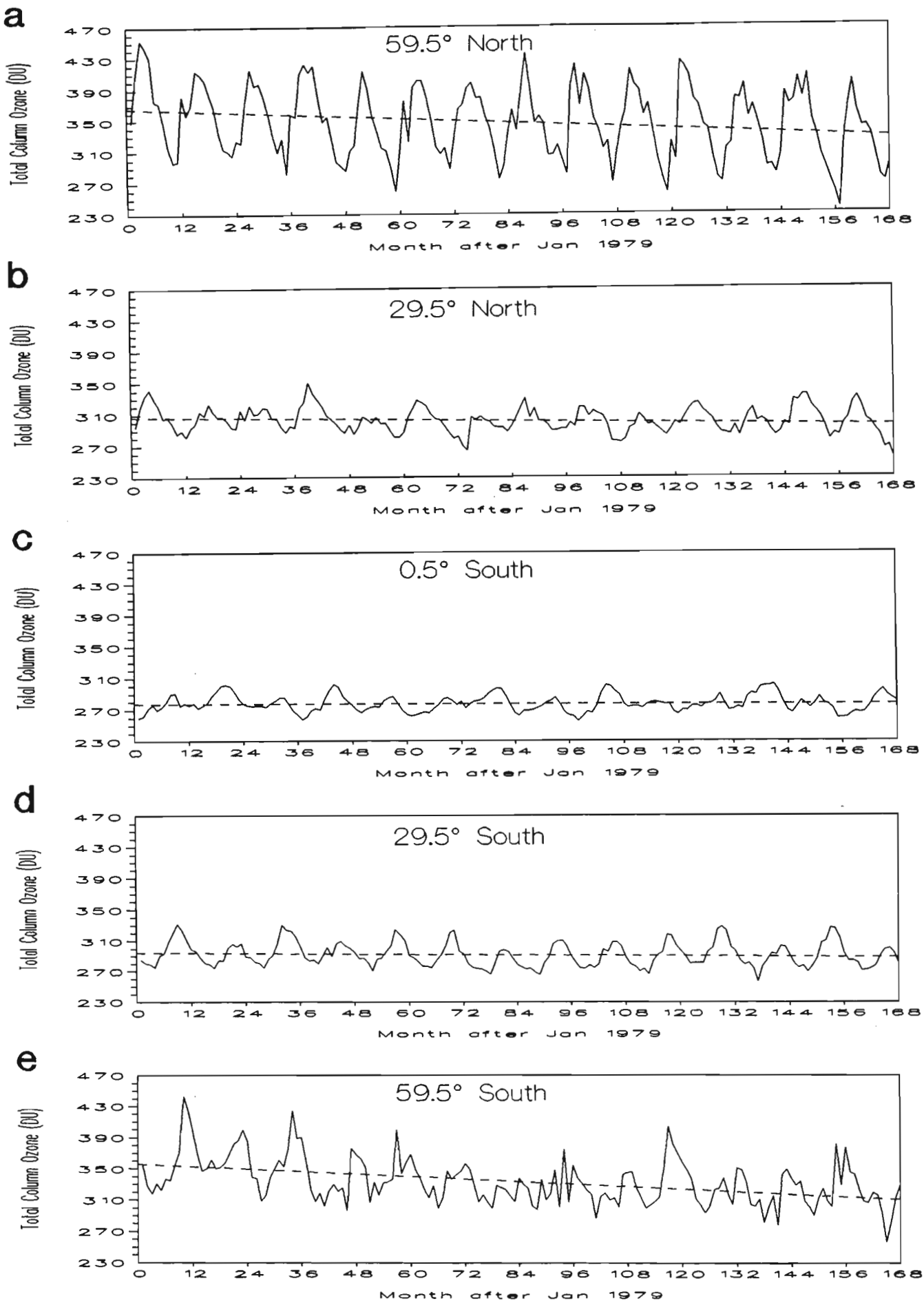


Figure 5.1: Monthly average total column ozone above each of the 5 example locations, from 1979 to 1992 (solid line). The long-term linear trend is also shown in each panel (dashed line).

in Table 5.1. The linear trends increase with latitude indicating that long-term

| Location | | $K_1$ | $K_2$ |
|---|---|---|---|
| Latitude | Longitude | (DU/month) | (DU) |
| 59.5°N | 0.625°E | -0.236 | 366.0 |
| 29.5°N | 0.625°E | -0.068 | 307.2 |
| 0.5°S | 0.625°E | -0.000 | 277.5 |
| 29.5°S | 0.625°E | -0.039 | 294.9 |
| 59.5°S | 0.625°E | -0.267 | 356.1 |

Table 5.1: The $K_1$ and $K_2$ model coefficients obtained from the regression analysis of the monthly averages for the 5 locations

ozone depletion occurs preferentially in the polar regions. Although trends are more negative at high latitudes, the offset values also increase with latitude and Northern Hemisphere locations have greater offset values than conjugate locations in the Southern Hemisphere (see $K_2$ values in Table 5.1).

The linear trends are removed from each data series leaving zero mean signals with no linear trend. A least squares technique, similar to that used in section 4.4 (see also Appendix A), is used to fit the annual variation sinusoidal term (Term 3) to these residual signals by locating the minimum of the sum of the squares of the differences between the sine curve and the data signal in the ($K_3$, $K_4$, $K_5$) space. No constraints are placed on allowed values of $K_3$, $K_4$ and $K_5$, for example the period is not fixed at 12 months for the annual variation. The residual data curves and the best fit annual variation signals are shown in Figures 5.2 a to 5.2 e. The fitted sine curves (dashed lines) track the data signals (solid lines) closely, showing differences in amplitude and phase between sites. The $K_3$ (annual variation amplitude), $K_4$ (annual variation phase) and $K_5$ (annual variation period) model coefficients obtained from the fitted sinusoidal terms are listed in Table 5.2. The amplitude of the annual variation increases with latitude with the high latitude Northern Hemisphere location showing greater seasonal variation compared to the conjugate Southern Hemisphere location. The annual variation amplitude of 54.6 DU at 59.5°N is more than twice the value of 23.5 DU at 59.5°S. Note that the phases ($K_4$ values) are not constant and Northern Hemisphere signals are not simply six months out of phase with Southern Hemisphere signals. The periods ($K_5$ values) are very close to the expected value of 12.00 months indicating the validity of labelling these signals as the annual variations.

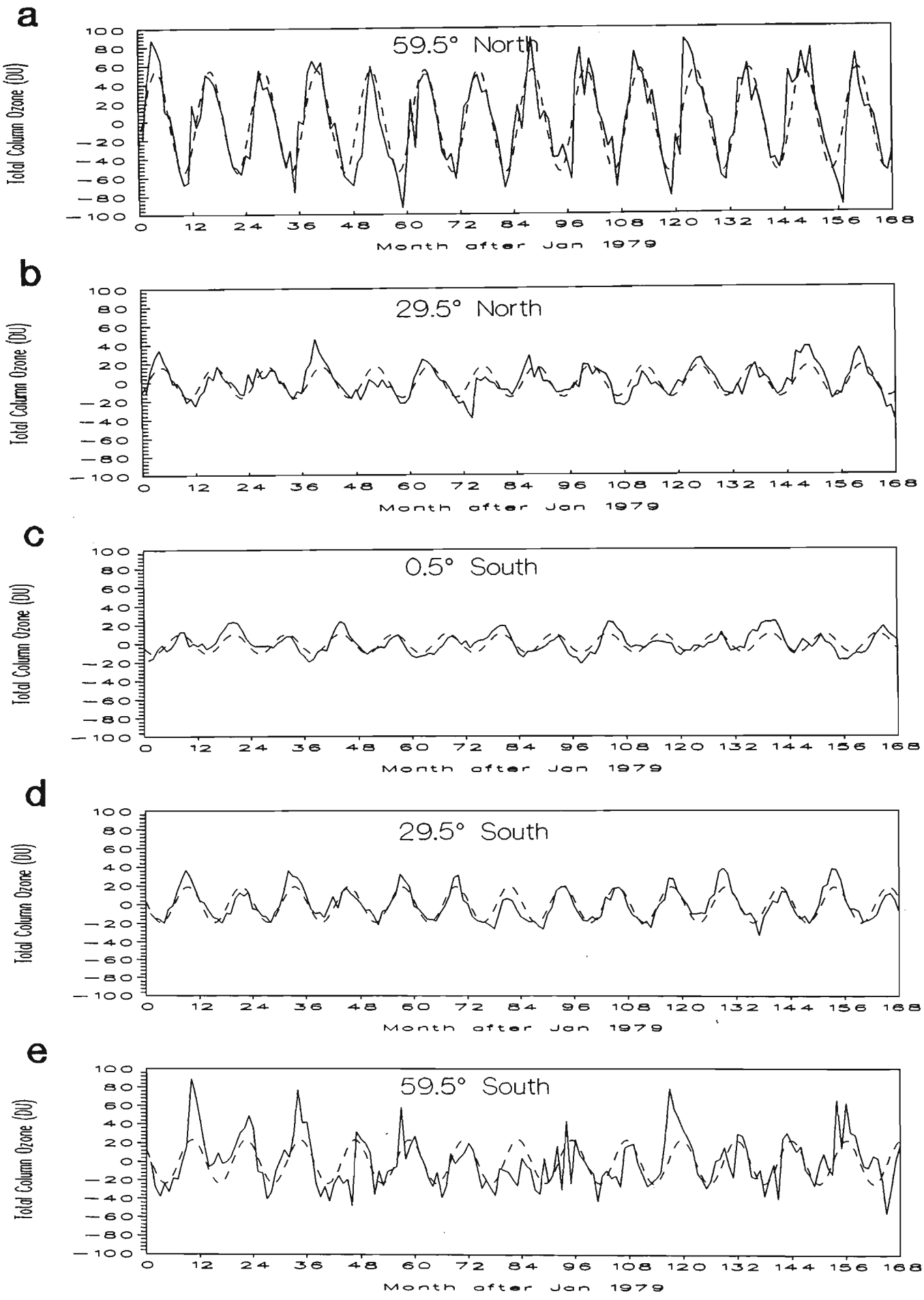


Figure 5.2: The detrended monthly mean total column ozone above each of the example locations (solid line). Also shown for each location is the best fit sine function (dashed curve) which represents the annual variation.

| Location | | $K_3$ | $K_4$ | $K_5$ |
|---|---|---|---|---|
| Latitude | Longitude | (DU) | (months) | (months) |
| 59.5°N | 0.625°E | 54.6 | 11.02 | 12.04 |
| 29.5°N | 0.625°E | 16.5 | 10.57 | 12.02 |
| 0.5°S | 0.625°E | 10.3 | 6.86 | 11.98 |
| 29.5°S | 0.625°E | 19.1 | 5.66 | 12.02 |
| 59.5°S | 0.625°E | 23.5 | 4.99 | 12.16 |

Table 5.2: The $K_3$, $K_4$ and $K_5$ model coefficients obtained from the sinusoidal fits to the detrended data

The fitted sine curves are then subtracted from the data signals to produce a second set of residual curves. Sinusoidal functions are again fitted to these residual curves as for the annual variation. Figures 5.3 a to 5.3 e show the second set of residual data curves (detrended and deseasonalised) and their best fit sine functions. The fitted sine curves track the data signals well at low latitudes (Figures 5.3 b to 5.3 d). At high latitudes however (Figure 5.3 a and 5.3 e), the signals show greater variability with the result that the tracking of the QBO signal may not be as accurate. The $K_6$ (QBO amplitude), $K_7$ (QBO phase) and $K_8$ (QBO period) model coefficients obtained from the fitted sine curves are listed in Table 5.3.

| Location | | $K_6$ | $K_7$ | $K_8$ |
|---|---|---|---|---|
| Latitude | Longitude | (DU) | (months) | (months) |
| 59.5°N | 0.625°E | 7.2 | 7.50 | 35.56 |
| 29.5°N | 0.625°E | 6.1 | 4.55 | 29.91 |
| 0.5°S | 0.625°E | 7.7 | 19.57 | 30.17 |
| 29.5°S | 0.625°E | 6.6 | 2.07 | 29.80 |
| 59.5°S | 0.625°E | 8.2 | 36.56 | 45.39 |

Table 5.3: The $K_6$, $K_7$ and $K_8$ model coefficients obtained from the sinusoidal fits to the detrended and deseasonalised data

Previous studies have quoted QBO periods of 26 to 30 months (Zerefos et al., 1992) and 22 to 34 months (Gray and Ruth, 1993). Other than the two high latitude locations, the periods of the signals fall within this range. Inaccuracies at the high latitude sites may result from poor model tracking of the highly variable signals. It is also possible that discrepancies at the poleward sites may

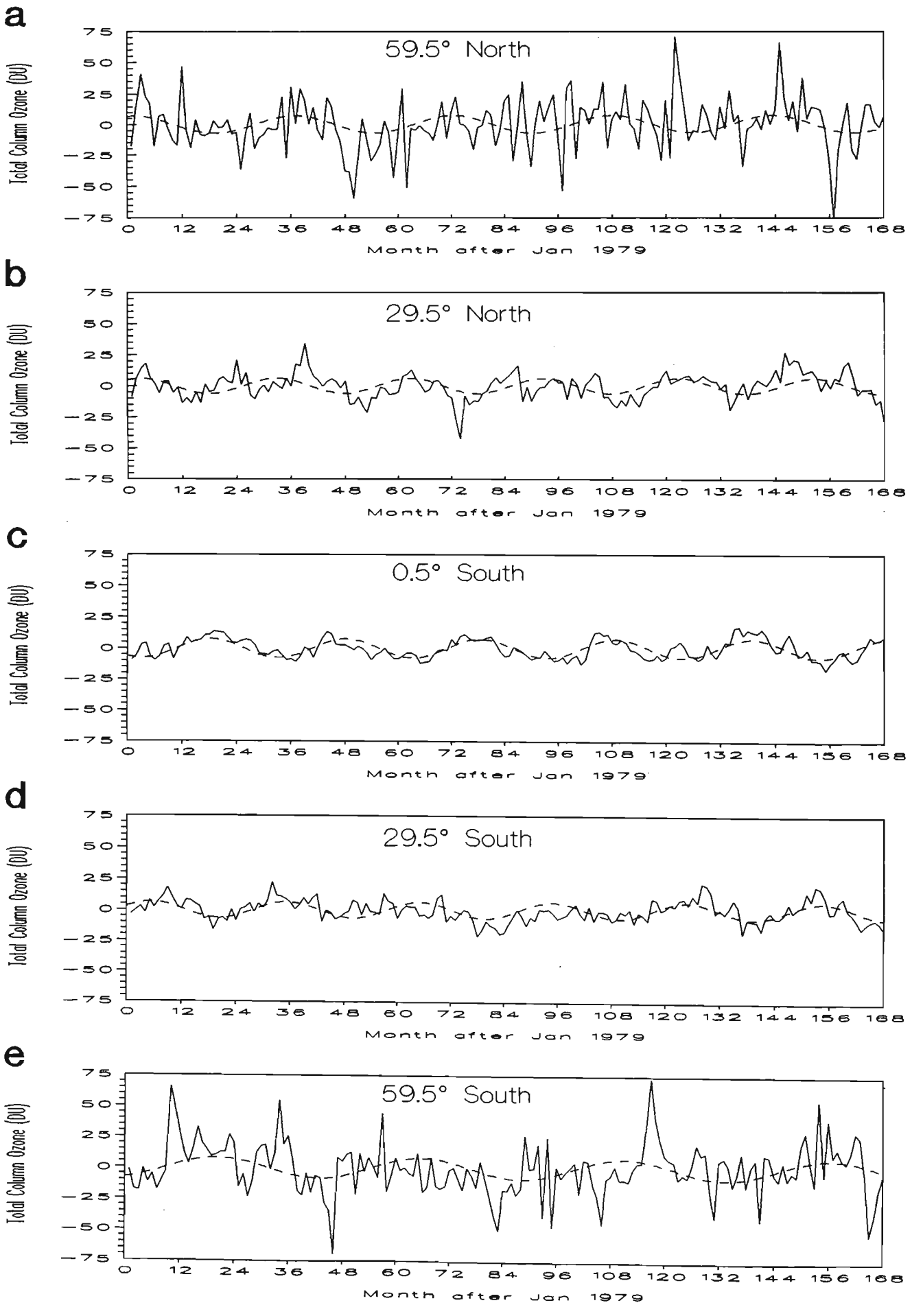


Figure 5.3: The detrended and deseasonalised monthly mean total column ozone above each location (solid line). The best fit sine curves, representing the QBO, are shown for each location using dashed lines.

be related to ENSO influences (Henderson-Sellers and Robinson, 1987) where the model tends to track the ENSO signal rather than the QBO signal. Once the coefficients from the two poleward sites are rejected due to their erroneous periods, no latitudinal trends may be inferred from the remaining three sets of coefficients.

As before, the QBO signals are subtracted from each data series to leave residual data curves to which sinusoidal functions are fitted. These fitted curves, thought to characterize the influence of the solar cycle, are shown in Figures 5.4 a to 5.4 e (dashed lines) together with the residual data signals (solid lines). The coarse scale on the vertical axis obscures the fitted sine curves at low latitudes (Figures 5.4 b to 5.4 d). The tracking at high latitudes (Figure 5.4 a and 5.4 e) may be erroneous due to the large variability in the signals which may also exaggerate the solar cycle amplitudes. The $K_9$ (solar cycle amplitude), $K_{10}$ (solar cycle phase) and $K_{11}$ (solar cycle period) model coefficients obtained from the sinusoidal terms are listed in Table 5.4. Although it is not statistically sound to expect to

| Location | | $K_9$ | $K_{10}$ | $K_{11}$ |
|---|---|---|---|---|
| Latitude | Longitude | (DU) | (months) | (months) |
| 59.5°N | 0.625°E | 7.3 | 54.59 | 141.52 |
| 29.5°N | 0.625°E | 3.1 | 38.11 | 154.69 |
| 0.5°S | 0.625°E | 1.6 | 56.25 | 187.50 |
| 29.5°S | 0.625°E | 2.7 | 26.48 | 150.00 |
| 59.5°S | 0.625°E | 6.6 | 37.50 | 150.00 |

Table 5.4: The $K_9$, $K_{10}$ and $K_{11}$ model coefficients obtained from the sinusoidal fits to the residual data curves

accurately track a theoretical 11 year solar cycle signal in a 14 year data series, the periods obtained ($K_{11}$ values) are generally close to the expected 132 month period. Amplitudes ($K_9$ values) increase with latitude, possibly as a result of the increased variability at the poleward sites, as mentioned earlier. Solar cycle phases ($K_{10}$ values) appear to show no latitudinal dependence.

The solar cycle signals are removed from each data series, creating another set of residual curves to which sinusoidal functions are fitted. These sine functions, thought to characterize semi-annual oscillation observed in previous studies (Perliski et al., 1989), are shown in Figures 5.5 a to 5.5 e (dashed lines) together with the residual data signals (solid lines). Again the coarse scale on the vertical

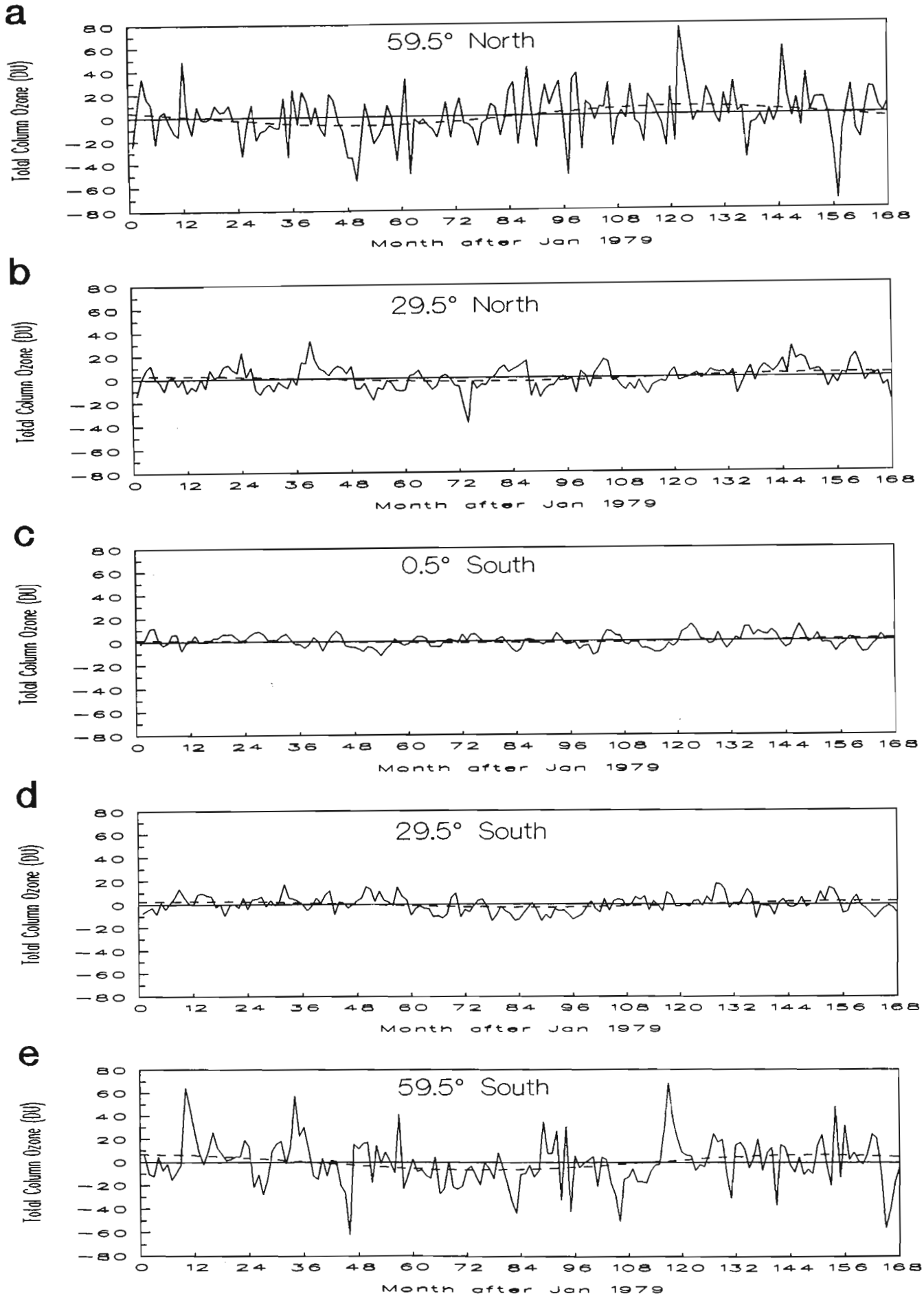


Figure 5.4: The residual data curves (solid lines) showing the influence of the solar cycle. The large variability at high latitudes obscures the low amplitude fitted sine functions (dashed curves).

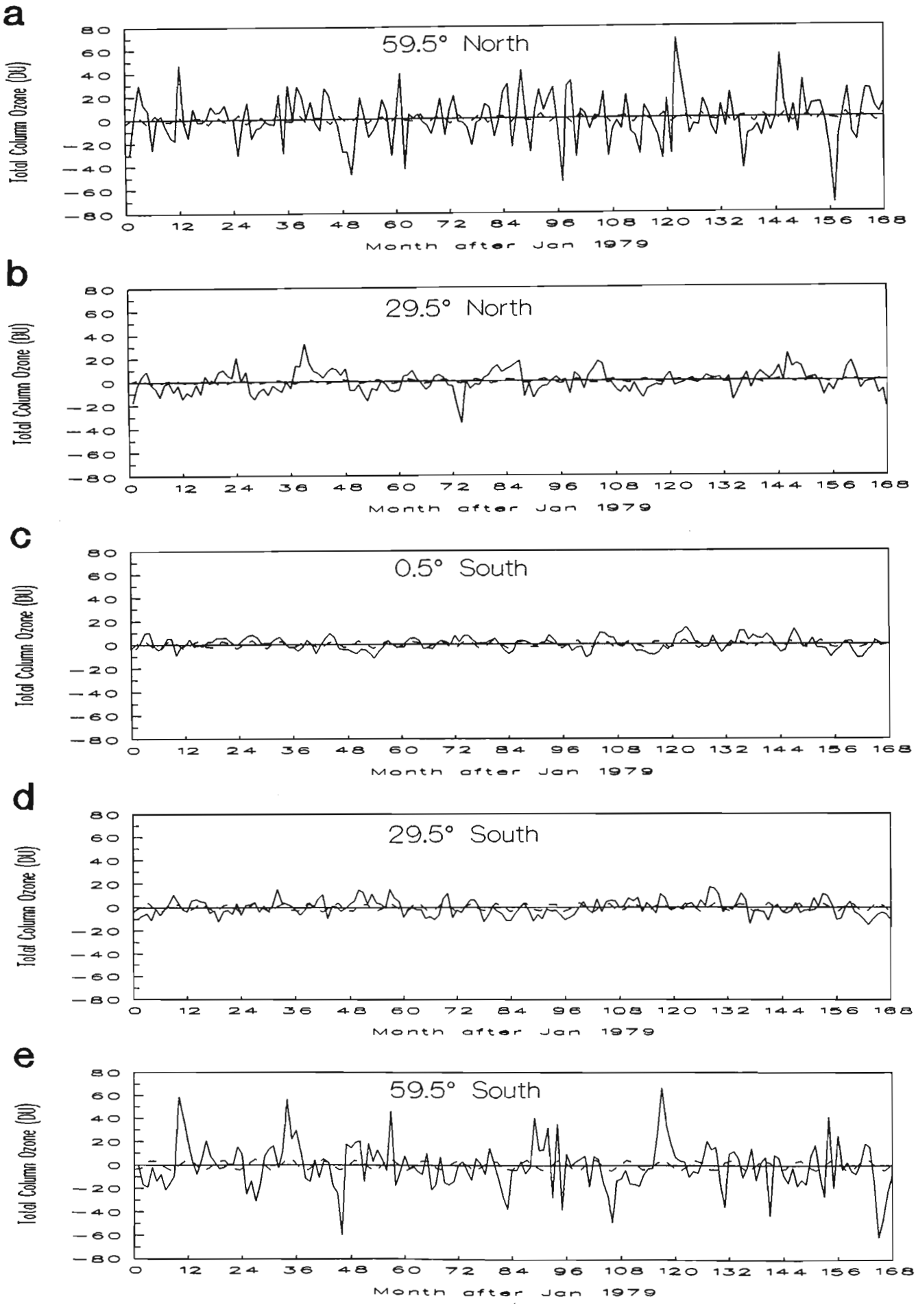


Figure 5.5: The residual data curves (solid lines) showing the influence of the semi-annual cycle. The large variability at high latitudes obscures the low amplitude fitted sine functions (dashed curves).

axes obscures the fitted curves. The tracking appears to be valid at 0.5°S and 29.5°S while at the other 3 locations no clear semi-annual signal is present. The $K_{12}$ (semi-annual variation amplitude), $K_{13}$ (semi-annual variation phase) and $K_{14}$ (semi-annual variation period) model coefficients obtained from the fitted sinusoidal terms are listed in Table 5.5. In 3 out of the 5 data sets, the model

| Location | | $K_{12}$ | $K_{13}$ | $K_{14}$ |
|---|---|---|---|---|
| Latitude | Longitude | (DU) | (months) | (months) |
| 59.5°N | 0.625°E | 4.1 | 3.05 | 4.48 |
| 29.5°N | 0.625°E | 1.8 | 7.01 | 7.49 |
| 0.5°S | 0.625°E | 3.4 | 4.75 | 6.00 |
| 29.5°S | 0.625°E | 3.9 | 4.36 | 6.00 |
| 59.5°S | 0.625°E | 4.6 | 5.06 | 7.52 |

Table 5.5: The $K_{12}$, $K_{13}$ and $K_{14}$ model coefficients obtained from the sinusoidal fits to the residual data curves

does not accurately track a 6 month period signal. This can either result from there being no 6 month signal present in the data, or the high variability at these sites results in poor model tracking. When the terms with periods significantly different from 6 months are discounted, no latitudinal trends may be inferred. At low latitudes this semi-annual signal is thought to arise from the meridional movement of the intertropical convergence zone (ITCZ) as discussed in section 5.5.18.

The removal of these 5 sine curves results in a set of final residual data signals, shown in Figures 5.6 a to 5.6 e. Fourier analyses of these data series indicates no remaining periodicity. The final residual data curves show the difference between measured monthly averages and values generated by the 14 model coefficients. These differences are analyzed further in section 5.6 where they are compared with the planetary wave signatures obtained in the previous chapter.

It was thought that it may be possible to see the effects of episodic volcanic eruptions in these residual data sets. The TOMS instrument data retrieval may be adversely affected after the injection of large volumes of aerosol into the stratosphere (Chapter 3). Furthermore, increases in stratospheric $H_2SO_4$ concentrations from volcanic eruptions may facilitate ozone depletion at low latitudes (section 5.5.1). It has been shown (Herman et al., 1993) that errors caused by the short-term presence of stratospheric aerosols in the TOMS zonally averaged ozone data are

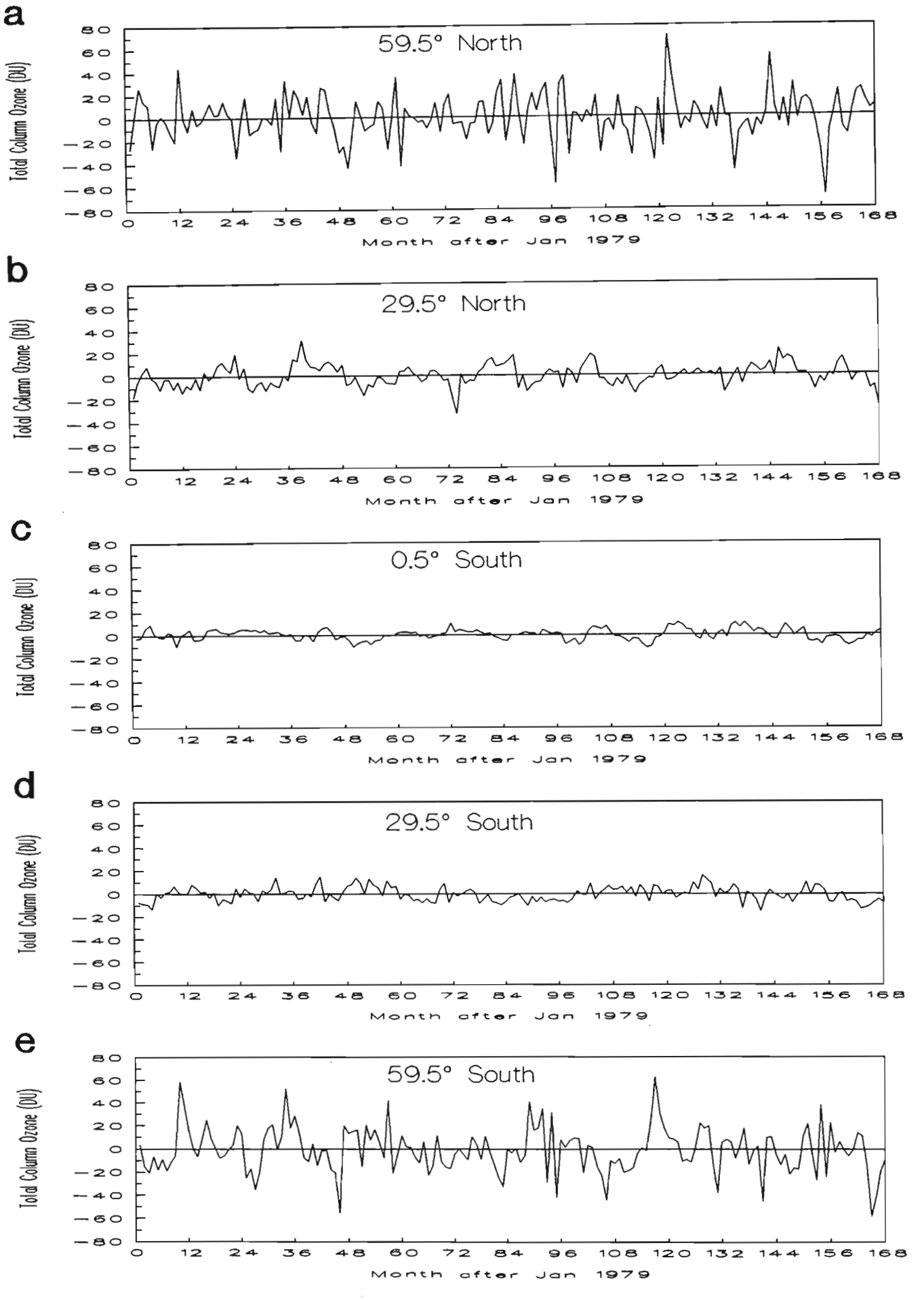


Figure 5.6: The final residual data curves for the 5 example locations.

less than 1% before correction, and have no significant effect on ozone trend determination. However, it may be possible that volcanic perturbations can be observed in the residual signals of the high spatial resolution data used in this model.

### 5.3.3 A note on the computational technique

In the development of this model, it was decided to use the least squares sinusoidal fitting method in preference to usual Fourier transform techniques since more precise amplitude, phase and period values may be calculated in this way. The results presented in this section are from version 2.00 of the model. This version improved on version 1.00 in the follows ways:

- It extended the data series by one additional year (1992).
- It included tracking of the semi-annual variation in total column ozone.
- It allowed the annual variation period to be determined by the model. This was fixed at 12 months in version 1.00.
- It made use of interpolation over the polar caps to allow calculation of global model coefficients.

Version 1.00 made use of Fast Fourier Transforms (FFTs) to determine the period of each signal and provide an initial estimate of the amplitude. The robust sine function fitting technique was then used to obtain the phase and a more accurate measure of the amplitude. The 168 data points in each time series required extensive zero padding to permit use of the FFT algorithm. Together with the required cosine window smoothing (Appendix A), this introduced measurable inaccuracies. The standard Fourier transform could not be used as it is too slow to permit processing of 51840 data signals. The data processing for version 1.00 of the model was computationally far less demanding than for version 2.00.

For the current version of the model, processing on a fast 486 dos-based PC takes approximately 10 minutes to obtain the 14 coefficients for a single TOMS cell. For the 51840 cells covering the globe, this would require almost a year of full-time processing. To reduce this prohibitive time period, a local area network (LAN) of computers was used to simulate a large parallel processing device. With access to

15 LAN machines during off-peak hours it was possible to complete the processing for the entire globe in 12 days. This is more than a factor of 15 faster than a single 486 due to advantages in disk access times on many different machines.

Three separate pieces of software were written to perform this parallel processing task: the control program, the message interface program and the calculation program. A brief description of each of these is given below.

The ozone model control program

This program runs on a single LAN machine acting as a task scheduler. It determines which data sets still require processing and where bottlenecks are occurring in the system. It then dynamically assigns tasks to different server computers to reduce delays and minimize processing time. It copies the required archived data sets from a remote harddrive to the LAN drive when requested by the message interface program.

The ozone model message interface program

This program runs on each of the server computers where the actual calculations are performed. It acts as a messenger between the control program and the calculation program. It sends messages to the control program to register the serving computer, to request raw data files and to return processed data. It changes the requests for data depending on the processing power of the local computer.

The ozone model calculation program

This is the program that does the actual calculations. It is called by the message interface program, obtaining its input parameters from a file created by the interface program. This program can also be run independently of the parallel processing system if only a few locations need to be processed.

Version 3.00 of the model will incorporate the following improvements:

- Tracking of the ENSO signal through the use of an ENSO index such as the Pacific Ocean sea surface temperature (SST).
- The use of data adaptive filtering to reduce noise levels at high latitudes.
- Calculation of the quality of the sinusoidal fits eg. determination of correlation coefficients between measured data and fitted sine curves to provide quantitative error estimates.

- Analysis of the autoregressive nature of all time series, especially the final residual data series.
- Investigation of local volcanic perturbations to the trends.
- Addition of 1993 Meteor-3 TOMS data.

## 5.4 Model results

The model described above has been used to calculate the 14 $K$ coefficients for each TOMS cell covering the globe. Analysis of the coefficients provides important insight into the mechanisms responsible for the global distribution of total column ozone. In addition they may be used in a predictive capacity to provide estimates of future monthly average total ozone.

Mercator projections of model coefficients are displayed below using colour codes to indicate the values in each TOMS cell. In each case the lowest values are shown in hues of blue with values increasing through hues of green, red and purple to the highest values shown in yellow. The colour codes for each coefficient are shown on colour bars on the right of each plot. Continental outlines are shown using solid lines and the 60°N, 30°N, 0°S, 30°S and 60°S parallels are shown with dashed lines. Each plot is discussed separately below.

As a result of the zonal symmetry exhibited in many of these plots, profiles of zonal means may be used to highlight differences between hemispheres. In some of these plots included in the following sections, mirror images of the profile about the equator are shown using dashed lines to further highlight hemispheric differences.

### 5.4.1 The long-term linear trend ($K_1$ coefficient)

Figure 5.7 shows the long-term linear trend (in DU/month) obtained from regression analysis of 168 monthly averages of total column ozone in each TOMS cell. The distribution is predominantly zonally symmetric although a region of reduced negative trends appears off the east coast of Canada while a region of enhanced decline appears over Siberia. The blue area over the Antarctic, showing large negative trends, is a consequence of the sustained increase in chlorine catalyzed

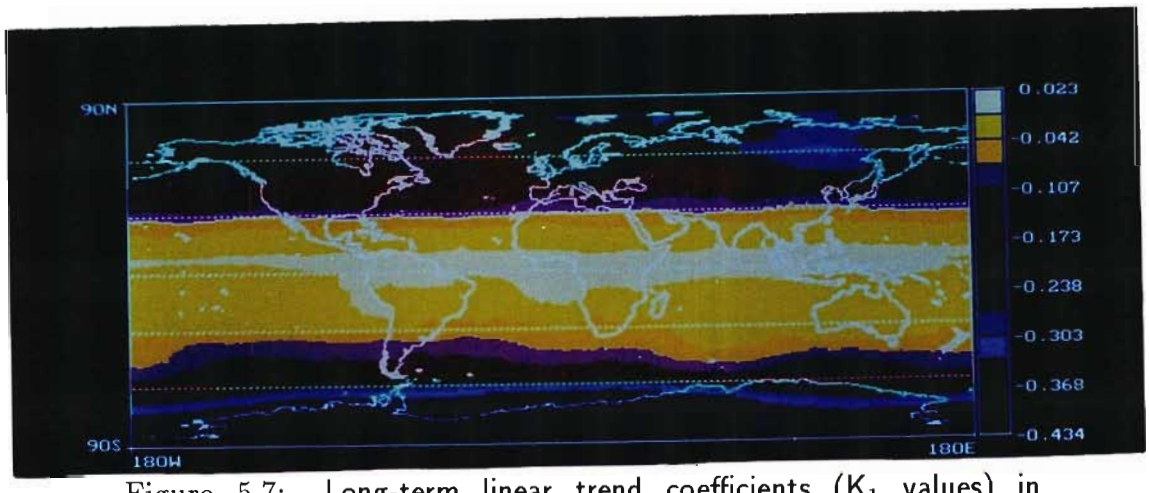


Figure 5.7: Long-term linear trend coefficients ($K_1$ values) in DU/month.

ozone depletion over this area in the spring of each year, as discussed extensively in Chapters 1 and 4. Linear trends in this region exceed -0.34 DU/month = -4.0 DU/year and can be as high as -5.2 DU/year. Although complete error analysis has not been included in this version of the model, investigation of regression $r^2$ values indicates that linear trend coefficients shown in yellow in Figure 5.7 may not be statistically significant.

To better illustrate the hemispheric differences in the linear trends, a meridional profile of the zonal means is shown in Figure 5.8 together with its mirror image. The large negative values over the Antarctic are almost twice as large as those over the Arctic. Note also that Southern Hemisphere trends are elevated above Northern Hemisphere trends between 20°S and 60°S, poleward of which they become more negative than Northern Hemisphere values. Comparisons of zonal means at midlatitudes (~40°) show Northern Hemisphere trends of -0.151 DU/month and smaller Southern Hemisphere trends of -0.067 DU/month. Similar meridional variations in linear trends have been observed in previous analyses (Stolarski et al., 1991).

### 5.4.2 The linear trend offset ($K_2$ coefficient)

Figure 5.9 shows the offset values obtained from the regression analysis. Ignoring the contributions from the sinusoidal terms, these values give an indication of the monthly average total column ozone during the last month of 1978. This plot exhibits considerable zonal asymmetry with local zonal minima over Scandinavia,

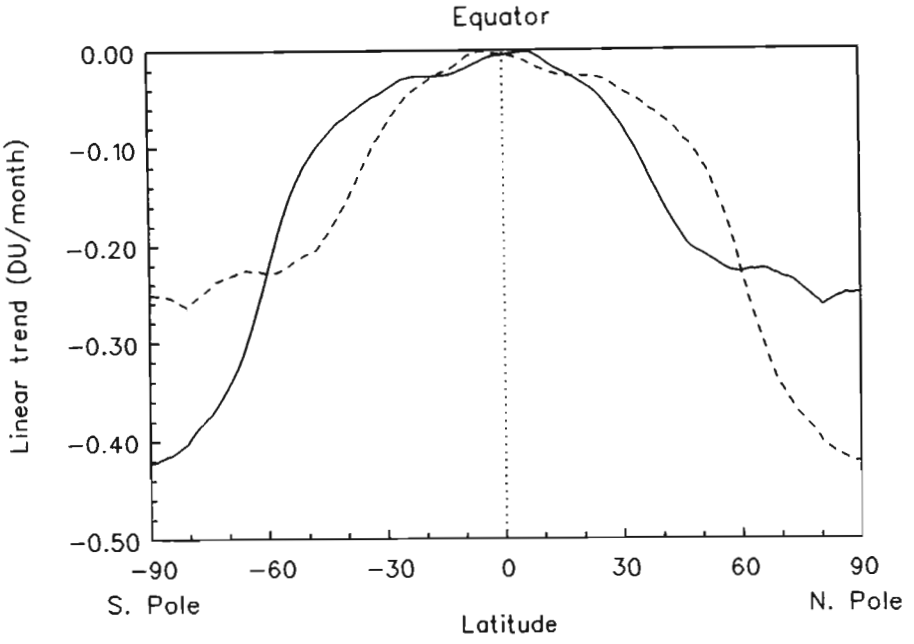


Figure 5.8: Meridional profile of the zonal means of the long-term linear trend model coefficients ($K_1$ values).

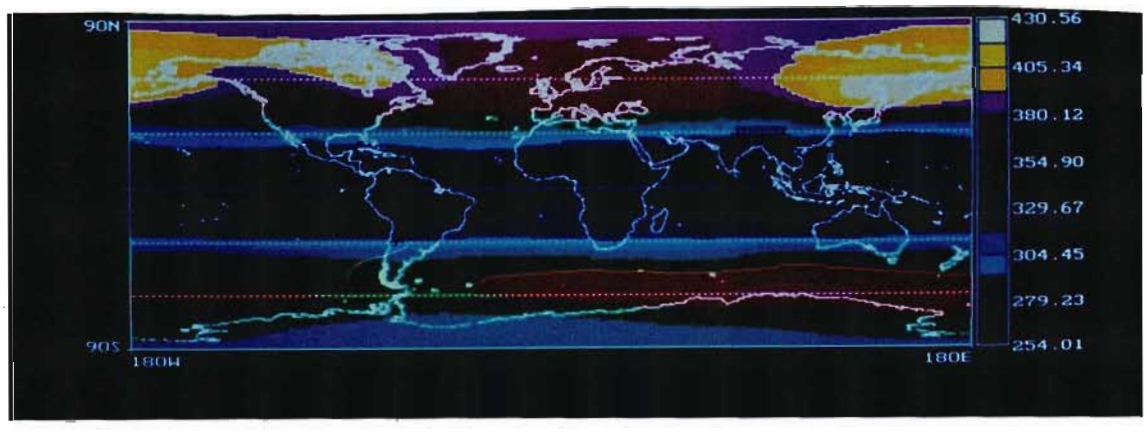


Figure 5.9: Linear trend offset coefficients ($K_2$ values) in DU.

the Pacific Ocean (minimally), and the Antarctic Peninsula. The zonal asymmetry along the 60°S parallel may result from the large quasi-stationary planetary wave 1 that forms during the winter of each year at these latitudes (Chapter 4). The wave structure usually has its maximum south of Australia and its minimum along the Antarctic peninsula and South America.

The meridional morphology of the distribution differs considerably between hemispheres. In the Northern Hemisphere, the offset value is low over the equator and increases through the midlatitudes to almost constant values over the Arctic. In

the Southern Hemisphere however, there is a local maximum near 60°S, poleward of which values decrease again. The feature is made more apparent in Figure 5.10 where zonal means of the offset are plotted. In the Northern Hemisphere

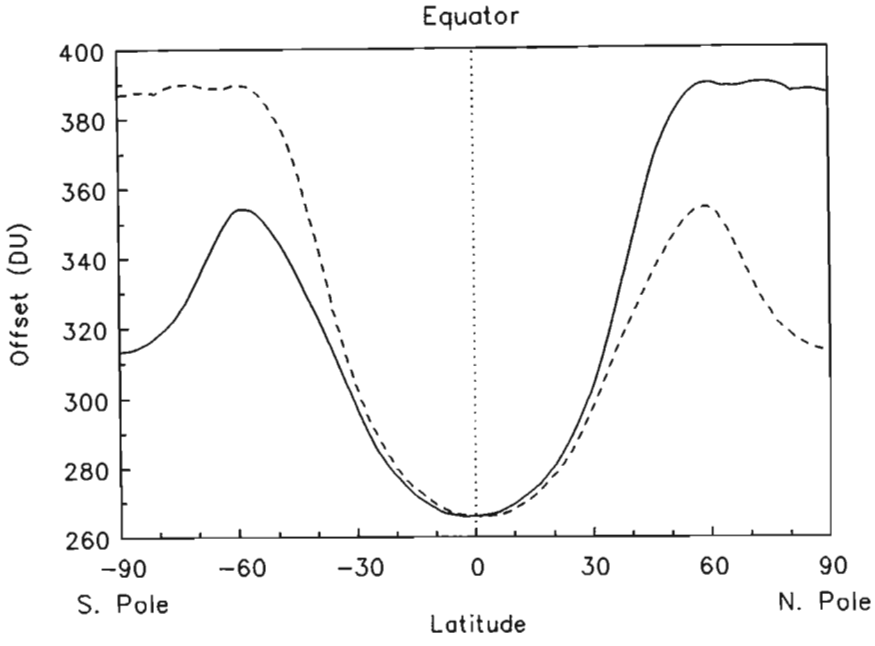


Figure 5.10: Meridional profile of the zonal means of the offset model coefficients ($K_2$ values).

values rise smoothly from the equator to 60°N, after which they remain constant or decrease slightly towards the pole. In the Southern Hemisphere, values rise from the equator to 60°S as for the Northern Hemisphere, though the meridional rate of increase is lower. This creates an offset of 353 DU at 60°S compared with a 60°N offset of 390 DU. Poleward of 60°S, values decrease sharply in comparison with the Northern Hemisphere where they remain almost constant. Comparisons of zonal means at the poles indicates a value of 313 DU at 90°S and 387 DU at 90°N.

### 5.4.3 The amplitude of the annual variation ($K_3$ coefficient)

Figure 5.11 shows the amplitudes of the annual variation signals in DU. These values must be doubled to obtain the typical annual range in monthly average total column ozone above each location. The most prominent feature of this plot is the steady meridional increase in annual amplitudes in the Northern Hemisphere while values in the Southern Hemisphere exhibit a local maximum at midlatitudes

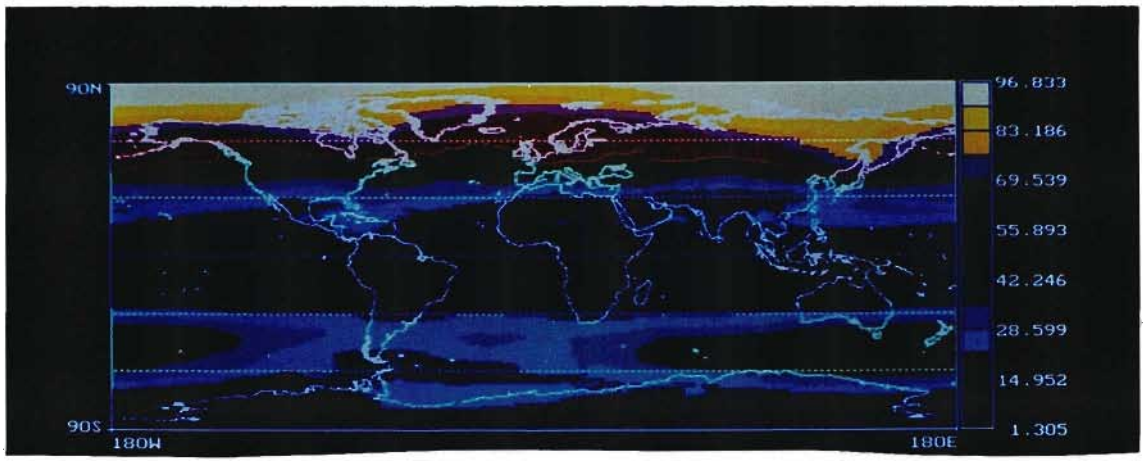


Figure 5.11: Annual variation amplitude coefficients ($K_3$ values) in DU.

before decreasing to equatorial values over the Antarctic. This may result from there being less ozone over the Antarctic than over the Arctic as indicated by the lower offset values ($K_2$ coefficients, see Figure 5.9) and steeper linear trends ($K_1$ coefficients, see Figure 5.7). However, even when calculated as a percentage of the offset value, the amplitude of the annual variation in the Northern Hemisphere far exceeds that of the Southern Hemisphere. An equatorward intrusion of higher values over the Far East perturbs the zonal symmetry of the Northern Hemisphere distribution, while reduced values around the South American continent perturbs the zonal symmetry in the Southern Hemisphere. The local maximum in annual amplitudes, centered near 40°S, is shown more clearly in Figure 5.12 where the zonal means are plotted. The mirror image curve (dashed line) shows that Northern Hemisphere annual amplitudes exceed those for the Southern Hemisphere at all latitudes. Differences become particularly apparent poleward of 40°. Note also that the absolute minimum in this curve does not occur at the equator but at about 8°S.

### 5.4.4 The phase of the annual variation ($K_4$ coefficient)

The data plotted in Figure 5.13 specify the month of the year when each location can be expected to experience its annual minimum in monthly average total column ozone. Interpretation of the phase of a sine function indicates that the phase of the annual variation ($K_4$) when subtracted from 9 gives the time into the year (in months) when the minimum in total column ozone occurs. For this

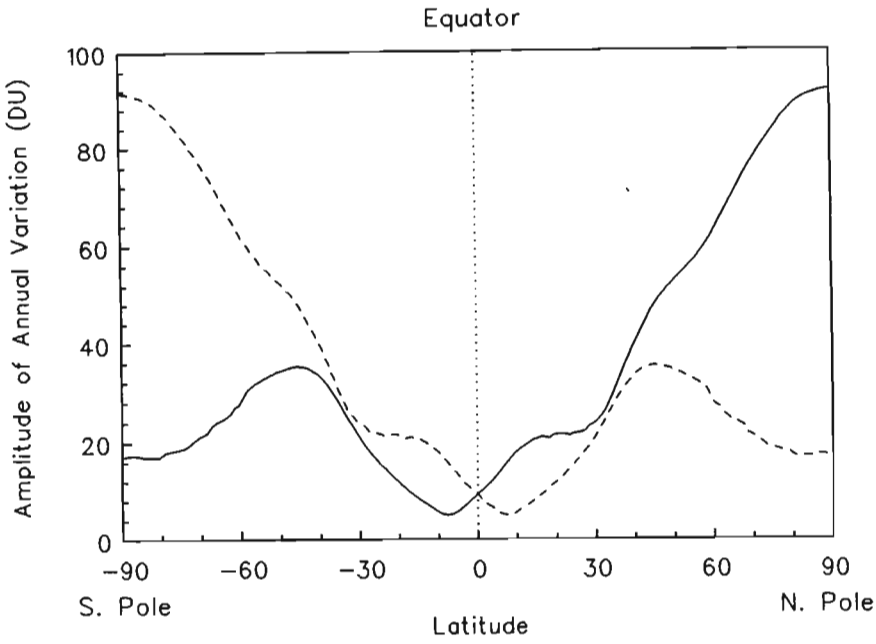


Figure 5.12: Meridional profile of the zonal means of the amplitude of the annual variation ($K_3$ values).

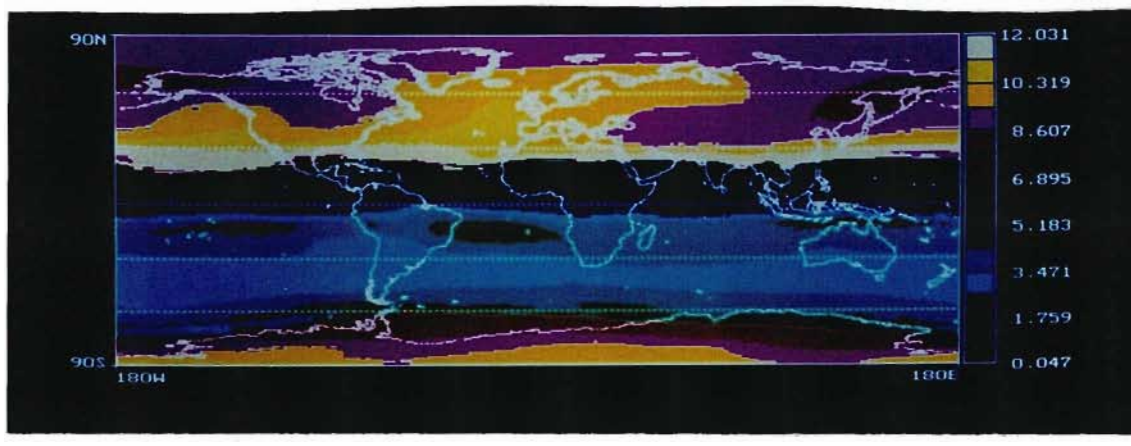


Figure 5.13: Annual variation phase coefficients in months. These data were obtained by subtracting the $K_4$ coefficients from 9 as discussed in the text.

reason phase data obtained from the model were subtracted from 9 to provide a direct indication of the timing of the annual minimum, for example a value of 1.0 is interpreted as 15 January, 1.5 as 31 January, 2.0 as 15 February etc. Although the interpretation of the phase is tied to its definition, the definition itself is arbitrary. Had cosine functions been used rather than sine functions, these data would appear very different. It is in the comparison of phases between locations

that they become instructive. The distribution of the $K_4$ values is approximately zonally symmetric. The annual variation over the Antarctic appears to be more in phase with the Northern Hemisphere than with other Southern Hemisphere locations. This supports the hypothesis that the Antarctic stratosphere is almost completely decoupled from the rest of the Southern Hemisphere stratosphere. Note also that the phase values appear to be more variable over the Southern Hemisphere where a local minimum is centered near 40°S. Together with results from section 5.4.3 it appears that locations near 40°S not only experience a greater annual variation in total column ozone compared with other Southern Hemisphere locations, but in addition the minimum in total column ozone occurs earlier in the year. This may have important consequences for surface UV radiation as discussed in section 5.5.12. The Northern Hemisphere annual variation phases appear to remain relatively constant with latitude.

The zonal means of the annual cycle phase are plotted in Figure 5.14. All

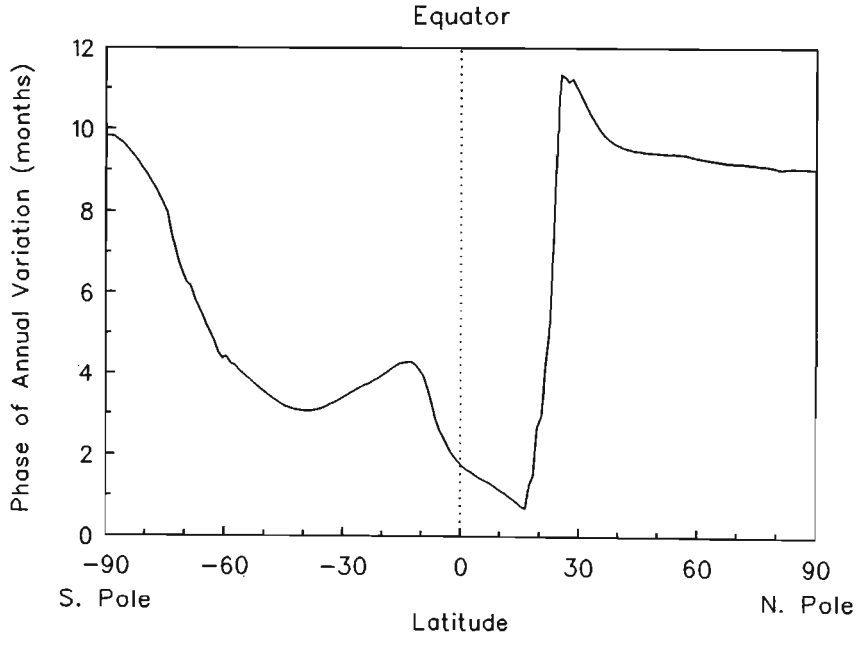


Figure 5.14: Meridional profile of the zonal means of the phase of the annual variation ($K_4$ values).

other factors being equal, it would be expected that the seasonal cycle in solar irradiance would cause the annual variation in Northern Hemisphere total column ozone to be 6 months out of phase with that of the Southern Hemisphere with a rapid phase change between the two signals across the equator. The fact that the steep gradient in phase values (see Figure 5.14) does not occur at the equator may be explained in terms of the arbitrary definition of the phase. The local

minimum in phase at 40°S indicates that locations at these latitudes can expect annual minima in monthly average total column ozone around 17 March while more equatorial stations near 15°S can expect annual minima near 23 April. The steep gradient in phase poleward of 60°S may be indicative of the isolation of the Antarctic stratosphere as discussed further in section 5.5.11.

### 5.4.5 The period of the annual variation ($K_5$ coefficient)

The periods of the annual variation are generally very close to the expected value of 12 months as shown in Figure 5.15. Areas shown in red in this plot have

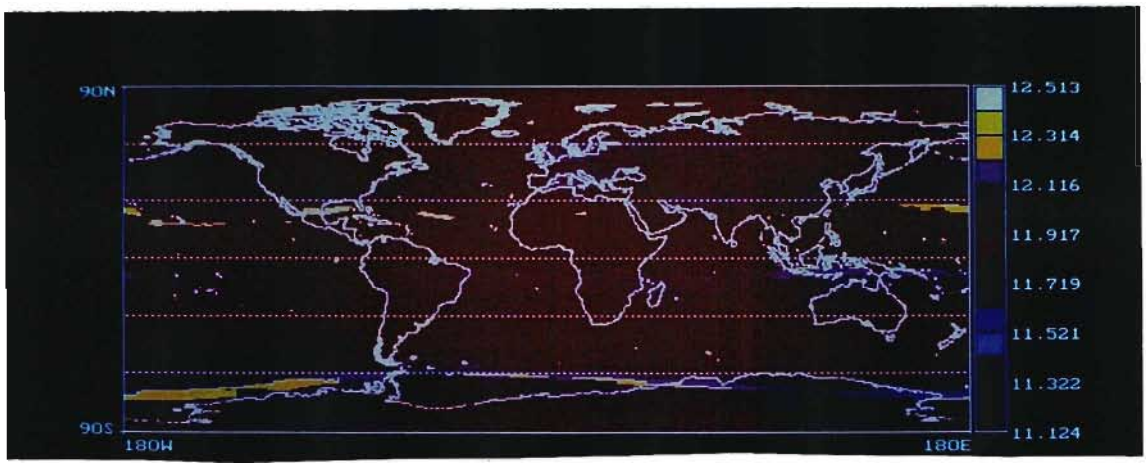


Figure 5.15: Annual variation period coefficients ($K_5$ values) in months.

periods expected for an annual variation and occur almost everywhere except for three zones near 70°S, 15°S and 45°N. Values do not differ from 12 months by more than 1 month for any location. When they do occur, the discrepancies are not large and result from errant model behavior. The three zones where this occurs are illustrated more clearly in Figure 5.16 where the zonal means are plotted. The horizontal dashed line in this figure shows the expected value of 12 months. Zonal mean values do not differ from 12 months by more than 0.1 months confirming the accuracy of the model when tracking the annual variation.

### 5.4.6 The period of the QBO ($K_8$ coefficient)

The period of the QBO is discussed before its amplitude and phase since large discrepancies between measured and expected values may cast doubt on the va-

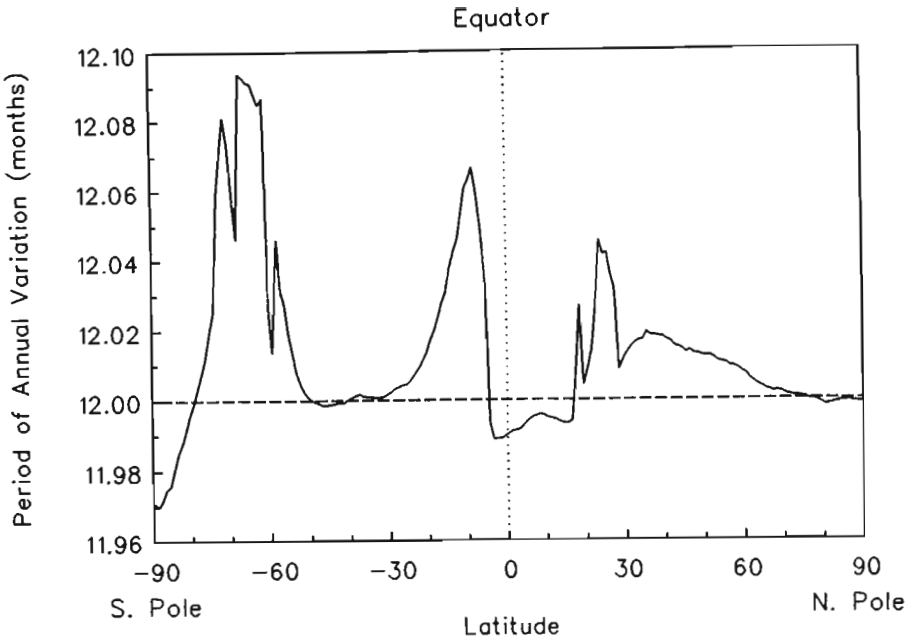


Figure 5.16: Meridional profile of the zonal means of the period of the annual variation ($K_5$ values).

lidity of the other two coefficients. Figure 5.17 shows a plot of the global QBO periods (in months) obtained from the model. Assuming valid QBO periods

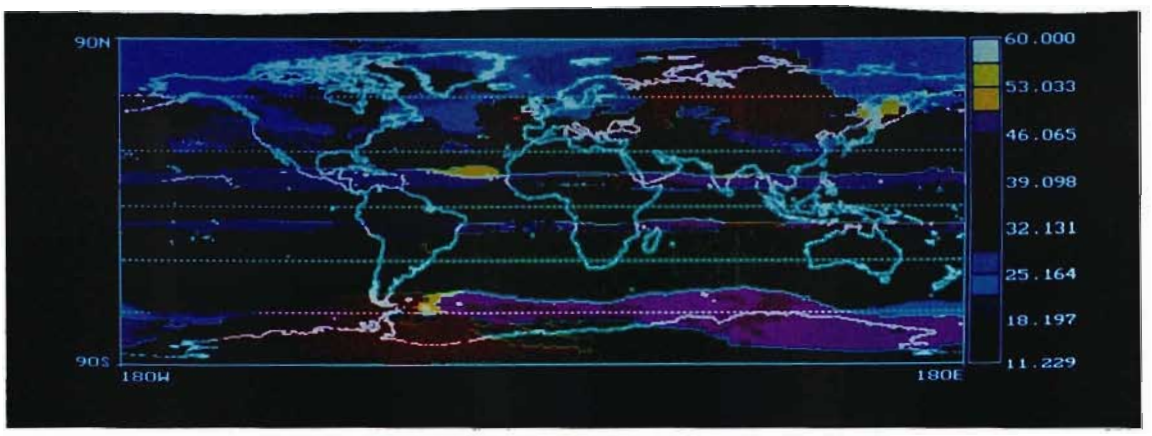


Figure 5.17: QBO period coefficients ($K_8$ values) in months.

of 22 to 34 months (Gray and Ruth, 1993), areas of Figure 5.17 shaded in hues of cyan and green indicate acceptable model behaviour. Analysis of these model coefficients indicates that 24% of the data fall outside of the range of 22 to 34 months. There are two possible reasons for this discrepancy:

1. Inaccurate model response to low amplitude signals: Investigation of the least squares algorithm used in this model shows that for QBO signals of

less than 3.5 DU amplitude, determination of the QBO period and phase is erratic. The model is unable to distinguish the signal from noise, resulting in poor tracking. It will be shown later that the two narrow zones of erroneous QBO periods near 10° to 15°N and 10°S result from low QBO amplitudes at these latitudes, as observed in previous QBO studies (Gray and Ruth, 1993).

2. Preferential tracking of the ENSO: The ENSO event, an episodic fluctuation in sea surface temperature off the South American coast, causes a reversal in the Walker cell circulation over the Pacific and an enhanced Hadley cell circulation (Henderson-Sellers and Robinson, 1987). These events, occurring every 2 to 7 years, disrupt the global mean circulation and may introduce longer period variations in the total column ozone record above certain locations. The ENSO influence was excluded from the current version of the model due to its episodic nature, although previous studies have modelled the ENSO as a four year cycle.

The zonal means of the QBO periods are plotted in Figure 5.18 together with the upper and lower bounds of previously reported values (Zerefos et al., 1992; Gray and Ruth, 1993). As discussed above, the two regions of anomalously high values bordering the equatorial region result from poor model tracking of low amplitude signals, while the midlatitude (Northern and Southern Hemispheres) and high latitude (Southern Hemisphere) discrepancies are thought to result from ENSO influences.

Based on these observations, the QBO amplitude and phase data sets have been screened, rejecting data at those locations for which QBO periods fall outside the range of 22 to 34 months. However, all signals of 3.5 DU amplitude or less are accepted, irrespective of the period, as it is the low amplitude of the signals that causes the period to be in error.

### 5.4.7 The amplitude of the QBO ($K_6$ coefficient)

Figure 5.19 shows the calculated QBO amplitudes where the data have been screened as described in section 5.4.6. Areas shown in black result from data being excluded by the screening process. These areas occur predominantly over the high latitudes of the Southern Hemisphere, the midlatitudes of the northern

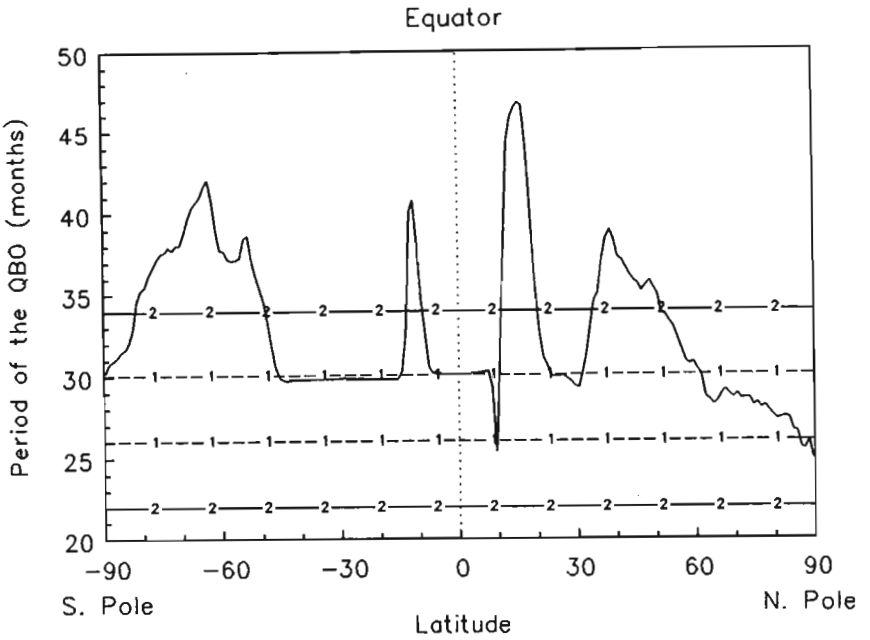


Figure 5.18: Meridional profile of the zonal means of the period of the QBO ($K_8$ values). Horizontal lines indicate upper and lower bounds of acceptable QBO periods as cited by 1) Zerefos et al., 1992: 26 to 30 months and 2) Gray and Ruth, 1993: 22 to 34 months.

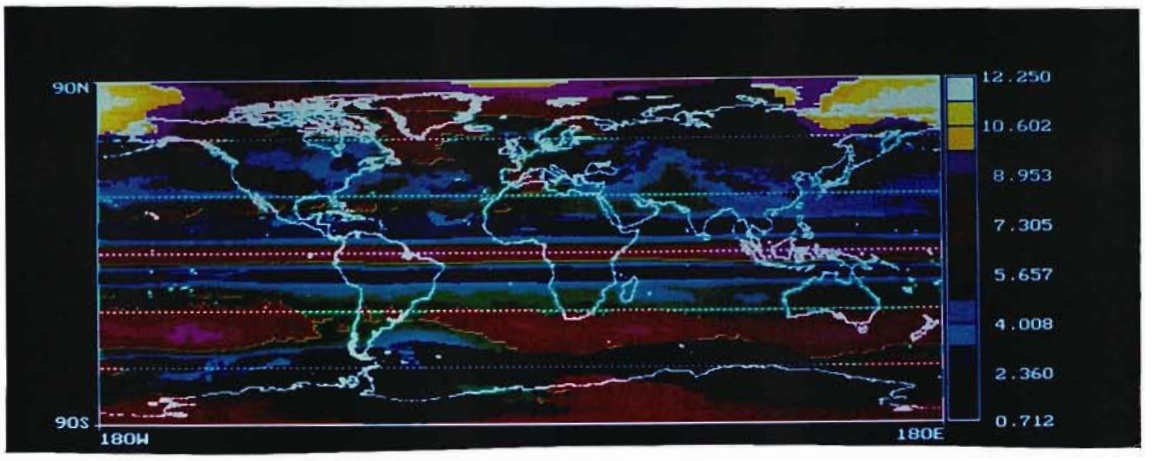


Figure 5.19: QBO amplitude coefficients ($K_6$ values) in DU. The black areas result from the data screening process.

Pacific and high latitudes over Asia. The most striking features of this plot are the two zones of weak QBO influence centered near 10° to 15°N and 10°S. This phenomenon, also evident in the plot of the QBO period coefficients, will be discussed further in sections 5.5.13 and 5.5.14. The influence of the QBO falls from between 5 and 9 DU in regions neighbouring these two zones to close to zero

within the zones. The zonal symmetry of the Southern Hemisphere distribution is perturbed by a local minimum in QBO amplitudes near the midlatitudes of the South American continent. As discussed in the previous chapter, the formation of a wintertime planetary wave 1 results in reduced variability in this region. The lower QBO amplitudes observed here may simply be an artifact of this lower variability.

The symmetry of the Northern Hemisphere distribution is disrupted by a local maximum in QBO amplitudes bordering the international date line north of 60°N. The greater background variability in total column ozone over the high latitudes of the Northern Hemisphere may account for the increase in QBO amplitude with latitude but does not explain the zonal asymmetry observed in this plot.

The two regions of low QBO amplitude bordering the equator, and the increase of QBO amplitude with latitude, particularly over the Northern Hemisphere, is shown more clearly in Figure 5.20 where the zonal means of the screened data are plotted together with the mirror image of the profile about the equator. The

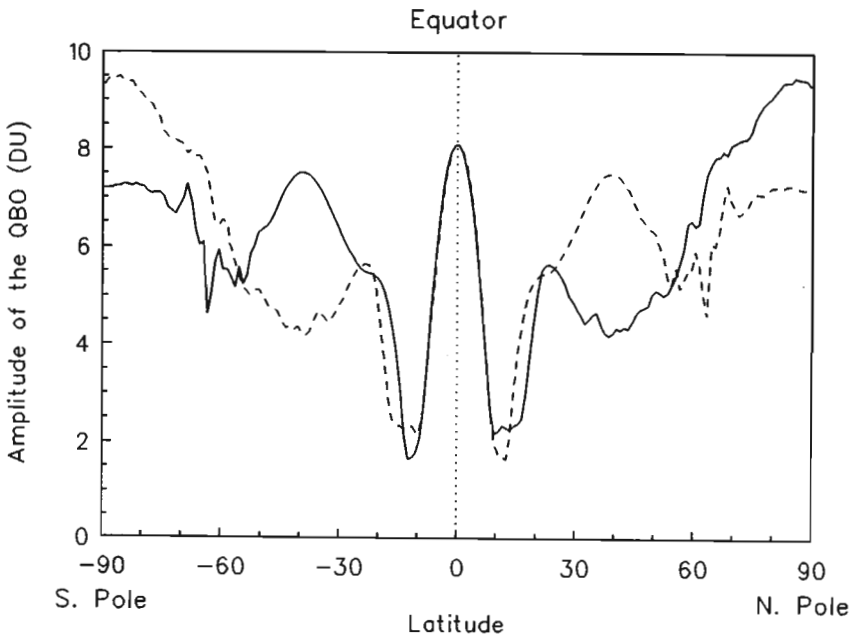


Figure 5.20: Meridional profile of the zonal means of the amplitude of the QBO ($K_6$ values).

$K_6$ coefficients show almost exact symmetry about the equator before reaching minima near 10°S and 10° to 15°N. Moving poleward, values rise, passing through local maxima in the midlatitudes where Southern Hemisphere QBO amplitudes are higher than those in the Northern Hemisphere. At high latitudes however, Northern Hemisphere QBO amplitudes exceed those of the Southern Hemisphere,

possibly as a result of inherently greater Northern Hemisphere variability in total column ozone. Note that the QBO amplitudes near the equator (~8 DU) are higher than those of the rest of the Southern Hemisphere.

### 5.4.8 The phase of the QBO ($K_7$ coefficient)

The model coefficients for the QBO phase are plotted in Figure 5.21, where again screened data have been used, creating black areas of excluded values. The most

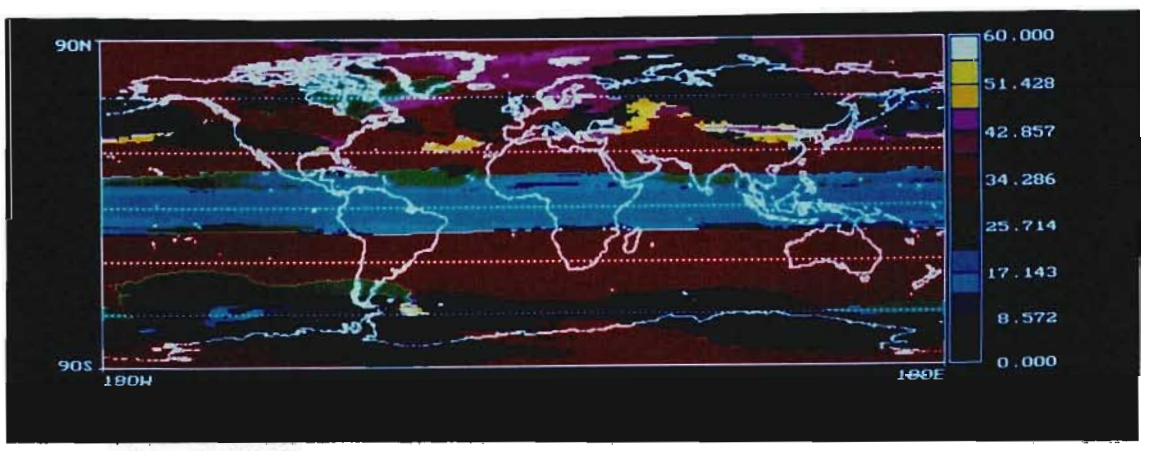


Figure 5.21: QBO phase coefficients ($K_7$ values) in months. The black areas result from the data screening process.

important feature of this distribution is the area bordered by approximately 15°S and 20°N which is 13 to 15 months (half a QBO cycle) out of phase with most locations over the rest of the globe. Again the feature is highlighted in the plot of the zonal means shown in Figure 5.22. Steep gradients in QBO phase are seen at the borders of this equatorial region. Other than a small increase in QBO phase near 40°N and a decrease near 65°S, phase values throughout the rest of the globe are nearly constant at between 30 and 40 months. Comparisons of the plots of QBO amplitude (Figures 5.19 and 5.20) and Figures 5.21 and 5.22 suggests that there is an isolated equatorial region where the QBO has a large influence on total column ozone variability and is half a QBO cycle out of phase with the QBO signals of the remainder of the globe. This behaviour of the QBO is discussed further in sections 5.5.13 and 5.5.14.

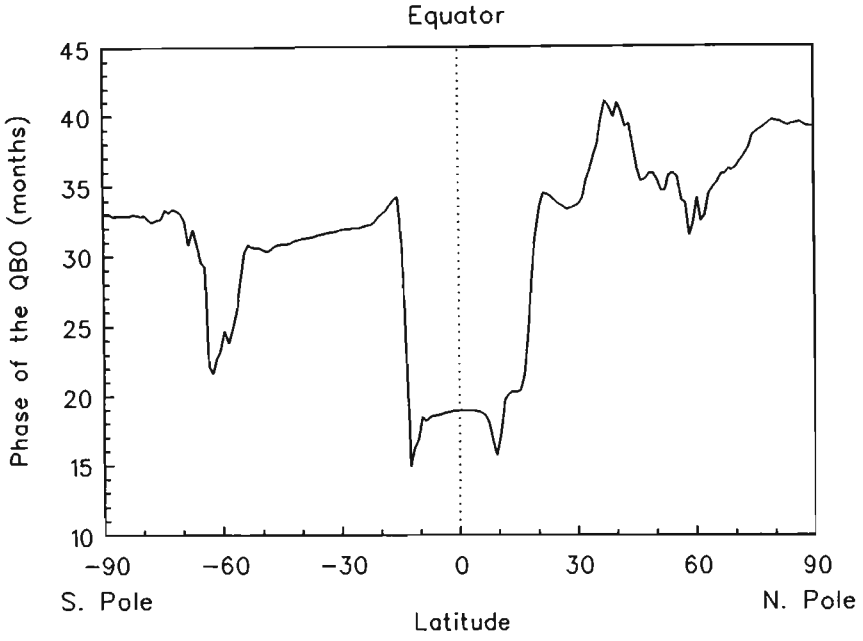


Figure 5.22: Meridional profile of the zonal means of the phase of the QBO ($K_7$ values).

### 5.4.9 The period of the solar cycle variation ($K_{11}$ coefficient)

As for the QBO, solar cycle periods are discussed before the phase and amplitude as erroneous periods indicate that the other two coefficients are unreliable. The $K_{11}$ model coefficients are plotted in Figure 5.23. The theoretical solar cycle

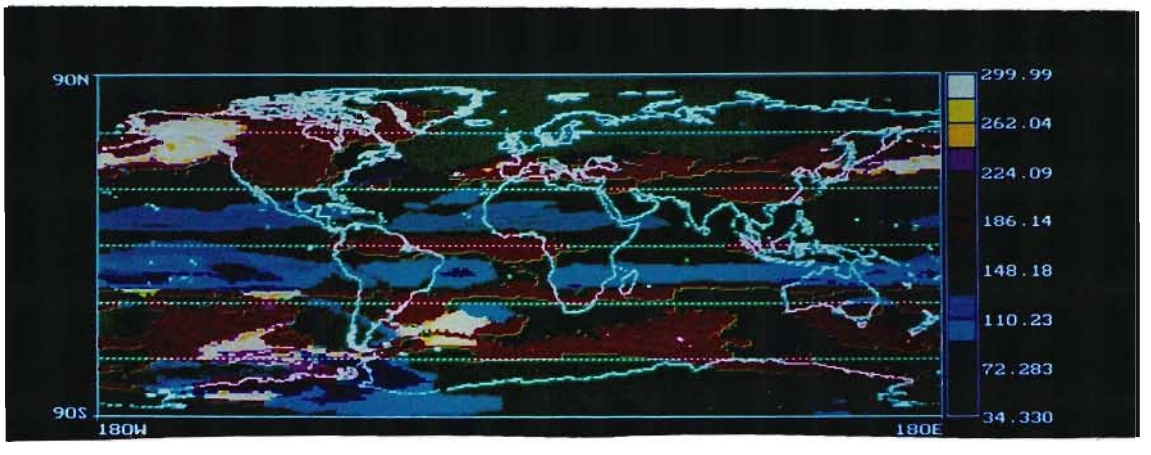


Figure 5.23: Solar cycle period coefficients ($K_{11}$ values) in months.

period of 11 years (132 months) is equivalent to a hue of olive green, observed mostly in the high latitudes of the Northern Hemisphere. The three solar cycle data sets have been screened, rejecting signals with periods falling outside the

range of 132 ± 6 months. However, as with the QBO, all signals of less than 3.5 DU amplitude were accepted. The resultant exclusion of 48% of the data is to be expected due to the statistical inaccuracy in tracking an 11 year signal in 14 years of data. The amplitude and phase distributions of more than half of the data points passing the screening process are discussed below. The zonal mean profile of the solar cycle amplitudes is not included here since no additional information on the features observed in Figure 5.23 may be derived.

### 5.4.10 The amplitude of the solar cycle variation ($K_9$ coefficient)

The solar cycle amplitude data set has been screened to exclude all data points whose periods differ by more than 6 months from the theoretical 132 month period. The amplitudes (in DU) for the remaining data points are plotted in Figure 5.24 where black areas result from the excluded data. The amplitudes tend to

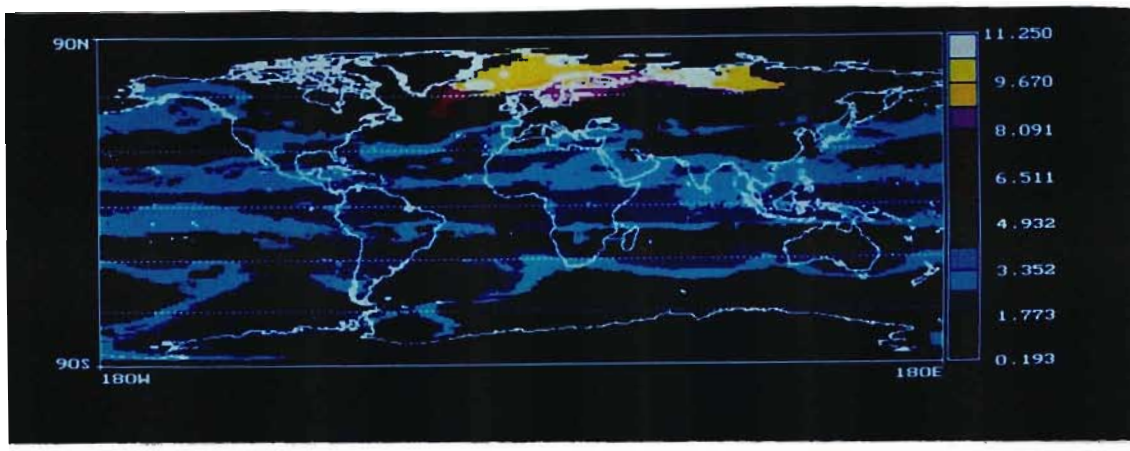


Figure 5.24: Solar cycle amplitude coefficients ($K_9$ values) in DU. The black areas result from the data screening process.

be low over most of the globe with enhanced values over the high latitudes of the Northern Hemisphere. These solar cycle amplitudes, as high as 11 DU (22 DU peak-to-peak), may again result from the high variability in total column ozone over the high latitudes of the Northern Hemisphere. Where data are available, the plot appears to be zonally symmetric. Zonal means of the screened data are plotted in Figure 5.25 where the large increase in solar cycle amplitudes in the Northern Hemisphere can be seen. The sparseness of the screened data in the Southern Hemisphere results in unrepresentative high latitude data in Figure

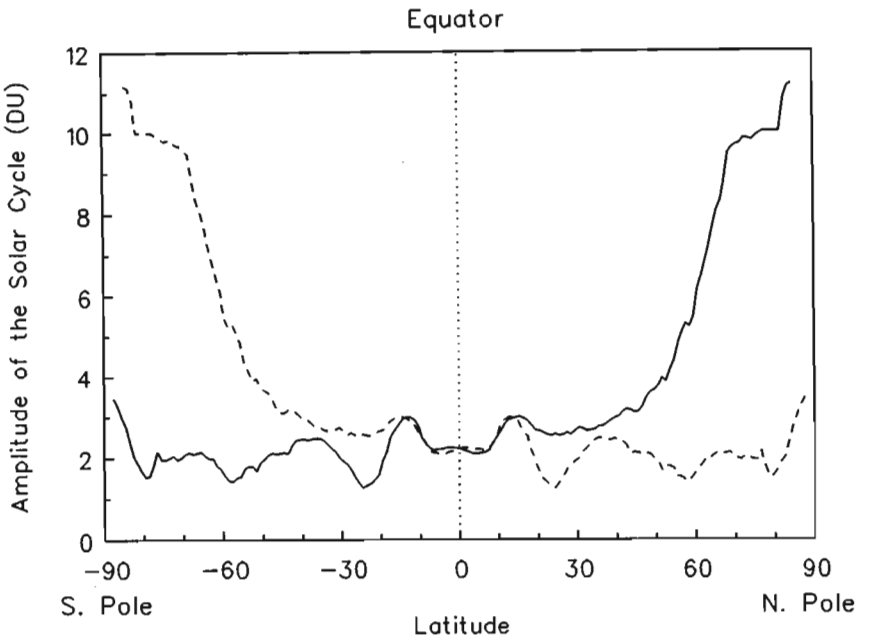


Figure 5.25: Meridional profile of the zonal means of the amplitude of the solar cycle ($K_9$ values).

5.25. A brief discussion of the expected effect of the solar cycle on total column ozone is given in section 5.5.16 below.

### 5.4.11 The phase of the solar cycle variation ($K_{10}$ coefficient)

Screened solar cycle phases are plotted in Figure 5.26 where black areas result from excluded data. The values appear chaotic as a result of the poor model tracking of the low amplitude solar cycle signals. The plot of the zonal means of the solar cycle phases shown in Figure 5.27 indicates that some Southern Hemisphere low latitude solar cycle signals may be out of phase by about half a solar cycle with locations at higher latitudes, although given the high rejection rate in the data, this is highly speculative. The high variability in monthly average total column ozone and the statistical inaccuracy in tracking an 11 year cycle in 14 years of data results in large uncertainties in the solar cycle model coefficients.

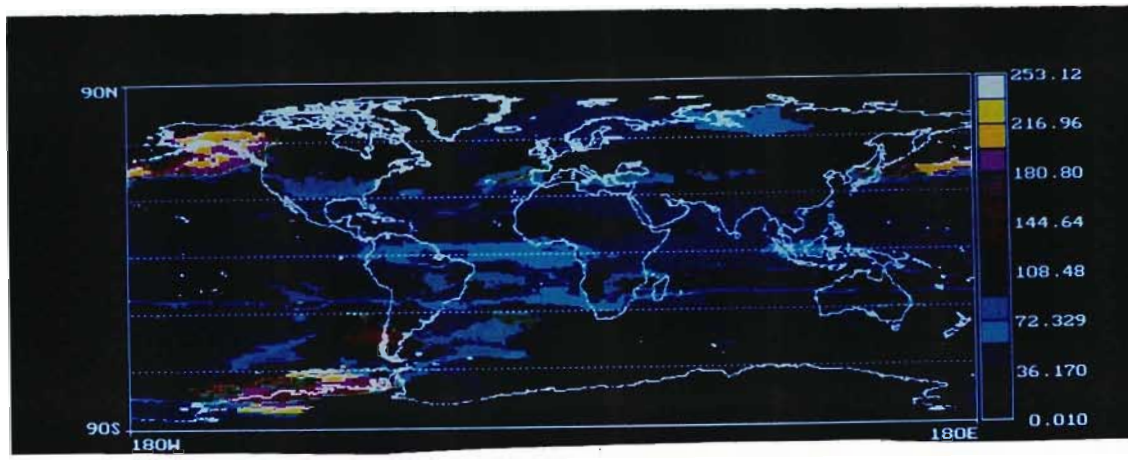


Figure 5.26: Solar cycle phase coefficients ($K_{10}$ values) in months. The black areas result from the data screening process.

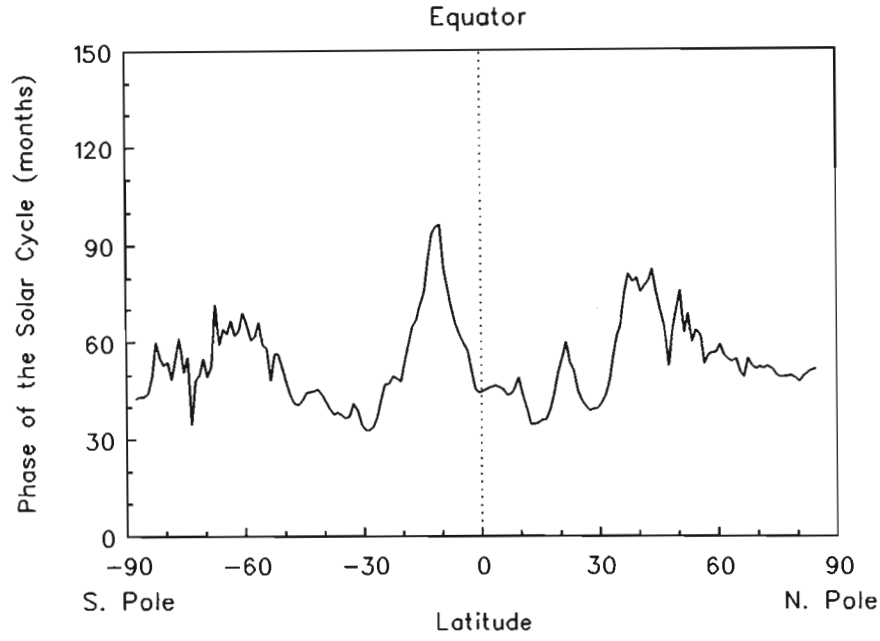


Figure 5.27: Meridional profile of the zonal means of the phase of the solar cycle ($K_{10}$ values).

### 5.4.12 The period of the semi-annual variation ($K_{14}$ coefficient)

The periods of the semi-annual oscillations are analyzed before the amplitudes and phases to provide validity checks for these data sets as for the QBO and solar cycle signals. The theoretical value of 6 months equates to the colour green in Figure 5.28 where the semi-annual oscillation periods are plotted. Much of the globe

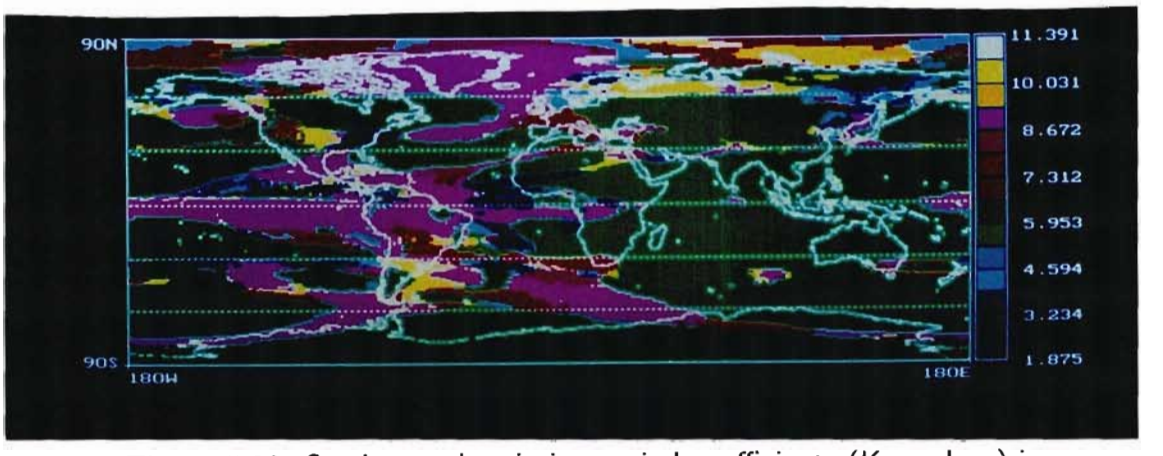


Figure 5.28: Semi-annual variation period coefficients ($K_{14}$ values) in months.

appears to contain acceptable data. The data set has been screened rejecting values outside the range of 6 $\pm$ 0.25 months (irrespective of the amplitude), resulting in 38% of the data set being ignored. The zonal mean profile has not been included here since it does not provide any additional insight into the semi-annual variation periods.

### 5.4.13 The amplitude of the semi-annual variation ($K_{12}$ coefficient)

Figure 5.29 shows the amplitude of the semi-annual variation in DU. Areas shown in black result from data being excluded by the screening process. Amplitudes are less than 12 DU over most of the globe but increase to approximately 40 DU over the Antarctic. The plot is zonally symmetric with small local maxima in the Northern Hemisphere. The zonal means plotted in Figure 5.30 show the rapid increase in semi-annual variation amplitudes over the high latitudes of the Southern Hemisphere. The presence of a signal at low latitudes may be attributed to the meridional movement of the Intertropical Convergence Zone (ITCZ) as discussed in section 5.5.18.

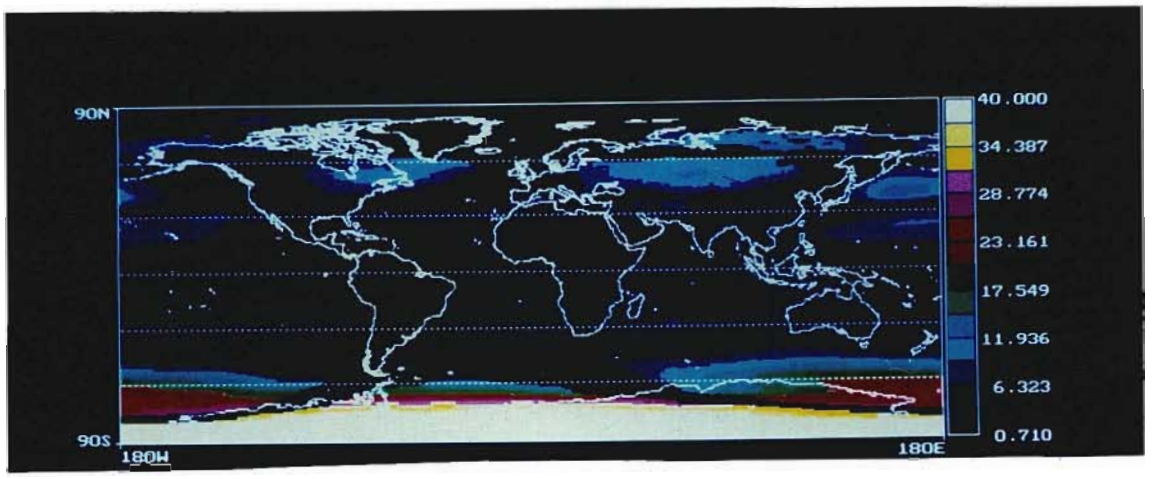


Figure 5.29: Semi-annual variation amplitude coefficients ($K_{12}$ values) in DU. The black areas result from the data screening process.

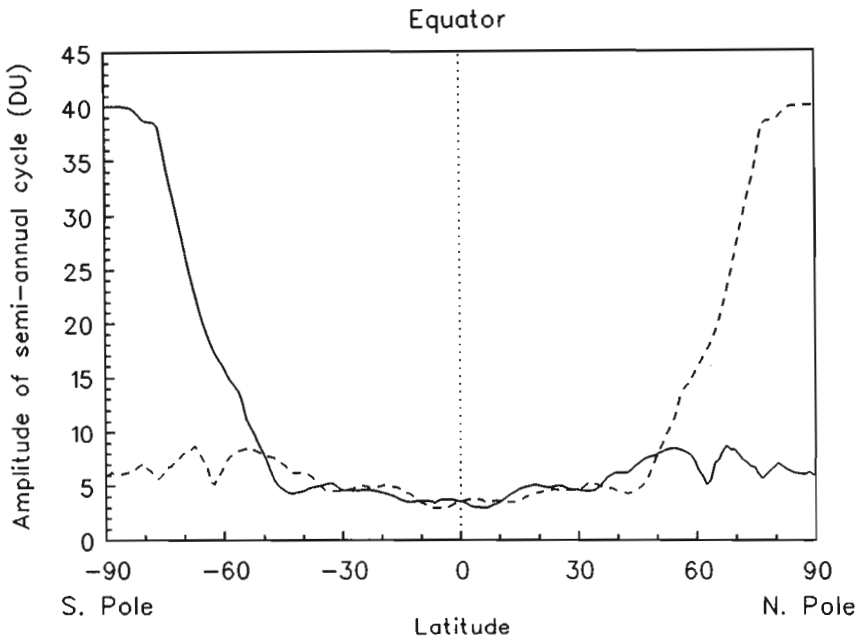


Figure 5.30: Meridional profile of the zonal means of the amplitude of the semi-annual variation ($K_{12}$ values).

### 5.4.14 The phase of the semi-annual variation ($K_{13}$ coefficient)

The phase of the semi-annual oscillation is plotted in Figure 5.31 where black areas result from excluded data. It appears that the phase of the semi-annual variation over the Antarctic differs from that over the rest of the globe. Together with results from section 5.4.13 is appears that total column ozone over the

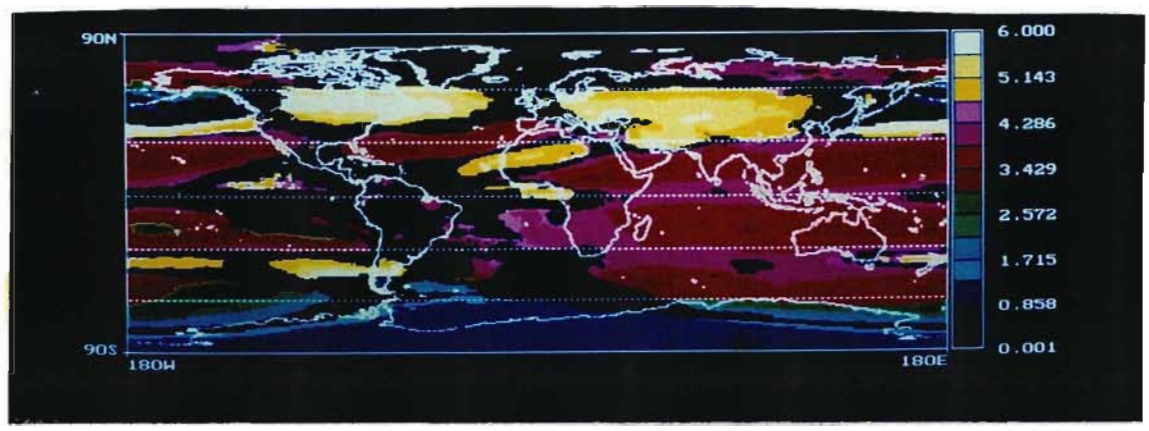


Figure 5.31: Semi-annual variation phase coefficients ($K_{13}$ values) in months. The black areas result from the data screening process.

Antarctic exhibits a large semi-annual variation which is out of phase with smaller amplitude semi-annual oscillations over the remainder of the globe. The phase difference is approximately half a semi-annual period (3 months) as shown in Figure 5.32 where the zonal means of the semi-annual phase are plotted. The

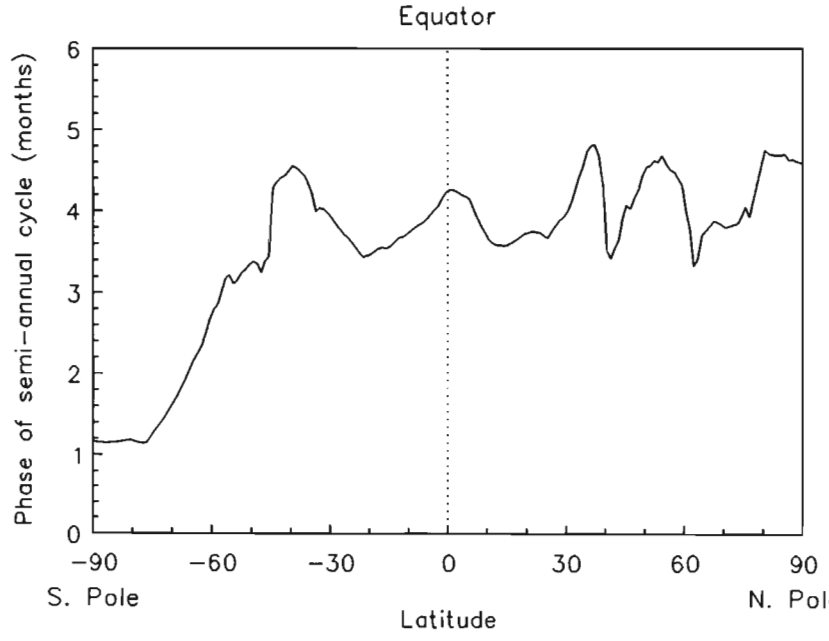


Figure 5.32: Meridional profile of the zonal means of the phase of the semi-annual variation ($K_{13}$ values).

large change in phase across the high latitudes of the Southern Hemisphere is also clear.

## 5.5 Discussion of model results

Although previous studies of long-term TOMS ozone variations have often neglected longitudinal resolution in favour of seasonal variability, this study has shown that in some cases the zonal variations in model coefficients are not insignificant and may be related to the occurrence of planetary waves. Furthermore, the high spatial resolution of the model has highlighted features in the coefficients that have been smoothed in previous studies that averaged input data series over many degrees of latitude. The drawback of the modelling technique used in this study is that no seasonal variability in model coefficients may be derived.

The observations reported in section 5.4 have raised a number of questions that require further discussion. These include:

1. What are the mechanisms accounting for the midlatitude ozone depletion observed in both hemispheres?

2. Why are negative trends at Northern Hemisphere midlatitudes (20°N to 60°N) greater than those in the midlatitudes of the Southern Hemisphere?

3. What causes regional differences in linear trends?

4. Can increases in tropospheric ozone offset decreases in stratospheric ozone so that no linear trend is measured and what implications does this have for surface UV radiation?

5. Are the hemispheric differences in model offset values related to dynamics or chemical ozone depletion before 1979 (see Figures 5.9 and 5.10)?

6. What are the sources of the zonal asymmetry in the model offset coefficients?

7. Why are annual amplitudes in the high latitudes of the Southern Hemisphere significantly smaller than those of the Northern Hemisphere?

8. What are the forcing mechanisms for the annual variation in Northern Hemisphere total column ozone?

9. What are the forcing mechanisms for the annual variation in Southern Hemisphere total column ozone?

10. What are the sources of the zonal asymmetry in the annual amplitudes?

11. Why is the annual variation in the high latitudes of the Southern Hemisphere out of phase with the rest of the Southern Hemisphere while Northern Hemisphere annual variation phases remain relatively constant with latitude?

12. What are the implications for surface UV radiation as a result of differences in the phase and amplitude of the annual variation?

13. What processes create a total ozone QBO at low, middle and high latitudes?

14. Why do QBO amplitudes increase with latitude through middle and high latitudes?

15. What is the mechanism for the phase change in the QBO near 10° to 15°N and 10°S?

16. Why are QBO periods highly variable?

17. In what way does the solar cycle influence the long-term variability in total column ozone?

18. Why are solar cycle amplitudes so much higher over the Arctic than over the rest of the globe?

19. What effect do seasonal changes in solar irradiance have on the global total column ozone distribution?

20. What is the forcing mechanism for the semi-annual oscillation?

21. Why are Antarctic semi-annual oscillation amplitudes so much higher than those over the rest of the globe?

Due to space constraints only the more important and interesting of these issues are expanded on below.

### 5.5.1 Mechanisms for ozone depletion at midlatitudes

Figures 5.7 and 5.8 showed that in addition to the high latitude ozone depletion, there are measurable negative trends at midlatitudes, while trends between 30°N

and 30°S are statistically insignificant. Antarctic ozone depletion takes place on the surface of PSCs (Chapter 1) while limited heterogeneous Arctic ozone destruction may occur on PSCs or on the surface of sulfate aerosols as discussed below. PSC dependent mechanisms cannot be invoked to account for midlatitude ozone depletion due to the elevated stratospheric temperatures preventing the formation of these clouds.

It has been suggested (Stolarski et al., 1991) that the negative ozone trends observed equatorward of the polar regions may result from export of chemically processed polar air to lower latitudes as the vortex is distorted by wave activity or as it breaks up. Chlorine rich air produced by heterogeneous processes could be transported away from the pole to regions where increased sunlight would enhance photochemically driven ozone destruction. It has been shown (Toumi et al., 1993) that high concentrations of $ClONO_2$ can be found at lower latitudes because of direct transport of polar air away from the vortex or mixing of ClO and $NO_2$ at the edge of the vortex. The role of $ClONO_2$ in the ozone depletion cycle was presented in Chapter 1.

A detailed summary of this 'ozone dilution effect', including model studies of its ability to support observed midlatitude ozone depletion, has been presented by Callis et al. (1991). However, this proposed mechanism may not account for the observed depletion at Northern Hemisphere midlatitudes unless there is sufficient chemical processing of polar air. It has been shown (Stolarski et al., 1991) that Northern Hemisphere trends are significantly larger than those predicted by models incorporating only homogeneous gas phase chemistry, suggesting that some form of heterogeneous processing must be taking place. Episodic PSC formation has been observed in the Arctic (Evans, 1990) together with associated denitrification of the stratosphere (Kondo et al., 1992). In general however, Arctic temperatures seldom fall low enough to allow the PSC formation required for heterogeneous chemistry. Whether the dilution effect could account for Southern Hemisphere midlatitude ozone depletion depends critically on the containment properties of the Antarctic vortex ie. if it acts as a 'flowing processor' or as an impermeable containment vessel for heterogeneous ozone destruction reactions (see review in Randel, 1993 a).

It has also been shown (Callis et al., 1991) that ozone depletion observed in Stratospheric Aerosol and Gas Experiment (SAGE) and SBUV data, between 1979 and 1985, at 50° latitude in both hemispheres, has occurred at altitudes

below 15.8 hPa. This further limits the dynamical and chemical processes that can be invoked to account for midlatitude ozone depletion.

In addition to the export of processed air following the breakup of the Antarctic vortex in late spring and early summer, ozone rich air over the midlatitudes mixes with ozone poor air over the pole, replenishing stratospheric ozone over the Antarctic. This results in a decrease in ozone at lower latitudes. Increases in the severity of the Antarctic ozone depletion may then result in a long-term net decline in total column ozone at all latitudes. However, the short ozone photochemical lifetimes at low latitudes permits rapid regeneration of ozone through the standard ozone production mechanisms.

Hofmann and Solomon (1989) showed that heterogeneous chemical processes occurring on middle to high latitude background and enhanced $H_2O$-$H_2SO_4$ aerosol surfaces in the region around 50 hPa may lead to $NO_2$ reductions and enhanced values of $HNO_3$. This in turn may perturb the ozone chemistry. Their model calculations of the effects of the March and April 1982 eruption of El Chichón suggested 10% reductions in local ozone at 50 hPa for January 1983. It has also been suggested (Prather, 1992) that these sulfate aerosols may act as substitutes for PSCs at midlatitudes, facilitating heterogeneous ozone depletion. Sulfate aerosols injected into the stratosphere by volcanic eruptions are able to suppress $NO_x$ concentrations by more than a factor of 10 relative to gas phase chemistry. When $NO_x$ levels fall below a critical threshold, eg. 0.6 p.p.b.v. at 24 km in the midlatitudes, the chlorine catalyzed loss of ozone can be has high as that occurring during the formation of the Antarctic ozone hole, ie. more than 50 p.p.b.v. per day (Prather, 1992).

In addition to long-term negative trends observed in the TOMS data, other studies using data sets from ground-based instruments have also observed midlatitude ozone depletion. Bojkov et al. (1990) made use of revised Dobson data from January 1958 to December 1986, from 29 Northern Hemisphere locations to derive total column ozone trends between 19°N and 64°N. Data for each month of the year were analyzed individually to determine seasonal trends. This analysis indicated significantly steeper negative trends during the winter (December to March) than during the summer (May to August), mostly at high latitudes, with trends in winter becoming more negative with increasing latitude. Both Figures 5.7 and 5.8 confirm Northern Hemisphere trends becoming more negative with increasing latitude.

### 5.5.2 Hemispheric differences in linear trends at midlatitudes

Figure 5.8 showed that the zonal means of the 14 year linear trends were more negative between 20°N and 60°N than over the same latitude range in the Southern Hemisphere. The dilution effect discussed in the previous section may account for these hemispheric differences. The greater integrity of the Antarctic vortex reduces export of processed air to midlatitudes, accounting for the smaller rates of depletion observed there. Larger Northern Hemisphere planetary wave activity and reduced vortex containment results in more frequent Stratospheric Sudden Warmings (SSWs) and greater transport of vortex air to midlatitudes.

Hemispheric differences in linear trends were also noted by Herman et al. (1993) who made use of V6 TOMS data from January 1979 to December 1991 to derive long-term trends in total column ozone. Zonal means were calculated for each month of the year to permit determination of linear trends as a function of season. Long-term trends from Herman et al. are compared with the results obtained from this study in Figure 5.33. Although the trends derived in this study contain no

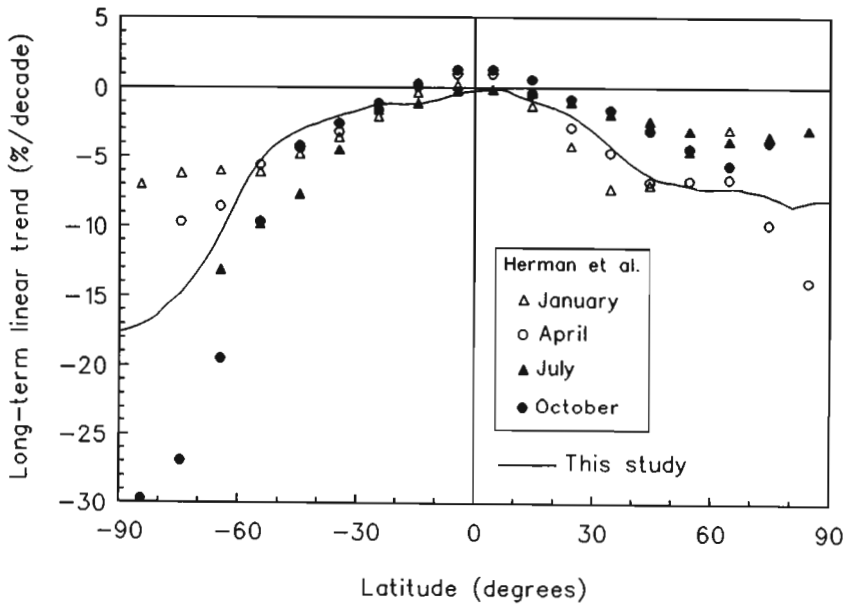


Figure 5.33: Comparison of long-term linear trends from Herman et al. (1993) and the results obtained from this study.

seasonal dependence, the trends should fall within the bounds of values quoted by Herman et al.. The data from this study, plotted in Figure 5.33, were derived from the model coefficients by first calculating zonal means of the long-term linear

trend ($K_1$) and offset ($K_2$) model coefficients. Trends (in % per decade, PD) were then calculated as a percentage of the long-term linear average value as follows:

$$PD = \left[ \frac{\overline{K_1} \times 120}{\overline{K_2} + (\overline{K_1} \times 84)} \right] \times 100 \tag{5.4}$$

where the overbar denotes zonal mean values. In general, the values obtained from this study fall within the bounds of values reported by Herman et al., although there appears to be some discrepancy between 30°S and 60°S where the derived trends are smaller than those of Herman et al..

The hemispheric differences in long-term trends appear to be strongly seasonally dependent, as shown by the Herman et al. (1993) data plotted in Figure 5.33. During the months of July and October, midlatitude Southern Hemisphere trends appear to be more negative than those of the Northern Hemisphere while the reverse is true for the months of January and April. This may be indicative of export of ozone poor air (or chlorine rich air) from the Arctic and Antarctic vortex as it is during the months of vortex formation and following vortex breakup that trends are more negative than those of the opposite hemisphere. Stolarski et al. (1991) also showed that the Northern Hemisphere midlatitude ozone trend shows a significant seasonal cycle with a maximum in late winter (February and March).

It is clear from Figure 5.33 that the large negative trends observed in the high latitudes of the Southern Hemisphere in this study are strongly influenced by the large October linear trend resulting from the annual formation of the Antarctic ozone hole. Niu et al. (1992) have shown that the most severe negative trends in total ozone are observed between 0° and 80°W during October.

### 5.5.3 Regional differences in linear trends

Figure 5.7 showed that there were regional differences in linear trends, particularly over the Northern Hemisphere. Similar observations have been reported previously (Bojkov et al., 1990), where trends in Japan were noted to be considerably less negative than those in North America and Europe. Investigation of Figures 5.7 and 5.9 shows that although the trend in terms of DU/month may be only slightly less negative in Japan than in North America, the offset value is higher in Japan and the trend in terms of %/year may be reduced. This introduces ambiguity as to whether long-term trends should be quoted in terms of DU/year or %/year.

The zonal mean trend coefficients obtained in this study have been calculated as a percentage of the zonal means offset coefficients and are plotted in Figure 5.34 together with the trend in DU/month obtained from Figure 5.8. The trend in

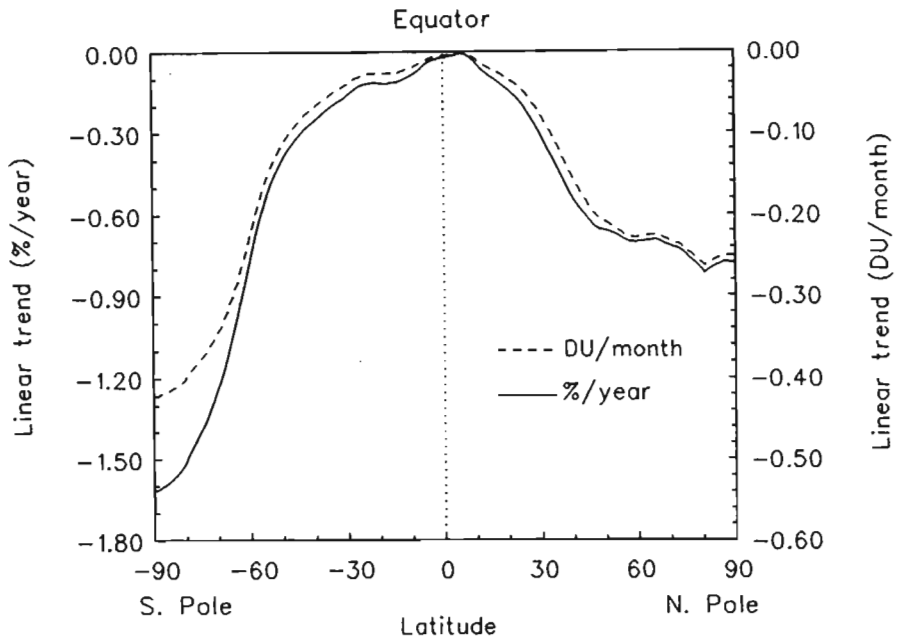


Figure 5.34: The zonal mean linear trends as a percentage of the zonal mean offset value (solid line) and in terms of DU/month (dashed line).

terms of %/year (solid line) shows enhanced Antarctic ozone depletion compared to the trend in terms of DU/month (dashed line). Previous studies, using version 6 TOMS data for 11.6 years from November 1978 to May 1990 (Stolarski et al., 1991), reported negative trends of about 0.4 to 0.5 %/year from 40°N to 60°N. These are smaller than the values of 0.55 to 0.70 %/year obtained in this study, possibly as a result of the anomalously low Northern hemisphere total column ozone during 1992 (Gleason et al., 1993). It is in this region of the globe where negative trends cause greatest concern since much of the world's population is concentrated here.

Niu et al. (1992) also showed that linear trends in total column ozone are dependent on geographical location. For January, February and March in the Northern Hemisphere, long-term trends are negative over the Eastern Hemisphere and positive over the Western Hemisphere. No mechanism was proposed to account for these regional differences. It is possible that during the winter months of the Northern Hemisphere, when some degree of ozone depletion occurs, topographic differences may result in enhanced export of processed air from sectors of the Arctic vortex and impeded transport from other sectors. No support for this

speculation has been found in the literature and suggests an avenue for further investigation.

### 5.5.4 Increases in tropospheric ozone

Tropospheric ozone levels appear to have been increasing between 50°N and 50°S over the period 1979 to 1987 (Fishman et al., 1990) although the predominant build-up has occurred in the Northern Hemisphere (Oltmans, 1985; Volz and Kley, 1988; Crutzen, 1988). This increase, mostly due to tropospheric ozone production through enhanced emissions of $NO_x$ by industry and transportation, may partially offset decreases in stratospheric ozone so that total column ozone amounts may stay relatively constant or decline less steeply. However, since TOMS shows lower sensitivity to tropospheric ozone compared to ground-based instruments such as the Dobson spectrophotometer (see Chapter 3), this may be a complicating factor when comparing long-term trends from these two data sets (Stolarski et al., 1991). Because of the increase in Northern Hemisphere tropospheric ozone, TOMS will slightly over-estimate the long-term trends in this region.

Relative increases in tropospheric ozone with respect to stratospheric ozone may have important implications for the determination of UVB trends from total column ozone trends. Tropospheric ozone can be disproportionately effective as a filter against solar UV-radiation under circumstances when most scattering of solar radiation by air molecules and dust occurs in the troposphere, ie. during the summer or at low latitudes (Brühl and Crutzen, 1989). It is therefore possible that Northern Hemisphere UV levels are decreasing in spite of the decrease in total column ozone due to increases in tropospheric ozone. The effect of changes in total column ozone on surface UVB irradiance will be discussed in greater detail in Chapter 6.

### 5.5.5 Differences in model offset coefficients at high latitudes

Figures 5.9 and 5.10 show that at high latitudes of the Southern Hemisphere, offset coefficients are considerably lower than those at high latitudes of the Northern Hemisphere. There are three possible reasons for this:

1. Antarctic ozone depletion occurred to some extent prior to 1979. This appears to be an unlikely source for the hemispheric differences in high latitude model offset values since no Antarctic ozone hole was reported before the early 1980s.

2. The current annual formation of the Antarctic ozone hole causes a net reduction in the offset value of the long-term linear trend.

3. The isolation of the Antarctic vortex impedes the poleward flow of ozone rich air from the tropics, although this would only occur during the winter months. In comparison, the relative permeability of the Northern Hemisphere vortex would allow transport of ozone rich air over the Arctic.

In examining total column ozone data between 1958 and 1962, Rood (1986) showed that very low north polar values were observed over this period while there was no indication of anomalously low Antarctic total column ozone. It was explained that the early 1960s was a period of abnormal climatology caused by persistent planetary wave activity. It was also suggested that any systematic trend in planetary wave activity could affect how much ozone is transported into the polar regions. This possibility was investigated thoroughly in Chapter 4.

### 5.5.6 Zonal asymmetry in model offset coefficients

Significant zonal asymmetry was observed in the model offset coefficients (see Figure 5.9). In the Southern Hemisphere, a local minimum at 60°S occurs across the Antarctic peninsula and South American continent. This region coincides with the minimum in the wintertime planetary wave 1 observed in the total column ozone distribution. Figure 5.35 shows the Southern Hemisphere total column ozone distribution on 16 October 1993 obtained from the TOMS instrument on board Meteor-3. The large wave 1 perturbation is clearly visible with its minimum over the South American continent. It is also evident that the Antarctic ozone hole has been distorted and pushed towards the Southern Atlantic.

The consistent annual formation of a planetary wave 1 in the midlatitudes of the Southern Hemisphere elevates ozone levels south of Australia and depresses ozone levels across the South American continent. Similar wave forcing may also account for the zonal asymmetry in the Northern Hemisphere distribution of model offset values. However, this is purely speculative since a detailed examination of

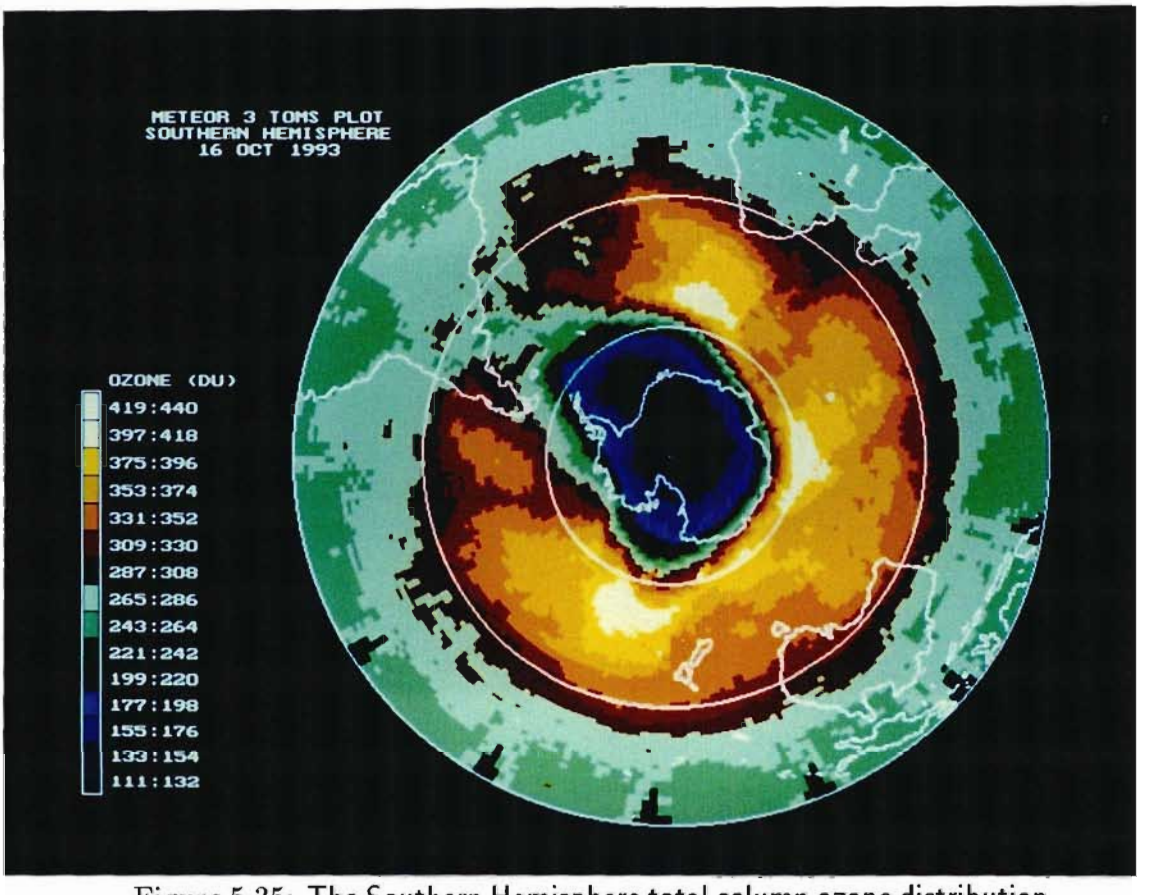


Figure 5.35: The Southern Hemisphere total column ozone distribution as seen by Meteor-3 TOMS on 16 October 1993.

Northern Hemisphere planetary wave morphology has not been conducted in this thesis. No conclusive mechanism can be proposed for the zonal minimum in model offset coefficients over the Pacific Ocean or equivalently, the zonal maximum in offset coefficients over the central Atlantic. However, it has been mentioned in Chapter 3 that the occurrence of persistent marine stratocumulus over this region of the Atlantic may cause TOMS to over estimate total column ozone, possibly accounting for the observed higher offset values.

### 5.5.7 Hemispheric differences in high latitude annual amplitudes

Figures 5.11 and 5.12 show that the amplitude of the annual cycle can be up to 5 times higher over the high latitudes of the Northern Hemisphere than over the same region of the Southern Hemisphere. It is unlikely that this discrepancy arises

from the increased high latitude variability in total column ozone as this would affect both polar regions equally. A more likely explanation is that the seasonal cycle in Antarctic total column ozone is altered fundamentally by the annual formation of the ozone hole. Before 1980, when Antarctic ozone depletion was minimal, the natural seasonal cycle in ozone exhibited a minimum during March and April and a maximum during November and December as shown in Figure 5.36. The zonal mean total column ozone is plotted as a function of latitude and

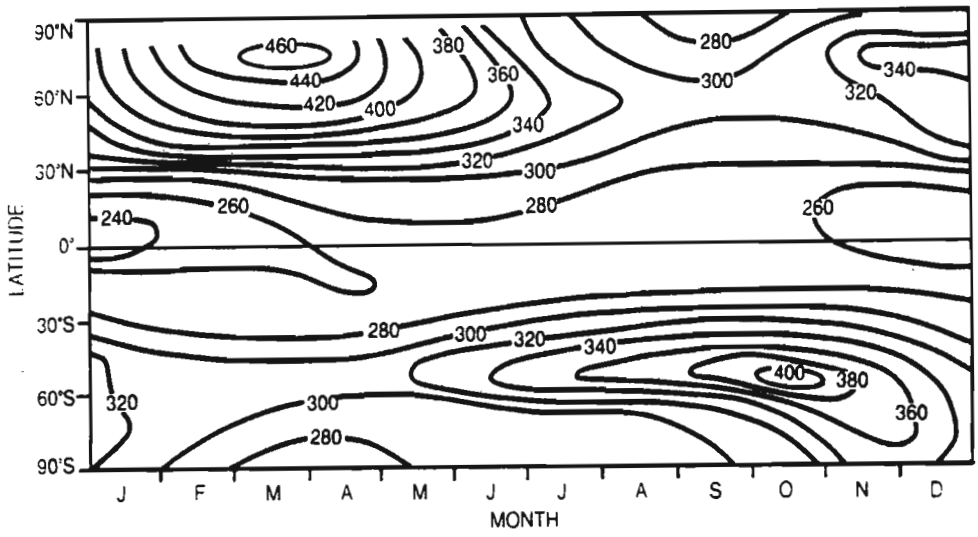


Figure 5.36: Climatological zonal mean total column ozone over the entire globe before the onset of the Antarctic ozone hole. Adapted from Salby and Garcia, 1990.

time. Southern Hemisphere midlatitude total ozone maximizes during October while high latitude values maximize during November and December.

However, since the 1980s the formation of the Antarctic ozone hole has occurred during those months when the annual ozone cycle would have reached its maximum. This significantly alters the observed seasonal variation. Figure 5.37 shows a plot of zonal mean total column ozone at 80°N and 80°S for each day in 1979 and 1992 obtained from the V6 TOMS data base. It is clear that in both hemispheres the depletion has occurred predominantly during the months following winter. The Northern Hemisphere zonal mean profile shows a steady decline through the summer, reaching its minimum just before the onset of polar night. The shape of this profile indicates a clear annual cycle once interpolation across the polar cap fills the missing data through the first and last 60 days of the year. The Southern Hemisphere zonal mean profiles suggest that the majority of the

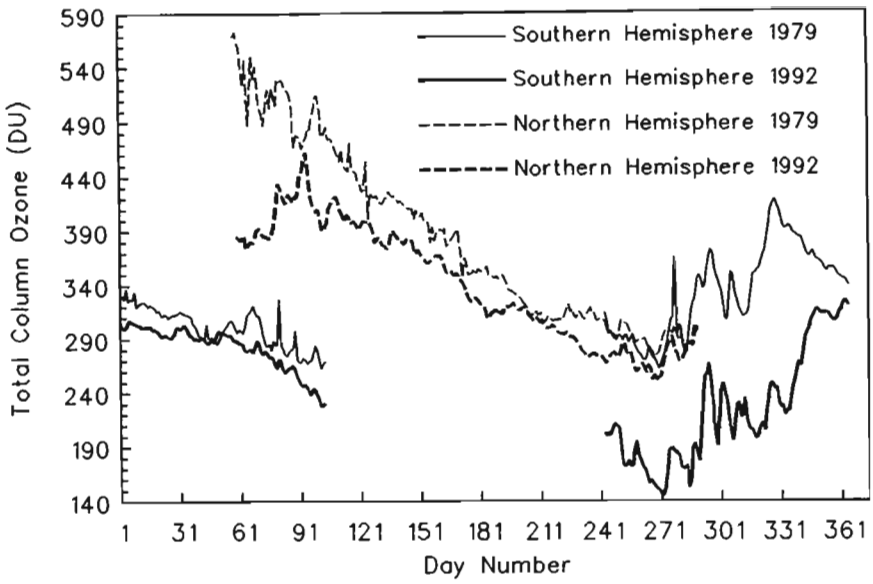


Figure 5.37: Zonal mean total column ozone along 80°N and 80°S for each day of 1979 and 1992.

depletion from 1979 to 1992 has occurred during those months where ozone previously maximized. The 1992 post-winter data are generally lower than values observed early in the year, unlike 1979. Together with the interpolation across the polar cap which tends to cause a local maximum in the profile during the winter, a semi-annual cycle is observed. Comparison of Figures 5.11 and 5.29 show that over the Antarctic, the semi-annual amplitude may be more than twice as large as the annual amplitude.

### 5.5.8 The Northern Hemisphere annual cycle in total column ozone

In this study it was noted that the amplitude of the annual variation in the Northern Hemisphere total ozone increases with latitude. Bojkov et al. (1990) reported the change in ozone between the annual minimum and the annual maximum to range from 15 to 25% in the 30° to 39°N zone to 35 to 50% in the 53° to 64°N zone. Comparative results obtained from this study are 9 to 28% in the 30° to 39°N zone and 24 to 42% in the 53° to 64°N zone.

Perliski et al. (1989) showed that at high latitudes and low altitudes, modelled ozone concentrations increase in the winter due to transport and decrease in the

summer due to chemical destruction. In the middle stratosphere, modelled ozone variations are large due to the annual variation in the odd-oxygen production rate, and in the upper stratosphere the modelled annual ozone variation is caused by the large annual oscillation in temperature. Most of these effects are related to seasonal differences in extraterrestrial solar irradiance which are discussed further in section 5.5.17.

### 5.5.9 Forcing mechanisms for the Southern Hemisphere total ozone annual variation

The annual variation in Southern Hemisphere total column ozone appears to be determined primarily by atmospheric dynamics and, to a lesser extent, through the seasonal shift in the ozone chemical equilibrium discussed in section 5.5.8. During the Southern Hemisphere winter, as the Antarctic ozone hole forms, distortion of the meridional circulation by the circumpolar vortex impedes the poleward flow of ozone rich air both through changes in the eddy transport (exclusion of planetary wave transport to high latitudes – see Chapter 4; Chen and Robinson, 1992) and through shifts in the mean circulation cells. This results in the formation of a midlatitude ridge of ozone rich air which is disturbed by zonally propagating planetary waves. Total column ozone above midlatitude locations therefore rises through the winter and spring as the ridge forms equatorward of the polar vortex. At the same time, chlorine catalyzed ozone depletion reduces total column ozone levels over the Antarctic so that a minimum is reached in October, November or December (depending on the lifetime of the vortex). Following the breakup of the vortex, ozone rich air from the midlatitudes mixes with ozone poor air over the pole, as discussed in section 5.5.1. This creates a decrease in middle and low latitude total column ozone during the last few months of the year while high latitude total ozone increases rapidly over this period.

Annual variations in Southern Hemisphere total column ozone will be examined in greater detail in Chapter 6 where 14 years of total column ozone above 5 South African cities are analyzed.

### 5.5.10 Zonal asymmetry in the annual amplitudes

Zonal asymmetries were observed in the distribution of the annual amplitude coefficients, particularly in the Southern Hemisphere. Between 30° and 60°S, annual amplitudes show a zonal minimum across South America and the Antarctic peninsula. The seasonal cycle at these latitudes arises predominantly from the formation of the wintertime crescent shaped midlatitude ridge as shown in Figure 5.35. The crescent shape, resulting from the wave 1 forcing, has its minimum over South America and the Antarctic peninsula. The amplitude of the annual cycle should therefore be a minimum over this region and show a maximum south of Australia and New Zealand (see Figure 5.11).

### 5.5.11 Meridional variations in the phase of the annual variation

It is clear from the discussion of the forcing mechanisms for the Southern Hemisphere annual variation that high latitude total column ozone will be out of phase with the annual variation at middle and low latitudes. For most of the Northern Hemisphere, the annual minimum in total column ozone occurs during early winter (October to December) causing it to be in phase with the Antarctic depletion.

### 5.5.12 Annual variation in total column ozone and its influence on surface UVB

Observations of the phase and amplitude of the annual variation (section 5.4.3 and section 5.4.4) indicate that there is a local maximum in amplitude and a local minimum in phase near 45°S. The large amplitude of the annual ozone cycle at these locations is expected to result in large variations in surface UV. Furthermore, since the phase values are lower, the minimum in total column ozone occurs closer to the summer solstice with the result that maximum UV levels will be even higher than if the phase were constant. This has important implications for the optimal siting of Southern Hemisphere UV observation stations for which large annual variations in surface UV are important (Bodeker and McKenzie, 1993).

### 5.5.13 The QBO in total column ozone at low, middle and high latitudes

The QBO in equatorial zonal winds is a cycling of stratospheric wind direction with a period of 22 to 34 months. A mechanism involving the transfer of momentum to the zonal flow through vertically propagating equatorial waves has been proposed to explain the oscillation (Lindzen and Holton, 1968; Holton and Lindzen, 1972). A review of these processes is given by Gray and Pyle (1989) together with results from a model in which the QBO in zonal winds was driven by momentum deposition associated with the presence of thermally dissipating Kelvin and Rossby-gravity waves. Any discussion of the QBO is therefore directly related to planetary wave propagation. The total ozone QBO has been observed at all latitudes and is well documented (see review in Gray and Pyle, 1989). Previous studies have also observed the largest amplitudes of the ozone QBO at high latitudes. In the work presented in this chapter, it was shown that QBO amplitudes were higher over the high latitudes of the Northern Hemisphere than over the equator, although this was not true for the Southern Hemisphere.

It has been found that during the westerly phase of the equatorial QBO, the Antarctic polar night jet is stronger resulting in a more severe winter ozone depletion. Rood (1986) reported a change in the QBO between the early and middle 1960s while Angell and Korshover (1973, 1976, 1978) attributed the global ozone minimum in the early 1960s to these exceptionally strong QBO events. Krueger et al. (1989) indicated that the shallow ozone holes of 1984 and 1986 may have been correlated with the QBO. Schoeberl et al. (1989) showed that variations in the maximum in total ozone between 50°S and 60°S are related to a QBO in the eddy transport.

Garcia and Solomon (1987) pointed out that both the Antarctic minimum temperature at 100 hPa and the minimum total ozone values were correlated with the phase of the equatorial QBO, as indicated by the sign of the 50 hPa zonal winds over Singapore, for the period 1979 to 1986. Using TOMS data and 70 hPa temperature data at the South Pole, Newman and Randel (1988) showed that the interannual variability in these quantities is dominated by an approximate 2 year periodicity associated with the QBO. This association shows cold temperatures during the westerly phase (1980, 1983, 1985, 1987) and warm temperatures during the easterly phase (1979, 1984, 1986) of the QBO. It was suggested that

any long-term trends be assessed in light of this association (ie. changes between either easterly or westerly QBO phases). Mancini et al. (1991) showed the possible effects of the QBO on the ozone distribution in a study including a 2-dimensional model parameterization of Kelvin and Rossby-gravity wave forcing in the lower equatorial stratosphere. The observed behaviour indicated a QBO induced temperature effect on PSC formation and the activation of heterogeneous chemistry. It was concluded that the high dynamical forcing of the ozone QBO at the middle to high latitudes through the secondary circulation induced by the equatorial temperature oscillation is comparable to the chemical forcing through heterogeneous processes.

Tolson (1981) presented an analysis of 7 years of total column ozone data obtained from the BUV instrument on board the Nimbus-4 satellite. Results of this analysis suggested that at 50°S, the QBO component of the total ozone variation had a magnitude (peak-to-peak) of 12 $\pm$ 2 DU. Values from Figure 5.20 show a QBO amplitude of 12.6 DU (peak-to-peak) comparing well with the results obtained by Tolson. Model studies by Gray and Pyle (1989) show QBO amplitudes of 25 DU at 50°S, approximately twice as large as the values observed by Tolson and in this model.

An explanation of the mechanism for the QBO in total column ozone begins with an investigation of the temperature changes associated with the equatorial QBO in stratospheric winds (Gray and Pyle, 1989). A forcing of the zonal wind at the equator requires a corresponding temperature adjustment to maintain thermal wind balance (see section 2.4.2). This coupling between the equatorial wind and temperature QBO in the lower stratosphere through the thermal wind balance holds at low latitudes only for processes of sufficiently long timescale, such as the QBO. The thermal wind balance is maintained by an induced meridional circulation with adiabatic heating (cooling) in the cold (warm) region driving upward (downward) motion (Reed, 1964). The direction of the induced circulation depends on the phase of the equatorial wind QBO. During westerly (easterly) phase, descending (ascending) motion is present at the equator just below the level of maximum vertical wind shear, with rising motion at midlatitudes. The induced meridional circulation does not extend beyond 30° from the equator (Gray and Pyle, 1989).

The QBO in ozone is a product of the interaction of the dynamics discussed above and photochemistry. Gray and Pyle (1989) have proposed the following

mechanism to account for the ozone QBO: The QBO in equatorial zonal winds produces a rising and sinking motion at the equator as described above. The induced meridional circulation may be expected to modulate the strength of both the vertical motion and the horizontal divergence of the Hadley circulation in the lower stratosphere. During the westerly phase of the QBO the strength of the equatorial ascent in the Hadley cell is reduced and during the easterly phase it is increased. This change in the vertical velocity of the Hadley cell circulation modulates the vertical transport of ozone poor air from the lower stratosphere and troposphere into the middle stratosphere. During the westerly phase QBO the equatorial total column ozone should be above the long-term mean. The return arm of the QBO induced circulation in the subtropics is opposite in direction to that at the equator. This causes a reversed ozone anomaly in the subtropics. In Figures 5.19 and 5.20 it was shown that there were regions near 10° to 15°N and 10°S where the effect of the QBO fell to close to zero. These areas fall between the two zones of vertical transport in the Hadley cell circulation creating a zero vertical transport regime. It is clear from the mechanism described above that if the Hadley cells are centered at 10°N and 10°S there should be no QBO influence there and QBO amplitudes obtained from the statistical model should rise rapidly both poleward and equatorward of these regions. This mechanism also accounts for the meridional QBO phase shift observed in Figures 5.22 and 5.22 as discussed in section 5.5.14. Gray and Pyle (1989) reported that the amplitude of the QBO wind is largest at the equator and drops off rapidly with an e-folding width of approximately 15°.

The mechanism outlined above explains the QBO at subtropical latitudes but it remains to be shown how this influence extends to higher latitudes. It has been demonstrated that high latitude zonal wind velocities are closely related to the phase of the equatorial QBO (Holton and Tan, 1980, 1982; Labitzke, 1982). Gray and Pyle (1989) made use of a model as an aid to understanding the mechanism which causes the QBO in total column ozone at high latitudes. The model produced QBO influences poleward of 30° latitude (the limit predicted by linear theory) and suggested that a combination of the mean circulation and eddy motions extended this influence to higher latitudes.

However, there is some uncertainty as to whether the combination of mean and eddy circulation is responsible for the high latitude QBO in ozone. Tung and Yang (1988) stated that it has not been firmly established that the high latitude QBO

signal in temperature and ozone is caused by the equatorial QBO in the zonal winds. They also commented that there is no definitive measure for the phase of the equatorial QBO. The 50 hPa wind direction used by Garcia and Solomon (1987) may be an imperfect measure because during some years the zonal winds in the equatorial lower stratosphere may have different signs at different levels.

### 5.5.14 The phase change in the QBO signal near 15°N and 15°S

Figures 5.21 and 5.22 show a marked change in QBO phase at approximately 15°N and 15°S. Similar signatures have been observed in previous studies. Figure 5.38 shows the total column ozone anomaly obtained from the model of Gray and Pyle (1989). The minimum in QBO amplitude and the change in QBO phase across the latitude lines at 15°N and 15°S is clearly evident. In the discussion above it was explained that the equatorial and low latitude QBO is produced through modulation of vertical velocities by a modified Hadley cell circulation. During the westerly phase of the equatorial wind QBO, descending motion at the equator results in an increase in column ozone while the ascending motion of the low latitude leg of the meridional circulation cell results in a decrease in total column ozone. It is clear that if the circulation cell is centered at 15° latitude, there will be a marked change in phase across this transition region. The phase reversal in the temperature QBO at about 15° has also been previously quoted (Angell and Korshover, 1962, 1964, 1967).

### 5.5.15 Variability in the period of the QBO

It is necessary to account for the large variability observed in QBO periods, both in this model and due to real physical processes. Gray and Pyle (1989) reported QBO periods varying between 22 and 34 months with an average of 28 months. It was also shown that QBO periods are generally integral multiples of the semi-annual oscillation (SAO) in the equatorial winds above 10 hPa. The suggestion that the period of the SAO may influence the period of the QBO was first made by Lindzen and Holton (1968). Gray and Pyle (1989) indicated that this influence of the SAO on the period of the modelled QBO may provide understanding of the large variation in the period of the QBO in the real atmosphere. While such

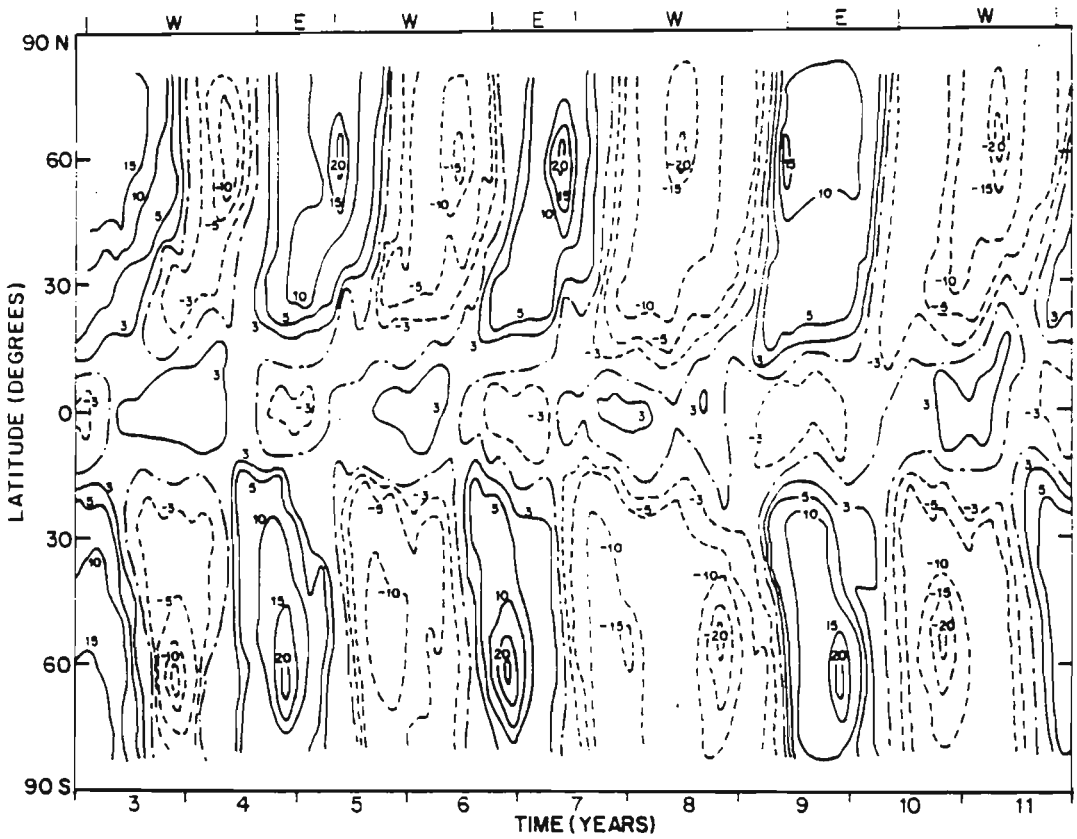


Figure 5.38: Latitude-time section of the modelled column ozone anomaly (DU) obtained from a model run from Gray and Pyle (1989). The direction of the equatorial wind at approximately 50 hPa is indicated at the top of the figure.

large variability may be explained in terms of the year-to-year variability in the amplitude of the wave forcing that gives rise to the QBO, it may be more easily understood in terms of a variability around a number of semi-annual quantized periods.

### 5.5.16 Effects of the solar cycle on the global total column ozone distribution

Although the variation in the total solar radiation incident at the top of the Earth's atmosphere is small over the period of a solar cycle (Brusa and Frolich, 1982), the variability increases with decreasing wavelength. Since it is the shorter wavelengths ($\lambda < 310\,\text{nm}$) that control the temperature and photochemistry of

the middle atmosphere, upper stratospheric and mesospheric processes may be significantly influenced by the solar cycle. Furthermore, since it is the UV portion of the solar spectrum that is responsible for the creation of ozone, it would be expected that variations in the solar UV component could cause variations in total column ozone levels. Blackshear and Tolson (1978) estimated total column ozone derived from Nimbus-4 IRIS data and found that it had characteristics in common with variations in the 10.7 cm solar flux, the Zurich sunspot number, and the solar lyman $\alpha$ flux. It was shown that the highest correlation ($r^2 = 0.94$) existed between Lyman $\alpha$ flux and total ozone.

The basis for a solar cycle influence on total column ozone is the notion that reactive nitrogen compounds ($NO_x$) can destroy ozone quite effectively under certain photochemical conditions. Recall from Chapter 1 that it is also $NO_x$ that participates in the chlorine interference reaction and can indirectly inhibit ozone depletion under different conditions. The factors contributing to a solar cycle influence on total column ozone are summarized in Figure 5.39. In the middle

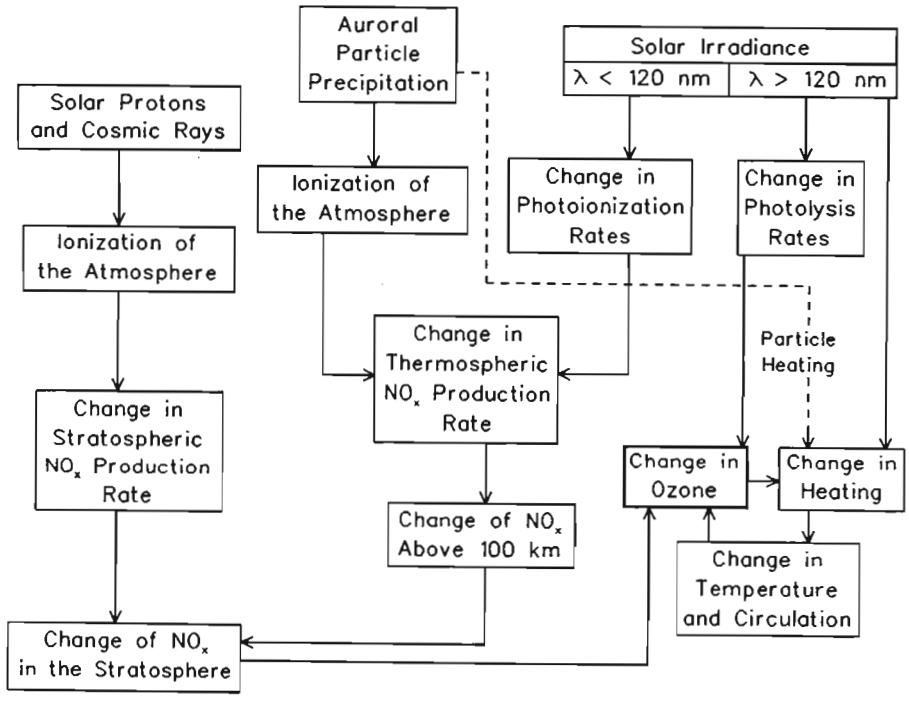


Figure 5.39: Potential responses on the middle atmosphere to variability in solar activity over the 11 year solar cycle. Adapted from Brasseur and Solomon (1986).

atmosphere, the solar variability has two primary effects:

- It can produce a change in the atmospheric heating rate by molecular oxygen and ozone. The resulting temperature variation is small for altitudes below 90 km, where the major heating occurs through ozone absorption in the Hartley band (near 250 nm). However since the solar variability is small at these wavelengths, the solar cycle induced temperature variation below 90 km is small. Note the temperature variation feedback in Figure 5.39 which has been previously discussed in Chapter 4. A decrease in middle atmosphere temperatures during a solar minimum may result in enhanced wintertime Antarctic stratospheric cooling, a more isolated vortex and greater ozone depletion. In turn the reduced heating through UV absorption by ozone lowers the temperature still further.

- Solar variability can modify the rates of photodissociation and photoionization of trace species, particularly in the upper atmosphere where shortwave radiation has not been absorbed. Since radiation at these wavelengths shows greatest variability with the solar cycle, it may be expected that variation in these rate constants will be large. A further feedback mechanism is involved here since ozone in the high stratosphere absorbs the short wavelength radiation which drives the photoionization and photodissociation.

The maximum in the 11 year solar cycle is accompanied by increased injection of energetic particles, causing aurora and solar proton events. These have been shown to have a significant impact on Antarctic ozone depletion (Stephenson and Scourfield, 1991). In addition, the events increase atmospheric ionization and cause Joule heating. The precipitation of high energy solar protons over the polar caps produce secondary electrons which ionize and dissociate molecular nitrogen leading to the formation of nitric oxide. The nitric oxide can then play an important role in ozone depletion through the following catalytic reactions:

$$NO + O_3 \rightarrow NO_2 + O_2 \tag{5.5}$$

$$NO_2 + O \rightarrow NO + O_2 \tag{5.6}$$

These processes can influence the chemistry of the thermosphere and possibly lower levels through long range transport.

### 5.5.17 Seasonal changes in solar irradiance

Seasonal changes in solar irradiance may indirectly influence the global total column ozone distribution through changes in atmospheric dynamics. As a result of the ellipticity of the Earth's orbit, it is closer to the sun during the Southern Hemisphere summer and further from the sun during the Southern Hemisphere winter. This effect can result in a 7% difference in solar radiation between hemispheres. This may also partially account for colder Antarctic winter temperatures compared to those of the Arctic. If Antarctic summer temperatures are higher than those of the Arctic and winter temperatures are lower, winter radiative cooling above the Antarctic must be more rapid than over the Arctic. Meridional temperatures gradients should be greater in the Southern Hemisphere creating a stronger geostrophic wind and a more isolated vortex. As discussed in detail in previous chapters, this has important consequences for the global total column ozone distribution.

### 5.5.18 Forcing mechanisms for the semi-annual oscillation

The middle and low latitude semi-annual oscillation observed in Figure 5.29 probably results from the meridional movement of the Intertropical Convergence Zone (ITCZ). The position of the ITCZ varies seasonally in approximate accordance with the location of maximum solar heating (Henderson-Sellers and Robinson, 1987). In equatorial and tropical regions, this could result in a 6 month period cycle. A zenith noon sun position, corresponding to maximum heating, occurs twice every year over the equator. As a result of the close correlation between solar irradiance and total column ozone discussed above, it can be expected that maxima in tropical total column ozone may also follow the meridional movement of maximum solar irradiance. A semi-annual variation in tropical total column ozone may influence subtropical regions as ozone rich air is transported poleward.

It has also been shown (Perliski et al., 1989) that the equatorial semi-annual oscillation in ozone above 10 hPa is caused mainly by the semi-annual temperature and wind oscillation (SAO). It was also shown that below 10 hPa the modelled equatorial ozone variation is caused by increased rates of odd-oxygen production associated with the semi-annual zenith crossings of the sun. The forcing mecha-

nisms for the semi-annual signal observed over the Antarctic is discussed further in section 5.5.19.

### 5.5.19 Large semi-annual oscillation amplitudes observed over the Antarctic

The large amplitudes of the semi-annual oscillation observed over the Antarctic have three possible sources:

- High latitude Southern Hemisphere total column ozone experiences a minimum during March and April and without the effects of chlorine catalyzed ozone depletion would experience a maximum during November and December (see Figure 5.36). However, the wintertime ozone depletion creates a chemically induced minimum during October and November. The occurrence of two minima each year produces a strong semi-annual signal confined to the Antarctic regions where chemical depletion takes place.

- Interpolation over the polar cap results in anomalously high midwinter total column. Since the polar night terminator, defining the regions where no TOMS measurements are possible, moves equatorward during the winter, interpolation across the polar cap may result in an increase in total ozone since values increase in the equatorward direction. This enhanced midwinter maximum, a function of the data processing and not a real physical effect, could further contribute to the observed semi-annual signal. Note however, that interpolation across the Arctic polar cap does not produce a similar effect.

- Perliski et al. (1989) showed that modelled polar semi-annual ozone oscillations are caused by modulation of radiatively driven middle stratospheric ozone variation by temperature dependent chemical destruction processes.

## 5.6 Comparison with measured data

The global coefficients derived from the model have been used to generate monthly mean total column ozone distributions for each month from January 1979 to December 1992. Measured monthly averages have then been subtracted from each

of these distributions to create 168 (12 months $\times$ 14 years) difference data sets. Due to space considerations, not all of these data sets can be shown. However, summaries of their general morphology are provided by means of plots of their zonal and meridional means. The zonal means are plotted in Figure 5.40 as a function of latitude and time using a colour code shown on the right of the figure. Hues of green, cyan, red and purple indicate values within 19 DU of that predicted

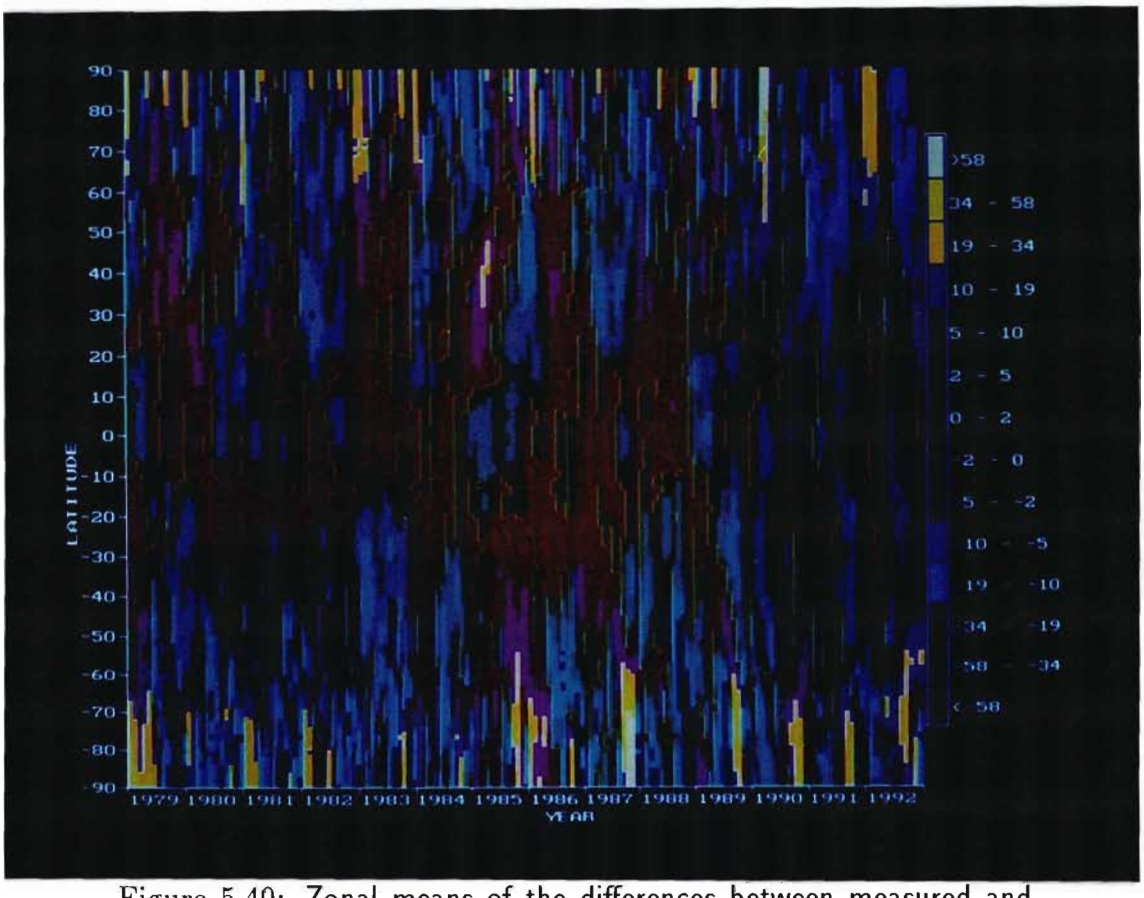


Figure 5.40: Zonal means of the differences between measured and modelled monthly average total column ozone. January 1979 to December 1992.

by the model, while hues of blue and yellow show zonal means differing by more than 19 DU from the model. Since these values were calculated by subtracting measured data from model derived values, hues of green and blue indicate under-prediction by the model (ie. abnormally high total column ozone) while hues of red and yellow correspond to over-prediction by the model (ie. unusually low total column ozone). It is clear from Figure 5.40 that between 60°S and 60°N zonal mean differences seldom exceed 19 DU. Total column ozone levels over the Arctic appear to be lower than normal during the winter of 1992 (yellow region in

Figure 5.40) as reported in other studies (Gleason et al., 1993). The high latitude Southern Hemisphere behavior in Figure 5.40 confirms the severe Antarctic ozone depletion of 1987 and the weaker depletion of 1988 as observed in Chapter 4. This suggests that the planetary wave activity indices derived in the previous chapter may be used to make a correction to the model at high latitudes to provide more accurate estimates of total column ozone.

Meridional means of each of the difference files have also been calculated and are plotted as a function of time and longitude in Figure 5.41. Because most of

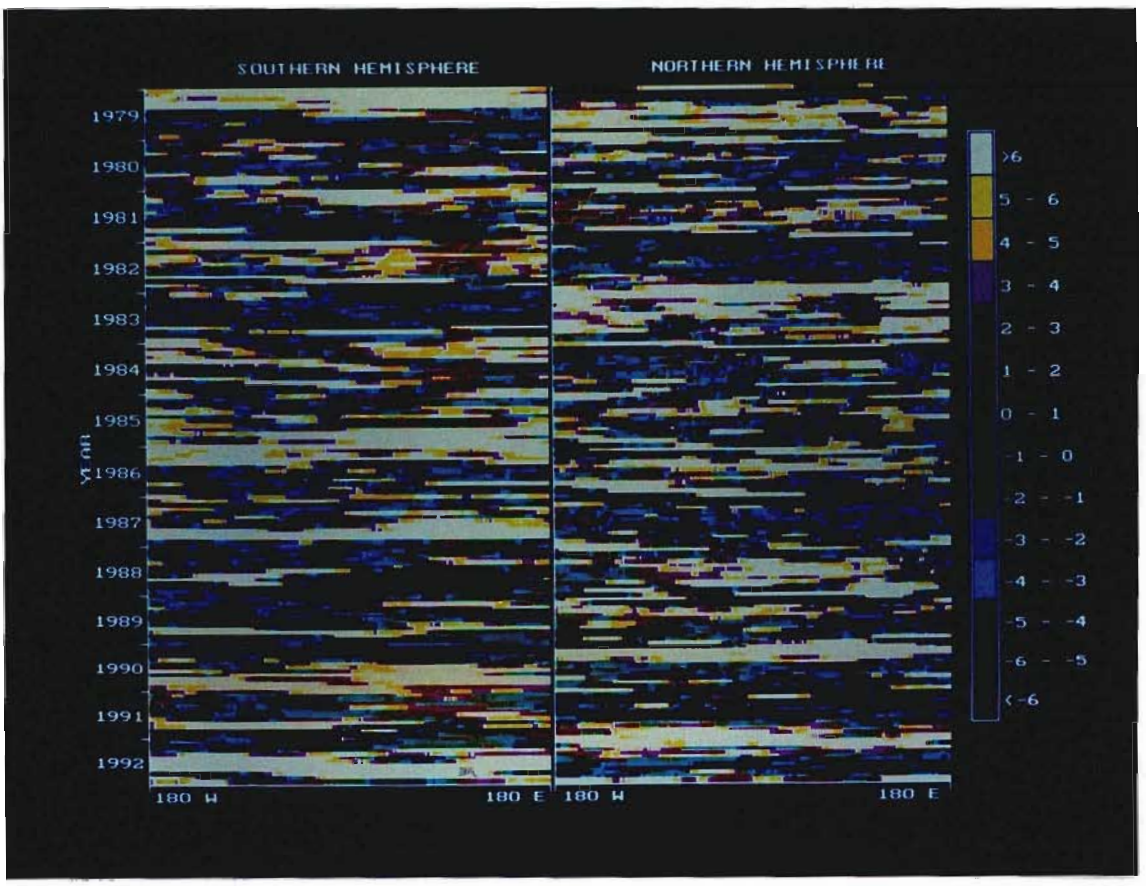


Figure 5.41: Meridional means of the differences between measured and modelled monthly average total column ozone, January 1979 to December 1992.

the variation in the difference files is in the meridional plane, meridional means show little variation and mask the underlying structure. It is also for this reason that the colour scale used is of higher resolution than that of Figure 5.40. The anomalously high winter Antarctic ozone levels for 1988 again appear as a dark blue area which appears to occur preferentially in the Eastern Hemisphere. It

was reported in a previous study (Schoeberl et al., 1989) that in 1988 the Southern Hemisphere total column ozone distribution showed a zonally asymmetric structure with the minimum of the crescent shaped midlatitude ridge near 90°W. This large planetary wave 1 perturbation, persisting throughout the winter, produces a maximum in total column ozone near 90°E as observed in Figure 5.41. Both Figures 5.40 and 5.41 again show that differences between measured and model data maximize during the winter months. The interannual variability in the duration and strength of the circumpolar vortex can result in large difference in polar total column ozone levels. These vortex properties are in turn strongly modulated by planetary wave activity as discussed in greater detail below.

## 5.7 The influence of planetary waves on long-term trends and model accuracy

To investigate the possibility that planetary wave powers derived in Chapter 4 may be used to make corrections to the zonal mean differences observed in Figure 5.40, meridional means of the zonal mean differences have been calculated between 60° and 90°. This provides two data series (one for each hemisphere) of 168 data points that characterize the differences between measured and modelled data at high latitudes. In addition, monthly averages of the midlatitude total planetary wave powers (see Chapter 4) have been calculated for both hemispheres. The appropriate monthly mean for each calendar month, averaged over the whole data series, has been subtracted from each data point in order to remove the annual cycle. Months of unusually high wave power then appear as large positive excursions from zero. The monthly mean total column ozone difference data series and the deviation signal of the total planetary wave powers for the Southern Hemisphere are plotted in Figure 5.42. The 60°S to 90°S meridional means from Figure 5.40 are plotted against the left ordinate (solid line) while the total wave power deviation from the annual mean is plotted against the right ordinate using an inverted scale (dashed line). Starting in 1985 wave power and ozone deviation events appear to be correlated in the winter of each year. During winters of reduced (enhanced) wave power, ozone deviations appear to be large and positive (negative) ie. the model over predicts (under predicts) total column ozone, indicating a year of enhanced (reduced) ozone depletion. Attention is drawn to the fact that the solid curve in Figure 5.42 is not a measure of the

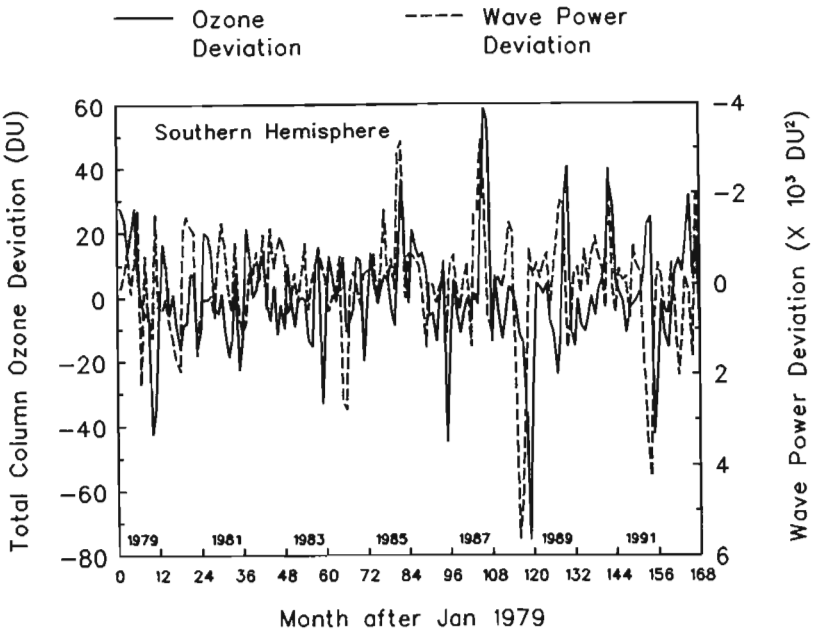


Figure 5.42: The total column ozone deviation signal obtained by calculating the meridional means from 60°S to 90°S in Figure 5.40 (solid line) and the deviation of Southern Hemisphere total planetary wave powers from the annual mean (dashed line) from January 1979 to December 1992.

depth of the Antarctic ozone hole. It indicates only deviations between actual depletion and model predicted depletions. Note also that there appears to be a time lag in many of the events where the deviation of the wave power precedes the total column ozone deviation by a few months. These events are discussed further below.

A similar plot for Northern Hemisphere total column ozone deviations and total wave power deviations is shown in Figure 5.43. However in this case, the signals appear to be far more variable and no clear correlation between the two signals exists.

## 5.8 Discussion

In Chapter 4 is was shown that during years of anomalously high (low) total wave power, ozone depletion over the Antarctic is considerably reduced (enhanced). This has been confirmed in Figure 5.42 where it seen that during years of high (low) wave power a statistical model of total column ozone variation tends to un-

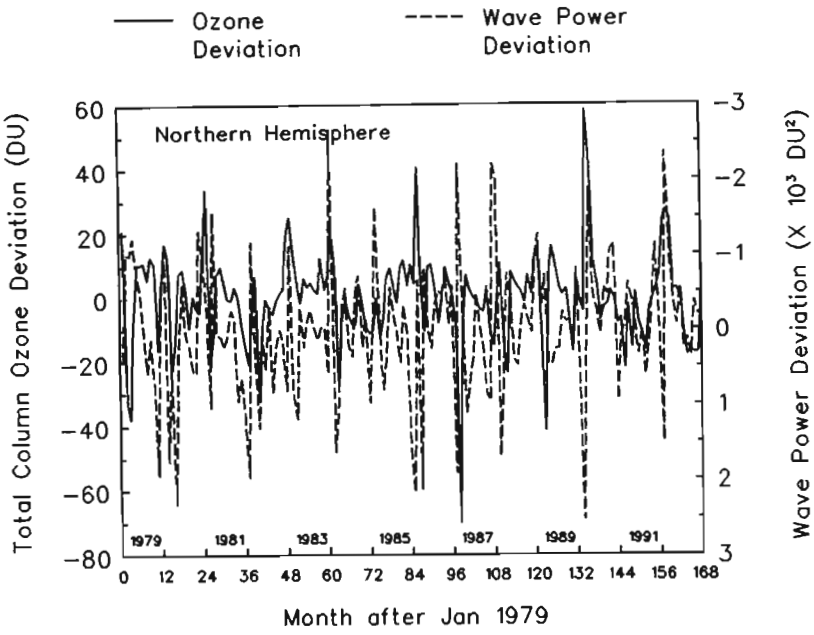


Figure 5.43: The total column ozone deviation signal obtained by calculating the meridional means from 60°N to 90°N in Figure 5.40 (solid line) and the deviation of Northern Hemisphere total planetary wave powers from the annual mean (dashed line) from January 1979 to December 1992.

der predict (over predict) Southern Hemisphere high latitude ozone levels. The characteristics of the events shown in Figure 5.42 from 1985 to 1992 are summarized in Table 5.6 where the total column ozone deviation (in DU), the wave power deviation (in $10^3$ $(DU)^2$) and the time lag between the two events (in months) are listed. Also listed are the months when the local maxima/minima were observed in each signal. The total column ozone deviation has been plotted as a function of total wave power deviation in Figure 5.44, ignoring the phase shifts listed in Table 5.6. The straight line fit to the data points ($r^2 = 0.87$) gives the following relationship between total column ozone deviation and total planetary wave power deviation:

$$\text{Total Ozone Deviation} = (-13.86 \times \text{Wave Power Deviation}) + 5.11 \qquad (5.7)$$

where the total column ozone deviation is measured in DU and the total wave power deviation is measured in $10^3$ $(DU)^2$. This relationship indicates that it should be possible to make use of the planetary wave analysis of Chapter 4 to make corrections to the high latitude behavior of the statistical model. Furthermore, it was shown in Chapter 4 that there is a direct correlation between planetary wave structures observed in the total column ozone distribution and wave structures

| Year | Ozone deviation | | Wave power deviation | | Time Lag |
|---|---|---|---|---|---|
| | Amp. (DU) | Month | Amp. ($10^3$ $(DU)^2$) | Month | (months) |
| 1985 | 36.402 | Oct | -3.165336 | Oct | 0 |
| 1986 | -44.718 | Nov | 1.077648 | Nov | 0 |
| 1987 | 58.627 | Oct | -3.194604 | Sep | 1 |
| 1988 | -74.976 | Nov | 5.631593 | Aug | 3 |
| 1989 | 40.48 | Oct | -1.842954 | Aug | 2 |
| 1990 | 35.802 | Oct | -2.550136 | Oct | 0 |
| 1991 | -42.387 | Nov | 4.262314 | Oct | 1 |
| 1992 | 31.239 | Sep | -0.191283 | Aug | 1 |

Table 5.6: Characteristics of total column ozone deviation and total wave power deviation events observed in Figure 5.42.

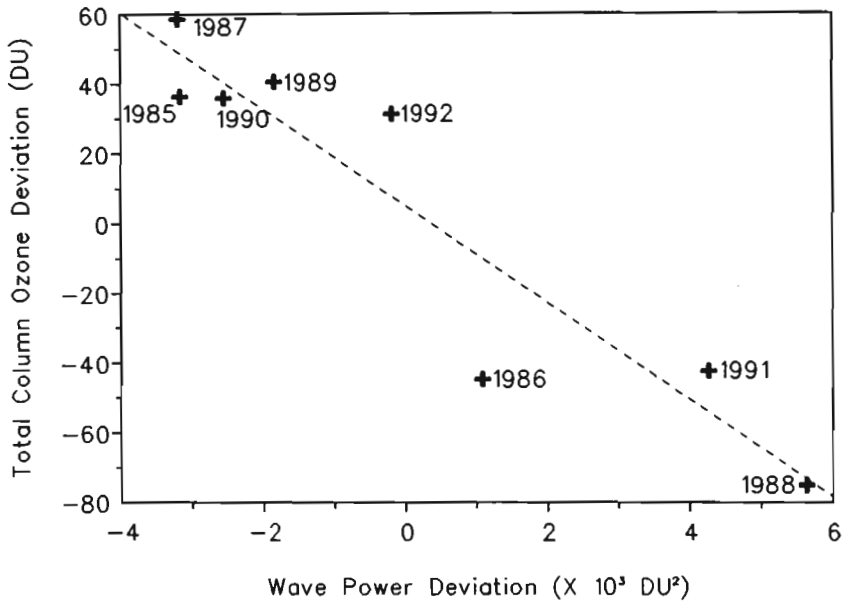


Figure 5.44: Total column ozone deviation (DU) as a function of total wave power deviation ($10^3$ $(DU)^2$). The phase shifts listed in Table 5.6 are ignored.

observed in the NMC geopotential height fields. The time lag observed in Figure 5.42 suggests that realtime calculations of planetary wave powers using NMC or ECMWF data may allow a 1 or 2 month forward prediction of the severity of the Antarctic ozone depletion for that year. Further studies are needed however to test the viability of this approach.

# Chapter 6

# APPLICATION OF THE STATISTICAL MODEL: SURFACE UVB RADIATION OVER SOUTH AFRICA

## 6.1 Introduction

The statistical model developed in the previous chapter has been used to study long-term trends in total column ozone above 5 South African cities. Interpolation between TOMS cells (as performed by the NASA/GSFC extraction software) was used to obtain daily TOMS data for each location from 1979 to 1992. Monthly averages of these data were used as inputs to the model to obtain the 14 coefficients (see section 5.1) characterizing the long-term trend and cyclical variations in the monthly averages. Annual mean total column ozone above each city was also calculated to provide further insight into long-term trends.

Since stratospheric ozone is the primary absorber of incident ultraviolet (UV) radiation, decreases in total column ozone will produce increases in the Earth's surface UV irradiance. To investigate this relationship, a surface irradiance model has been developed based on that of McKenzie (1991). The model output consists of a surface irradiance spectrum which, when multiplied by a human skin response spectrum (Diffey, 1991) and integrated from 280 to 340 nm, gives an erythemal

dose in $\mu$W cm$^{-2}$. Statistical model coefficients have been used to predict monthly average total column ozone levels to the year 2000. These data have then been used as inputs to the surface irradiance model to obtain an estimate of future UV levels above these 5 cities in South Africa.

## 6.2 The surface irradiance model

A simple clear-sky irradiance model has been developed to map total column ozone to surface erythemal irradiance. A complete description of this model, together with source code written to implement it, is given in Appendix B while a brief description is provided below. The model assumes a constant extraterrestrial solar irradiance with seasonal variations induced by changes in noon solar zenith angle. Given the extraterrestrial solar irradiance, the ozone absorption cross-section, total column ozone, location and day number, the model approximates surface irradiance between 280 and 340 nm, the UVB portion of the spectrum having greatest impact on biological systems. The model does not account for changes in surface pressure (altitude) or cloud cover.

The total irradiance at ground level can be divided into direct and diffuse components. The direct component is described by:

$$\text{Direct} = \varepsilon E_0 T_r T_a T_m \cos\chi \tag{6.1}$$

where $\varepsilon$ is a factor accounting for the ellipticity of the Earth's orbit, $E_0$ is the extraterrestrial solar irradiance, $T_r$ is the Rayleigh transmission factor and $T_a$ is the aerosol transmittance where the Ångström turbidity coefficient (Iqbal, 1983) has been taken to be 0.05 and the wavelength exponent for aerosol scatter is 1.3 (Iqbal, 1983). The molecular transmission $T_m$ was calculated using the Bouguer-Lambert law, assuming the ozone to be concentrated in a thin layer at 20 km. $\chi$ is the solar zenith angle (SZA). The diffuse component is given by:

$$\text{Diffuse} = \frac{(E_{dr} + E_{da} + \text{Direct}\, r_g\, r_a)}{1 - r_g\, r_a} \tag{6.2}$$

where the ground reflectance ($r_g$), and the effective sky reflectance ($r_a$), are both taken to be 0.1. $E_{dr}$ is the Rayleigh-scattered diffuse radiation and $E_{da}$ is the aerosol-scattered diffuse irradiance.

## 6.3 Observations

### 6.3.1 Daily total column ozone

TOMS total column ozone data have been obtained for Pretoria, Bloemfontein, Durban, Cape Town and Port Elizabeth for each day from 1979 to 1992. Geographic coordinates for each city are listed in Table 6.1 in order of increasing latitude.

| Location | Latitude | Longitude |
|---|---|---|
| Pretoria | 25.73°S | 28.20°E |
| Bloemfontein | 29.10°S | 26.23°E |
| Durban | 29.87°S | 30.98°E |
| Cape Town | 33.92°S | 18.37°E |
| Port Elizabeth | 33.97°S | 25.67°E |

Table 6.1: Geographical coordinates for the 5 cities examined in this study.

The temporal variation in the daily total ozone above each city is summarized in Figures 6.1 a to 6.1 e where total column ozone is plotted against day number through the year. The shaded region in each figure indicates the extreme values for each calendar day within the period 1979 to 1991, and a thick solid line shows the average for each calendar day over the same period. Data for 1992, the most recent complete year of Nimbus-7 TOMS data, have been plotted separately (dashed line) to show how this year compares with the baseline (bold line). Figures 6.1 a to 6.1 e reveal that total ozone levels in 1992 were generally lower than the previous 13 year mean for all 5 cities and on occasion fell lower than the minimum over the 1979 to 1991 period. Note also that the total column ozone levels generally increase with latitude. The 1979 to 1991 mean at Pretoria (thick solid line in Figure 6.1 a), the most northerly location, ranges from 260 to 305 DU while that for Port Elizabeth (thick solid line in Figure 6.1 e), the most southerly location, varies from 270 to 330 DU. The interannual variability in total column ozone above each city has been quantified by obtaining an annual mean of the differences between the two curves bounding the shaded regions in each of Figures 6.1 a to 6.1 e. The results listed in Table 6.2 indicate that the interannual variability generally increases with latitude.

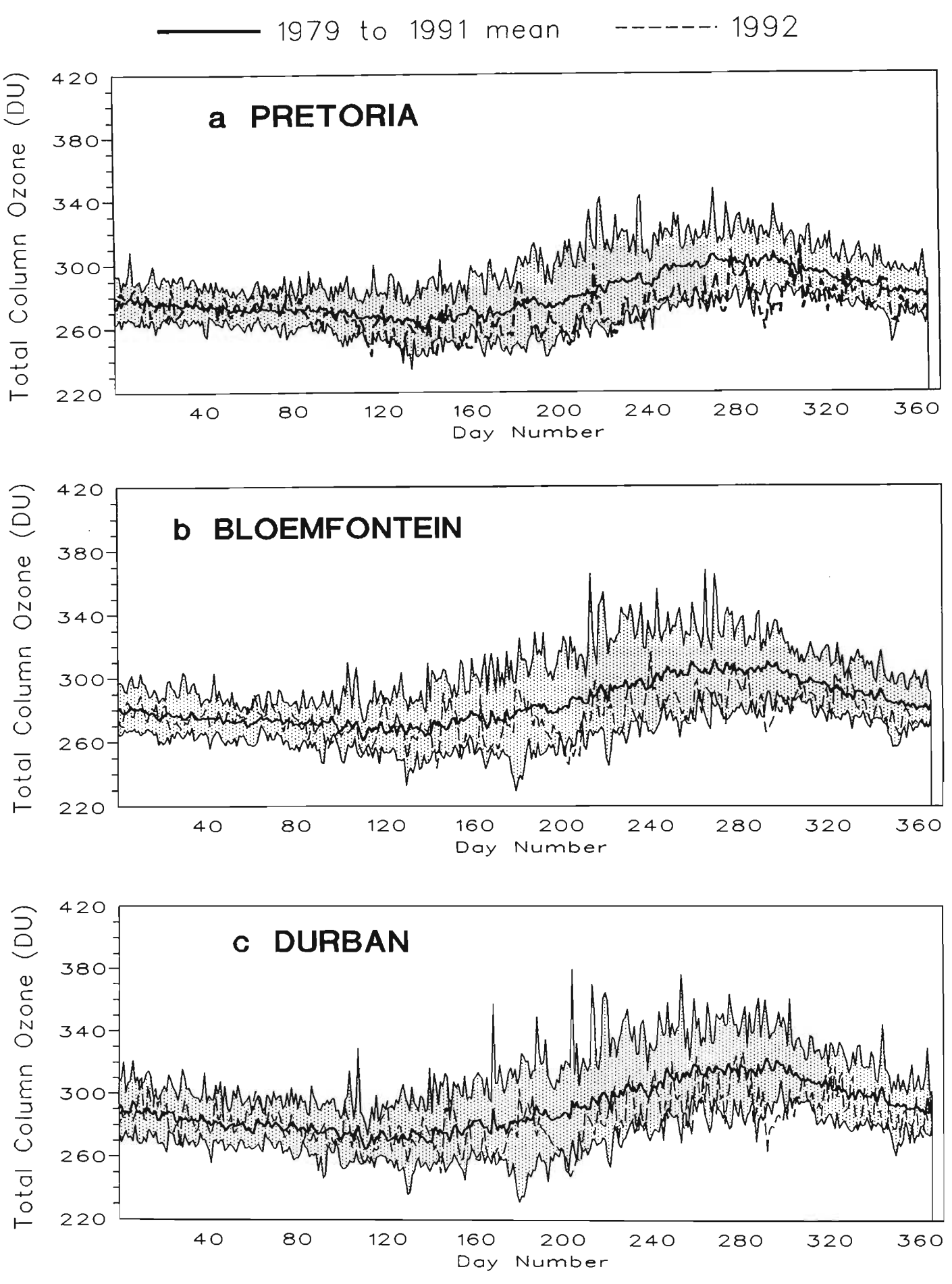


Figure 6.1: Interannual variation in daily total column ozone above a) Pretoria, b) Bloemfontein and c) Durban. The shaded areas show the range of daily values from 1979 to 1991 while the thick solid lines show the mean over the same period. The dashed lines show the 1992 daily total column ozone separately for comparison with previous years.

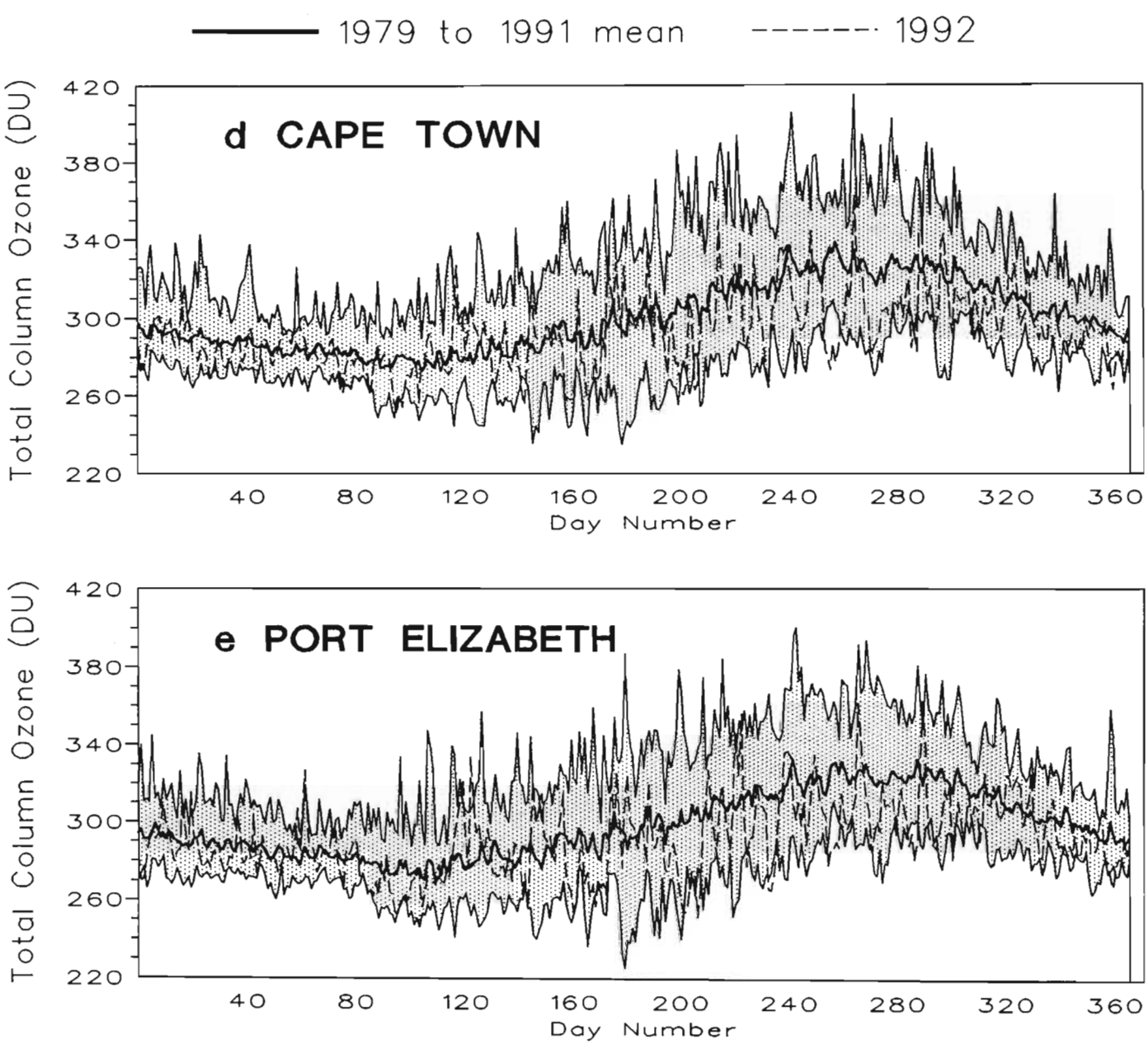


Figure 6.1 cont.: Interannual variation in daily total column ozone above d) Cape Town and e) Port Elizabeth. The shaded areas show the range of daily values from 1979 to 1991 while the thick solid lines show the mean over the same period. The dashed lines show the 1992 daily total column ozone separately for comparison with previous years.

| Location | Mean of Max - Min (DU) |
|---|---|
| Pretoria | 35.39 |
| Bloemfontein | 41.62 |
| Durban | 45.43 |
| Cape Town | 58.80 |
| Port Elizabeth | 57.50 |

Table 6.2: **Annual mean of differences between the maximum and the minimum of the calendar day averaged total column ozone from 1979 to 1991 for each of the 5 cities.**

For Port Elizabeth, the interannual difference in total column ozone near day 180 (29 June) exceeds 160 DU.

### 6.3.2 Monthly average total column ozone

Monthly averages of total column ozone above the 5 cities have been calculated from 1979 to 1992. The 14 values for each month (one value per year) have then been averaged to produce a representative seasonal profile for each location. Figures 6.2 a to 6.2 e show these profiles in order of increasing latitude for Pretoria, Bloemfontein, Durban, Cape Town and Port Elizabeth respectively. Maximum and minimum monthly averages for each month during the 14 year period are also shown. By ordering the cities in increasing latitude, it is clear from Figure 6.2 that the deviation from the 14 year mean, as shown by the maximum and minimum curves, generally increases with latitude.

The 14 year mean curves of Figures 6.2 show similar behavior with a decline during the first 4 or 5 months of the year, followed by a steady increase in total column ozone until a maximum is reached in September or October. During the last 2 or 3 months of the year total ozone falls to levels observed at the beginning of the year. The behavior of the 14 year mean for each of these cities is summarized in Table 6.3. Column one lists the cities in order of increasing latitude, the second and third columns show, respectively, the maximum and minimum DU values of the 14 year mean curve, the fourth column lists the percentage variation, (max-min)/max $\times 100$. The last two columns list the maximum variance and minimum

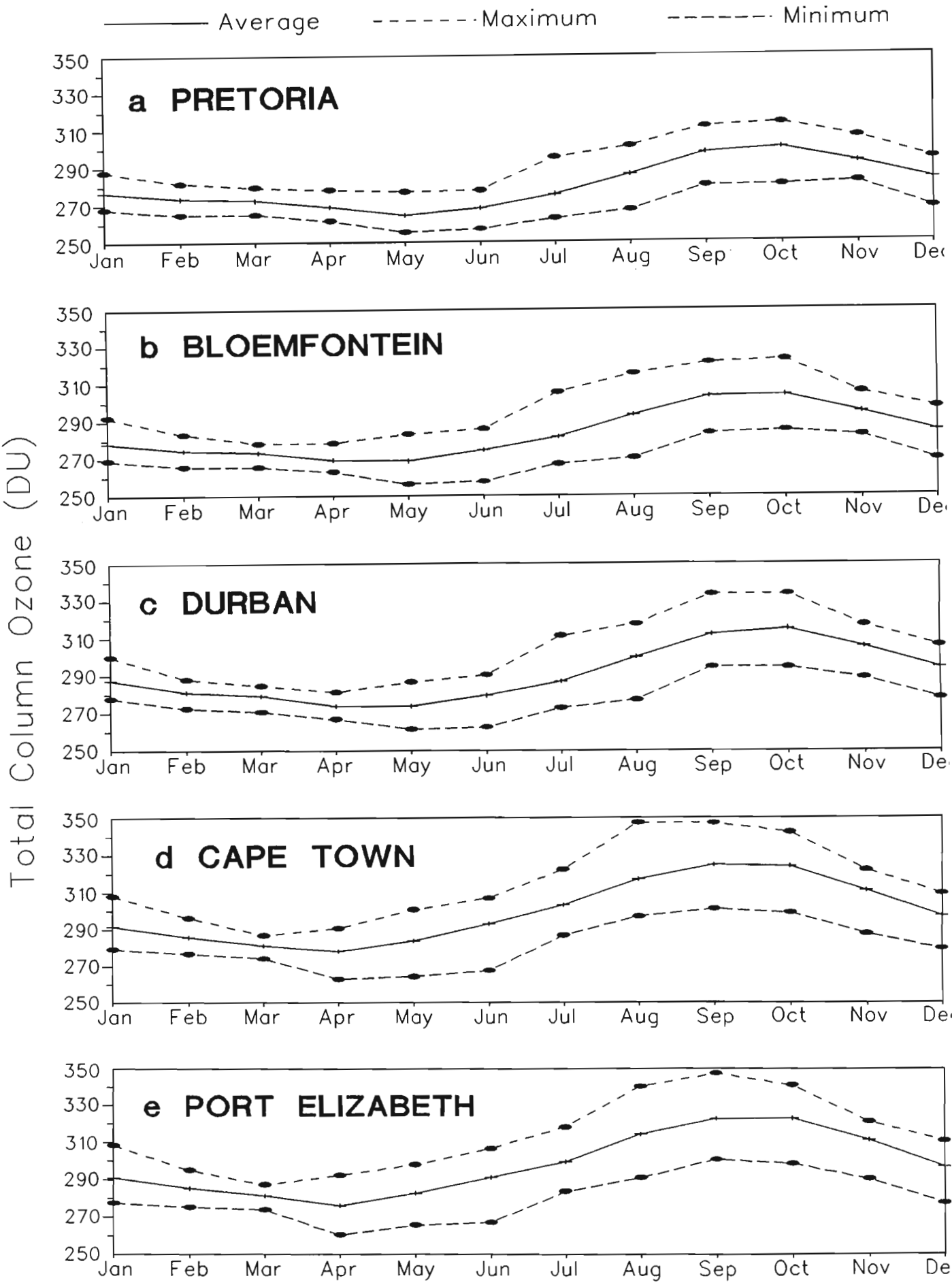


Figure 6.2.: The 1979 to 1992 average of the monthly mean total column ozone above a) Pretoria, b) Bloemfontein, c) Durban, d) Cape Town and e) Port Elizabeth. The solid lines show the mean, the short dashed lines show the minimum, and the long dashed lines shown the minimum experienced for each month through the period.

| Location | Max | Min | % variation | Max variance | Min variance |
|---|---|---|---|---|---|
| Pretoria | 299.70 | 264.56 | 11.7 | 156.38 | 158.12 |
| Bloemfontein | 303.66 | 268.61 | 11.5 | 229.39 | 208.71 |
| Durban | 314.98 | 273.54 | 13.2 | 216.03 | 215.97 |
| Cape Town | 324.64 | 278.42 | 14.2 | 286.57 | 337.02 |
| Port Elizabeth | 322.72 | 276.14 | 14.4 | 287.97 | 319.66 |

Table 6.3: Statistical data for curves shown in Figures 6.2 a to 6.2 e. Max and Min are in DU while the variances are in $DU^2$.

variance in $(DU)^2$. These variances have been determined using:

$$\mathrm{Var}(x_1 \cdots x_{12}) = \frac{1}{12}\sum_{i=1}^{12}(x_i - \overline{x})^2 \tag{6.3}$$

where the $x_i$ are either the maximum or minimum monthly means and the $\overline{x}$ are the values from the 14 year mean curve. From Table 6.3 it is clear that the annual variation, the variance of the maximum monthly means, and the variance of the minimum monthly means, generally increase with latitude although there are some exceptions. These may result from the 12° spread in longitude between these cities.

The monthly average analyses have been summarized in Figure 6.3 where the 14 year monthly mean total column ozone curves are shown for each of the 5 cities. Generally the amplitudes of these curves increase with latitude except for Cape Town and Port Elizabeth where Cape Town experiences larger monthly average total column ozone. Although there is a similarity in the shape of these profiles, Cape Town and Port Elizabeth (the more southerly locations) experience their annual minimum in April while the other locations experience their annual minimum in May. In addition, Cape Town and Port Elizabeth show a maximum in September while the other three cities show a maximum in October. In the previous chapter it was shown that between 20°S and 40°S the phase of the annual variation (after being subtracted from 9) decreases with latitude (see Figure 5.15) confirming these observations. These phase differences have important implications for the annual variability in surface UVB irradiance as discussed in section 5.5.12. Locations closer to the equator are expected to experience greater surface UVB irradiance than more southerly locations since the solar irradiance is incident at smaller solar zenith angles. However, since northern locations experience their minimum in total column ozone closer to the winter, the annual variation

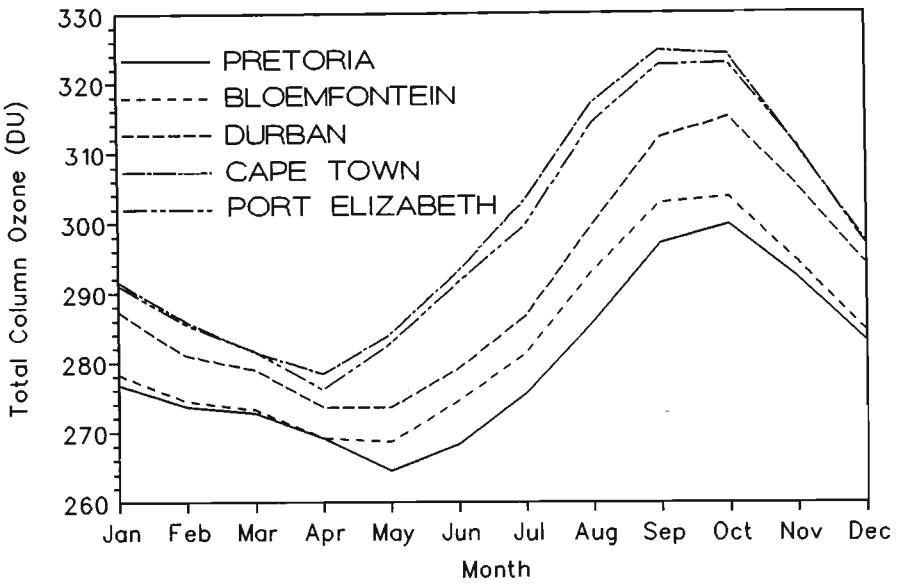


Figure 6.3: The 1979 to 1992 average of the monthly mean total column ozone for each of the 5 sites.

in surface UVB irradiance for all 5 cities could be similar. The phase differences obtained from the statistical models for each location are presented later.

Regression analysis of the 14 monthly mean total column ozone values for each month and for each city indicates that the maximum long-term decrease occurs during the month of May while during October the long-term trend for each city is statistically insignificant. This has important implications for long-term trends in surface UVB irradiance since during the month of May (winter) solar zenith angles tend to be large with the result that surface irradiances are low. Declines in total column ozone during this month will therefore have little effect on the annual maximum UVB trend. Long-term declines during the summer months would have far more serious consequences for surface UVB irradiances.

### 6.3.3 Yearly average total column ozone

Yearly averages of total column ozone have been calculated for each of the 5 cities from 1979 to 1992 as shown in Figures 6.4 a to 6.4 e. When calculating yearly averages, missing data points resulting from the NASA data validation process (see Chapter 3) have been interpolated to prevent preferential biasing in certain periods of the year. These figures also show the linear trends (dashed lines) determined from the regression analyses. The trends, in DU/year and as a

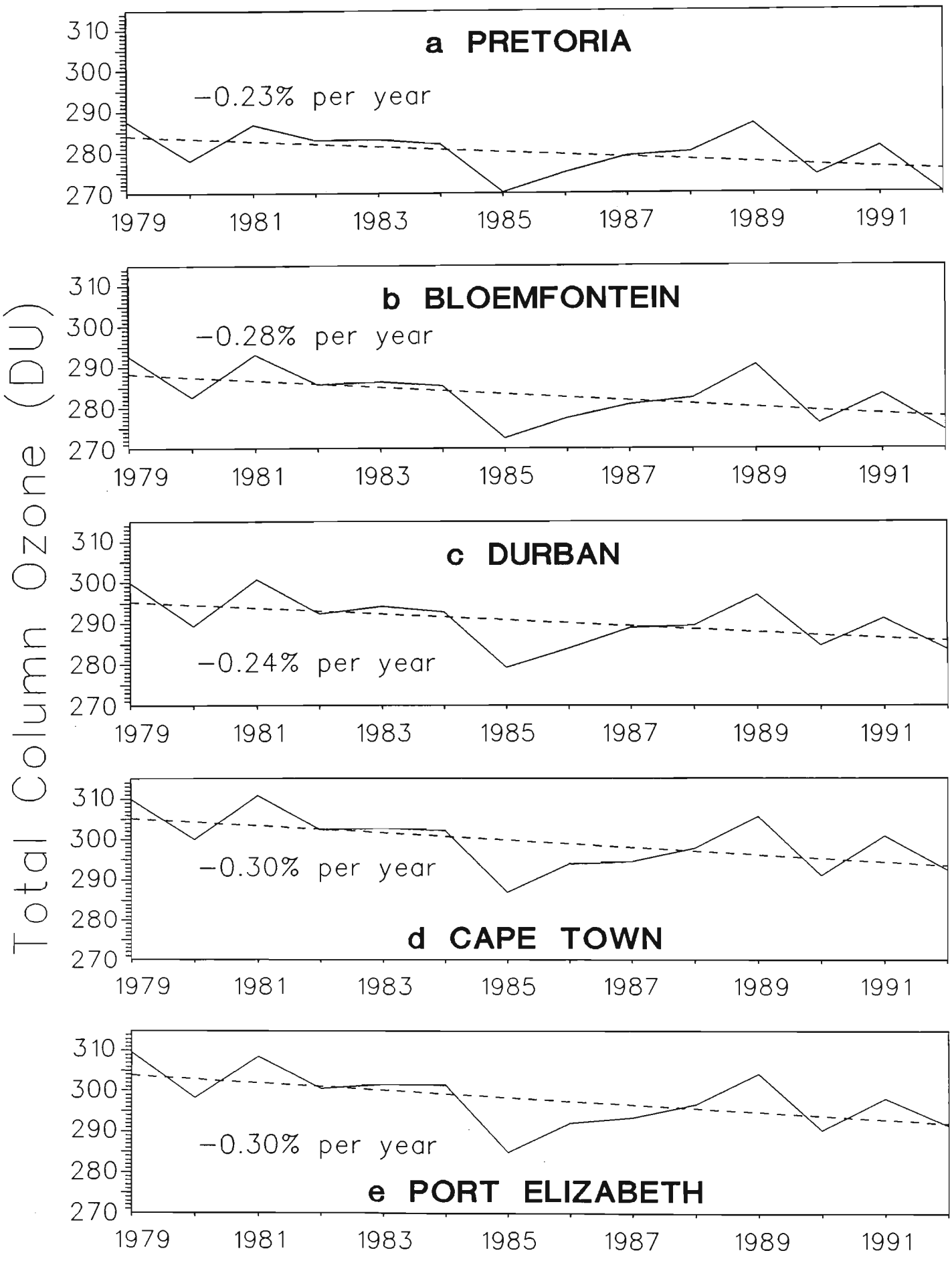


Figure 6.4: Yearly average total column ozone above each of the 5 cities. The long-term linear trend is shown in each panel using a dashed line and the value of the trend as a % of the 1979 value is also shown on each plot.

percentage of the 1979 yearly average total column ozone, are shown on the plots and listed in Table 6.4.

| Location | Annual trend (DU/year) | Annual trend (%/year) | r value |
|---|---|---|---|
| Pretoria | -0.650 | -0.23 | 0.533 |
| Bloemfontein | -0.822 | -0.28 | 0.533 |
| Durban | -0.721 | -0.24 | 0.483 |
| Cape Town | -0.904 | -0.30 | 0.540 |
| Port Elizabeth | -0.912 | -0.30 | 0.543 |

Table 6.4: Trends in yearly mean total column ozone over the 5 sites in DU/year and as a percentage of the 1979 value.

The regression coefficients (r values) for the straight line fits are also listed in Table 6.4. All cities show a decline in yearly average total column ozone with values varying from -0.23% per year (-0.650 DU/year) at Pretoria (the minimum) to -0.30% (-0.912 DU/year) at Port Elizabeth (the maximum). The low r values indicate the high statistical uncertainty in these results. Over the period 1979 to 1992, each of the 5 cities experienced its minimum yearly average total column ozone in 1985.

### 6.3.4 Results from the monthly average total ozone statistical model

The monthly average total column ozone data for each city have been modelled statistically resulting in 5 sets of 14 model coefficients (see Chapter 5). Table 6.5 lists the values which characterize the long-term linear trend and offset. The location names are listed in order of increasing latitude. The negative trends generally increase with latitude indicating that the more southerly locations in South Africa are experiencing a more rapid decline in total column ozone than more northerly locations. The Port Elizabeth value of -0.0594 DU/month is more than 1.5 times as large as the -0.0393 DU/month trend observed at Pretoria. Although the negative trends are greater for the more southerly locations, the initial offsets are also higher for these locations as shown in Table 6.5. The result is that in general more southerly locations have higher total column ozone in spite

| Location | Linear trend | |
|---|---|---|
| | Trend (DU/month) | Offset (DU) |
| Pretoria | -0.0393 | 283.2 |
| Bloemfontein | -0.0525 | 287.6 |
| Durban | -0.0455 | 294.4 |
| Cape Town | -0.0585 | 304.3 |
| Port Elizabeth | -0.0594 | 303.0 |

Table 6.5: **Model coefficients for the long-term linear trend and initial offset.**

of the greater rates of decline.

Table 6.6 lists the model coefficients for the annual variation and the QBO signal.

| Location | Annual variation | | | QBO variation | | |
|---|---|---|---|---|---|---|
| | Amp. (DU) | Phase (months) | Period (months) | Amp. (DU) | Phase (months) | Period (months) |
| Pretoria | 15.0 | 4.87 | 12.00 | 5.8 | 1.64 | 29.77 |
| Bloemfontein | 16.5 | 5.38 | 12.02 | 6.1 | 1.97 | 29.91 |
| Durban | 19.2 | 5.19 | 12.02 | 6.0 | 2.05 | 29.94 |
| Cape Town | 22.1 | 5.58 | 11.99 | 6.9 | 1.99 | 29.89 |
| Port Elizabeth | 21.6 | 5.57 | 12.01 | 6.9 | 1.94 | 29.94 |

Table 6.6: **Model coefficients for the annual variation and the QBO signal.**

The amplitude of the annual variation generally increases with latitude as shown previously in the daily total column ozone and monthly average analyses. Investigation of the phase coefficients for the annual variation shows that the more southerly cities (Cape Town and Port Elizabeth) experience their annual minima earlier in the year (28 March) than the other three locations (19 April for Pretoria, 3 April for Bloemfontein and 9 April for Durban). This phenomenon has been noted previously in Figure 6.2. The fact that the annual variation periods are very close to 12 months provides confidence in the ability of the model to track the annual cycle. The amplitude of the QBO signal also generally increases

with latitude, while its phase and period show no latitudinal dependence. Note that the periods of the QBO signals fall within the expected range of 22 to 34 months (Gray and Pyle, 1992).

Table 6.7 lists the model coefficients for the solar cycle variation and the semiannual oscillation.

| Location | Solar cycle variation | | | Semiannual variation | | |
|---|---|---|---|---|---|---|
| | Amp. (DU) | Phase (months) | Period (months) | Amp. (DU) | Phase (months) | Period (months) |
| Pretoria | 1.1 | 58.59 | 187.50 | 5.1 | 4.08 | 6.00 |
| Bloemfontein | 2.1 | 56.25 | 187.50 | 4.5 | 4.20 | 6.00 |
| Durban | 2.7 | 29.30 | 150.00 | 4.7 | 4.10 | 6.00 |
| Cape Town | 3.4 | 24.81 | 150.00 | 1.9 | 7.51 | 7.50 |
| Port Elizabeth | 3.3 | 26.15 | 150.00 | 1.9 | 7.38 | 7.51 |

Table 6.7: **Model coefficients for the solar cycle variation and the semiannual signal.**

The anomalously high periods obtained for the solar cycle signal for Pretoria and Bloemfontein (187.50 months) result from poor resolution of the signal within the model and are not statistically significant. After rejection of these two data points it is not possible to derive latitudinal dependences in the solar cycle coefficients. However, the 150 month periods for the remaining three locations are close to the expected 132 month solar cycle period. As already emphasized in chapter 5, doubt exists concerning these solar cycle coefficients since it is not statistically sound to attempt to identify an 11 year signal in 14 years of data.

The anomalously high values obtained for the period of the semiannual variation for Cape Town and Port Elizabeth also result from poor model behavior and these signals should be rejected. No latitudinal trend in the remaining semiannual coefficients is evident. The periods of the three remaining signals are exactly that of a semiannual signal.

The model coefficients make it possible to predict future monthly average total column ozone above each city. Figures 6.5 a to 6.5 e show measured TOMS total column ozone from 1979 to 1992 (solid line) together with statistical model predictions from 1979 to 2000 (dashed line) for each location. Model values agree well with measured data from 1979 to 1992, although large departures are seen during

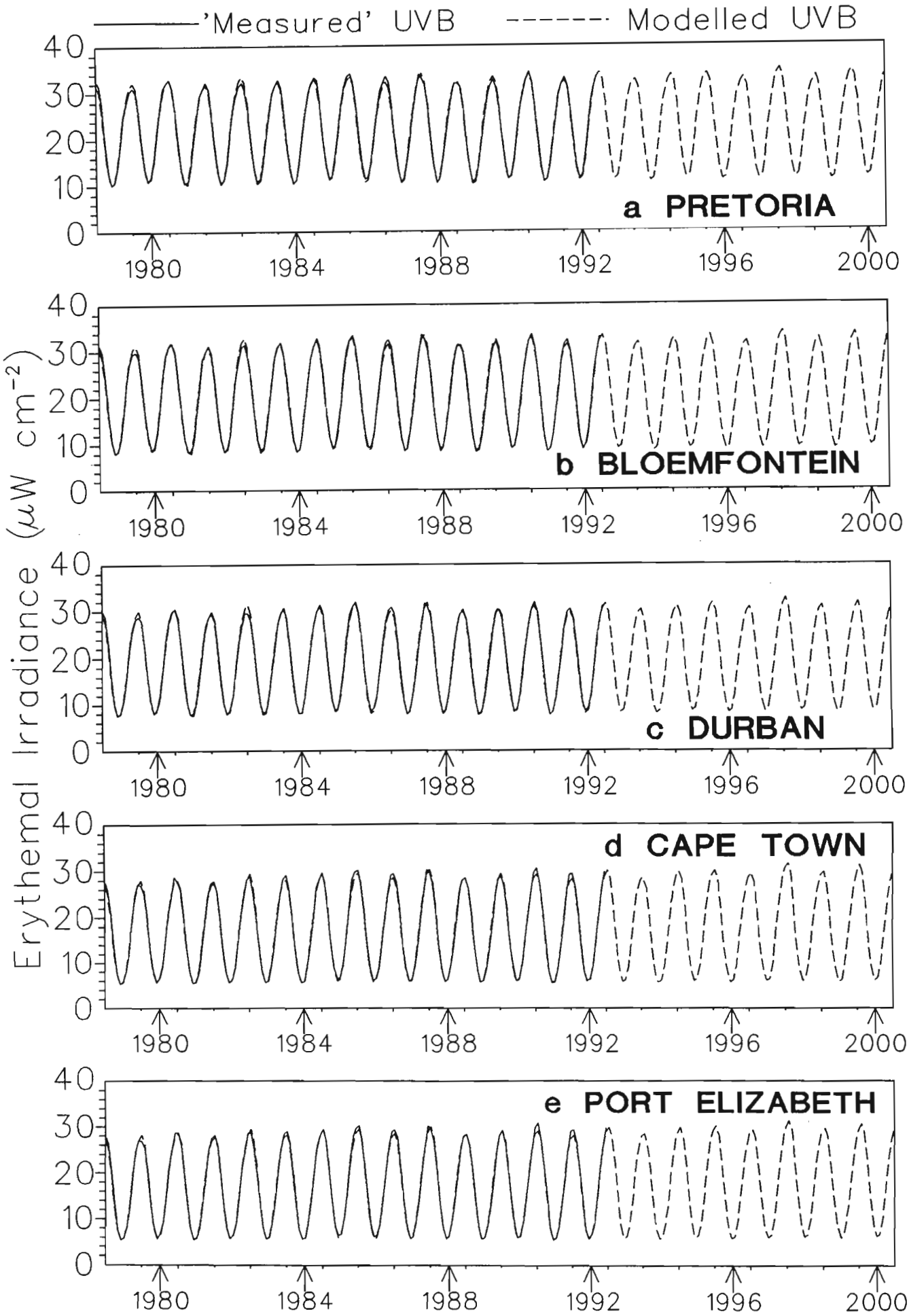


Figure 6.5: Measured (solid line) and model predicted (dashed line) monthly mean total column ozone above a) Pretoria, b) Bloemfontein, c) Durban, d) Cape Town and e) Port Elizabeth. The measured data are shown from 1979 to 1992 while the modelled data are plotted from 1979 to 2000.

1992. Previous studies (Gleason et al., 1993) using TOMS data have shown 1992 to be a year of record low total ozone and values are expected to deviate from statistical model predictions. The close fits observed between measured data and predicted values from 1979 to 1992 in Figure 6.5 provide confidence in the ability of the model to accurately forecast monthly average total column ozone. Care must be taken however since the model cannot account for nonlinear or episodic influences such as volcanic eruptions or changes to long-term local circulation etc.. Predictions from Figure 6.5 indicate that for all 5 cities the years of 1998 and 2000 could have a lower annual minimum in monthly average total column ozone by as much as 15 DU compared to previous years.

### 6.3.5 Results from the surface irradiance model

The data plotted in Figure 6.5 have been used as inputs to the surface irradiance model to predict erythemal doses to the year 2000. Monthly average total column ozone data were used to predict the noon surface irradiance spectrum for the middle day of each month. Modelled irradiance spectra were multiplied with the erythemal action spectrum of Diffey (1991) and integrated from 280 to 340 nm to obtain a measure of monthly erythemal dose. The results are shown in Figures 6.6 a to 6.6 e where monthly surface erythemal irradiances (in $\mu$W cm$^{-2}$) are plotted from 1979 to 2000 for each location. Solid lines show modelled surface erythemal irradiance based on measured TOMS total column ozone while the dashed curves show surface erythemal irradiances based on modelled total column ozone. The variation in surface erythemal irradiance is dominated by the seasonal cycle in solar irradiance and the variability in total column ozone makes only small contributions. Note that both the annual maximum and minimum surface erythemal irradiances decrease with increasing latitude. This results primarily from solar zenith angles being smaller for more northerly locations and less importantly, the more southerly locations experience greater total column ozone (as seen in Figures 6.3 and 6.4). Furthermore it can be seen that the dynamic range in surface erythemal irradiance increases with latitude partially as a result of an increase in the dynamic range of the total column ozone (as seen in Figure 6.1) and, more importantly, due to the increase of the seasonal range in the solar insolation with latitude. Projections to the year 2000 indicate that surface erythemal irradiances may be higher than normal during the summers of 1997/1998 and 1999/2000 for all 5 cities. However, since the annual variation

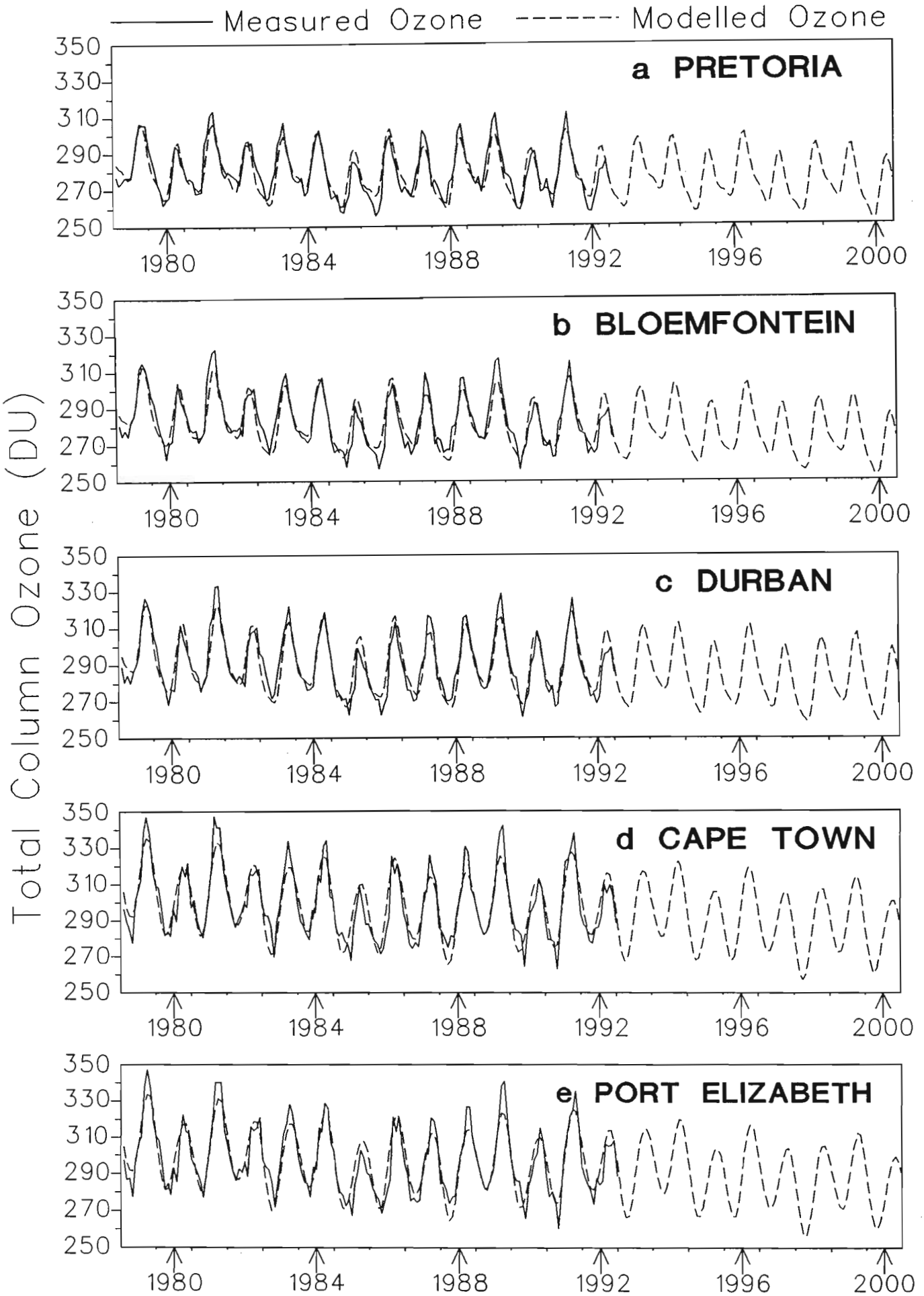


Figure 6.6: Modelled surface erythemal irradiances based on the data shown in Figure 6.5. The UVB irradiances based on the measured TOMS total column ozone (1979 to 1992) are shown using a solid line in each panel while the irradiances based on modelled total column ozone (1979 to 2000) are shown using dashed lines.

in surface UV is dominated by the seasonal solar cycle, UV trend detection is difficult.

## 6.4 Discussion

### 6.4.1 Long-term variability in South African total column ozone

The forcing mechanisms for the long-term trend, the annual variation, the QBO, the solar cycle variation and the semiannual oscillation have been discussed in detail in Chapter 5. The discussion is expanded below to provide an interpretation of long-term linear trends and annual variation specific to the South African context.

#### The long-term linear trend

Negative trends in total column ozone were observed above all 5 cities as summarized in Table 6.8 where the annual trends, obtained from annual means tend

| Location | Annual trend (DU/month) × 12 | Annual trend (DU/year) |
|---|---|---|
| Pretoria | -0.471 | -0.650 |
| Bloemfontein | -0.630 | -0.822 |
| Durban | -0.546 | -0.721 |
| Cape Town | -0.702 | -0.904 |
| Port Elizabeth | -0.713 | -0.912 |

Table 6.8: Long-term linear trends in total column ozone above each site obtained from monthly and yearly averages.

to be higher than those obtained from the monthly means for all 5 cities. These discrepancies probably result from the statistical uncertainty in the linear fits (low $r^2$ values). The predominant features of the linear trends are:

- The negative linear trend generally increases with latitude, although Bloemfontein experiences a more severe trend than Durban in both annual and

monthly means even though it is further north.

- The more southerly locations generally have higher total column ozone over the 14 year period that the more northerly locations.

Quantitative values of the observed linear declines may not be statistically significant although there is certainly a decline which tends to increase with latitude.

The mechanisms responsible for the observed subtropical decrease in ozone are not well understood. In situ chemistry may be responsible for the observed depletion only if sulfate aerosols can sufficiently stimulate heterogeneous chemistry or if there is considerable export of chemically processed air from the Antarctic vortex to midlatitudes (section 5.5.1). For any mechanism involving the transport of processed air away from the Antarctic vortex, it may be expected that locations closer to the vortex will experience enhanced ozone depletion as observed in Table 6.8. As mentioned in Chapter 1, it is possible that the Antarctic circumpolar vortex does not act as a containment vessel for the polar stratosphere, but rather acts as a 'flowing processor' (Randel, 1993 a). Under these circumstances, the vortex 'flushes' its whole volume over a 16 to 24 km altitude range with an approximate 30 day timescale. This would result in the export of processed chlorine rich and ozone poor polar air to lower latitudes. It has been noted (Randel, 1993 a) that the observed ozone depletion at midlatitudes in both hemisphere may be a signature of vortex processing. Being one the countries closest to the Antarctic continent, it behooves South Africa to embark on campaigns to measure local levels of stratospheric sulfate aerosol and active chlorine.

### The annual variation

The annual variation in total column ozone observed in Figures 6.1, 6.2 and 6.3 is caused primarily by seasonal variation in solar irradiance and changes in the southern hemisphere general circulation associated with the formation and breakdown of the Antarctic circumpolar vortex as discussed in section 5.5.9.

Figures 6.7 a to 6.7 d show National Meteorological Center (NMC) monthly average meridional cross-sections of the zonal mean wind for the entire globe, from 0 km (1000 hPa) to 50 km (1 hPa) altitude, for the months of January, April, July and October respectively (Randel, 1992). Units are $m\,s^{-1}$ with contours in $5\,m\,s^{-1}$ intervals below 70 hPa and $10\,m\,s^{-1}$ intervals above 70 hPa. Shaded areas

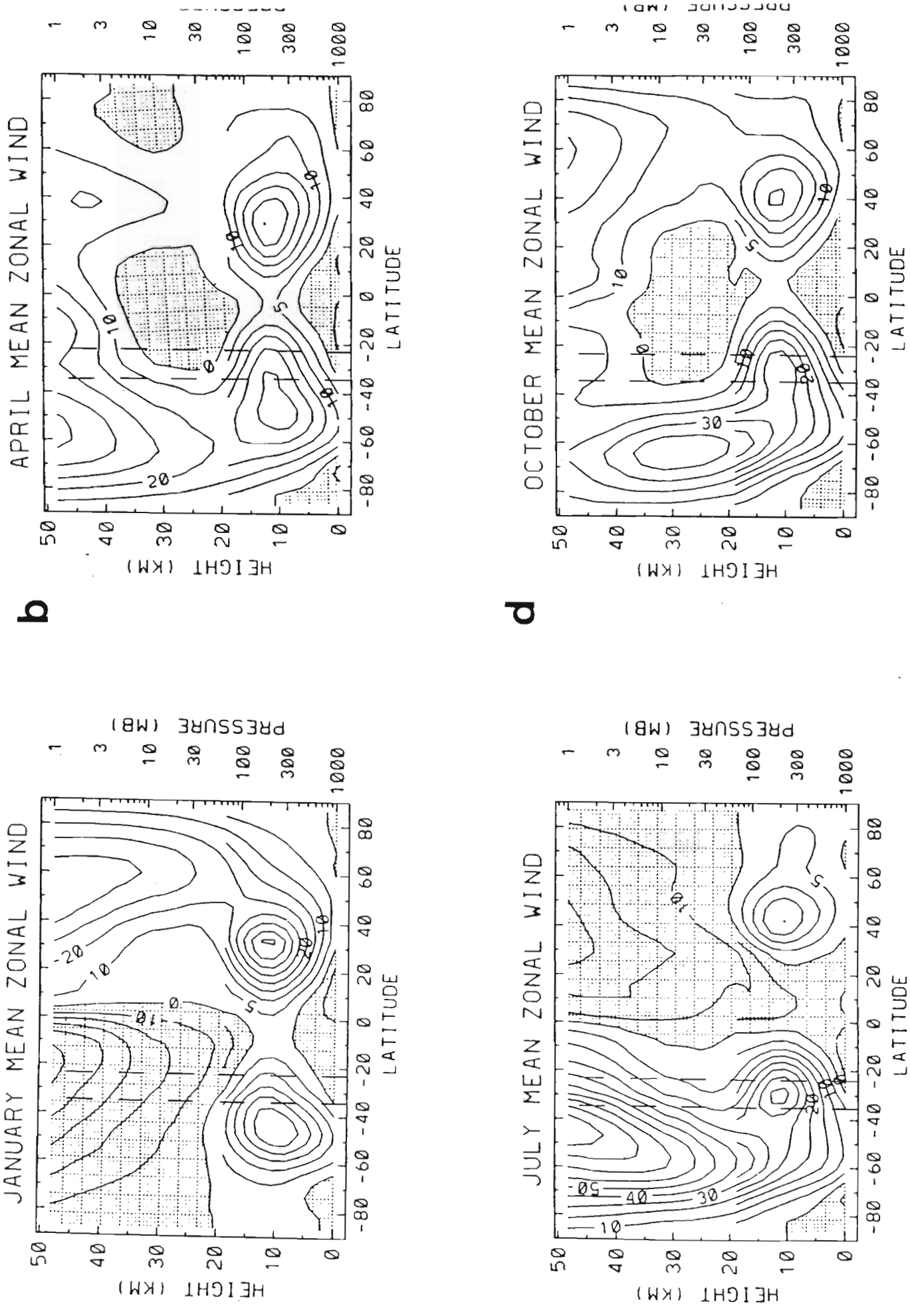


Figure 6.7: NMC monthly mean zonal wind speeds for a) January, b) April, c) July, and d) October, from Randel (1992). The northernmost and southernmost borders of South Africa are shown by vertical dashed lines.

show easterly flow while white areas show westerly flow. The northernmost and southernmost borders of South Africa are shown by two vertical dashed lines in each figure. Early in the year (Figure 6.7 a, January) the zonal flow is predominantly weak easterly with a small westerly jet at around 10 km (300 hPa). South Africa lies just equatorward of this 25 $m\,s^{-2}$ jet. By April (Figure 6.7 b) this jet is still at 10 km although a stronger westerly jet of around 50 $m\,s^{-1}$ has started to form at high altitudes (50 km, 1 hPa). It is thought that the formation of this stronger, high altitude jet may impede the poleward flow of ozone rich air from the source regions above the equator as discussed in Chapter 5. By July (Figure 6.7 c) the high altitude jet has increased greatly in strength reaching higher than 90 $m\,s^{-1}$ between 40° and 50°S above 45 km. The lower jet has moved directly over South Africa and increased slightly in strength to 30 $m\,s^{-1}$. Note that through the months of January, April and July, zonal winds at 300 hPa above South Africa increase steadily. By October (Figure 6.7 d) the high altitude jet has moved downward and decreased in strength to about 40 $m\,s^{-1}$. The collapse of this jet may now allow ozone rich air over the midlatitudes to flow southward and chemically perturbed air from within the vortex to drift to lower latitudes (South Africa).

The dynamical evolution of this stratospheric jet perturbs the general circulation of the Southern Hemisphere. The result is that a ridge of ozone rich air forms at midlatitudes which spreads equatorward during the winter, increasing total column ozone over South Africa. Following the breakup of the vortex in early summer, ozone rich air from the ridge flows southward to mix with ozone poor air over the pole, resulting in a decrease in total column ozone over South Africa. The proximity of the 5 South African cities, selected in this study, to this mid-latitude ridge determines the amplitude of the annual variation in total column ozone as shown in Figure 6.3 and in the annual amplitude coefficients of the statistical model.

It may also be possible that year-to-year changes in the position and amplitude of this mid-latitude ridge are responsible for the observed interannual differences in total column ozone over South Africa (see Figure 6.4). These characteristics of the mid-latitude ozone maximum may in turn be related to Southern Hemisphere planetary wave activity. Since planetary waves are primarily responsible for the meridional transport of ozone from low to middle latitudes, increased wave activity could result in enhanced transport of ozone rich air, increasing the

amplitude of the midlatitude ridge. This would result in elevated total column ozone over South Africa. This speculation requires further investigation to establish whether there is a connection between interannual differences in midlatitude planetary wave activity and South African total column ozone.

The study has also shown that the day-to-day variability in total column ozone above each of the 5 cities tends to increase with latitude (see Figures 6.1a to 6.1e). This may result from large scale weather systems moving in an easterly direction south of the country. Such fronts are known to influence total column ozone at South African latitudes (Diab et al., 1992). Stations closer to these systems will experience greater interannual variability in daily total column ozone.

### 6.4.2 Long-term variability in South African erythemal irradiance

This study made use of a simple UV irradiance model to determine the effect of changes in total column ozone over the 5 cities on surface erythemal irradiance. It appears that surface UVB irradiances will increase as a result of the long-term linear decrease in total column ozone although, as shown in Figure 6.6, the latitudinal differences far exceed any temporal variations. Furthermore, the seasonal variation in solar irradiance dominates the annual variation in surface erythemal UV. For Pretoria (the most northerly location) winter erythemal irradiance levels are typically 35% of the summer levels while for Port Elizabeth (the most southerly location) the winter values are about 22% of the summer maxima. For these 5 South African cities, trends in total column ozone until the turn of the century may have only a small effect on surface erythemal irradiances. However, such predictions must be treated with caution since the statistical ozone model only accounts for currently observed linear and cyclical trends. Catastrophic, unpredictable events such as the eruption of volcanos cannot be accounted for.

# Chapter 7

# INTERRELATIONSHIPS BETWEEN MIXING RATIOS OF LONG-LIVED CONSTITUENTS IN THE NCAR CCM2

## 7.1 Introduction

In Chapter 2 it was shown that planetary wave motion owes its existence to the conservation of potential vorticity (PV). Wave propagation arises from a balance between relative vorticity ($\zeta$) and planetary vorticity ($f$) in the expression for potential vorticity. However, to make use of PV conservation as a diagnostic for planetary wave propagation, it is necessary to know the conditions under which it is conserved.

Theoretical analysis by Plumb and Ko (1992) has shown that simple relationships exist between the concentrations of simultaneously measured long-lived stratospheric constituents. Under certain conditions of transport and chemistry, these relationships may be used to infer information concerning the global fluxes of the species as well as the ratio of their lifetimes. In the preliminary study presented in this chapter, data obtained from version 2 of the National Center for

Atmospheric Research (NCAR) Community Climate Model (CCM2) are used to test this theory by plotting simultaneously measured $CH_4$ concentrations against $N_2O$ concentrations for four months (January, April, July and October) at five different stratospheric pressure levels (105.4, 70.7, 31.0, 10.0 and 0.937 hPa). $N_2O$ has been used as the base species, as did Plumb and Ko (1992), since it has a sufficiently long lifetime in the lower stratosphere to meet the criteria imposed by the Plumb and Ko formalism. Furthermore, by treating PV as a long-lived tracer and plotting it against $N_2O$ concentrations as was done for $CH_4$, a method is suggested whereby PV lifetimes may be inferred. It must be stressed that this is only a preliminary study to examine the feasibility of the approach and to suggest directions for future work.

This investigation also has important implications for atmospheric measurement campaigns. To preserve the steep tracer gradients across the vortex edge, Schoeberl et al. (1992) used PV rather than geographic latitude as a spatial coordinate when analyzing data from airborne polar trace gas measurement missions. Since the position of the vortex oscillates about the pole in response to wave forcing (Chapter 4), the use of latitude as the spatial coordinate when averaging measurements over time, tends to flatten meridional tracer concentration profiles. However, since the PV distribution moves in tandem with the ozone hole, labeling the measurements with respect to PV preserves the steep tracer gradients. This approach is only justified if PV is conserved in the region of interest eg. across the vortex edge. It is hoped to investigate the high latitude behaviour of PV in this study.

It has been mentioned in previous chapters that the distribution of long-lived atmospheric trace gases is determined in part by planetary wave activity. Meridional cross-sections of tracer concentrations show the effects of the midlatitude surf zone where planetary wave breaking is frequent and tends to flatten isopleths (surfaces of constant tracer concentration). This phenomenon has been observed in $N_2O$ concentrations obtained from the Cryogenic Limb Array Etalon Sounder (CLAES) instrument on the Upper Atmosphere Research Satellite (UARS) (Randel et al., 1993). Schmidt (1982) and Garcia et al. (1992) have also demonstrated the importance of planetary waves in determining the global distribution of $N_2O$.

In addition to midlatitude surf zone effects, it has also been shown that planetary wave erosion of the vortex edge strongly influences meridional gradients of trace species (Schoeberl and Hartmann, 1991). During the midwinter and spring,

meridional gradients across the vortex boundary are much sharper than would be expected from simple radiative heating and cooling arguments. Observations during the Airborne Antarctic Ozone Experiment (AAOE) confirmed steep gradients across the vortex boundary for the long-lived species of $N_2O$, $O_3$ and $H_2O$ (Schoeberl et al., 1992).

## 7.2 Theory

The theoretical analysis of Plumb and Ko (1992) has shown that:

1. Two species whose local lifetimes are longer than quasi-horizontal transport time scales are said to be in climatological slope equilibrium. They then share surfaces of constant mixing ratio (isopleths) and scatter plots of the concentration of one species with respect to the other becomes a compact curve. It was also shown theoretically that the slope of this compact curve at any point is the ratio of the net global upward fluxes of the two species through the corresponding surface of constant mixing ratio.

2. Furthermore, two species whose local lifetimes are greater than vertical transport time scales are in gradient equilibrium and the scatter plot of the concentration of one with respect to the other becomes a linear curve. It was shown theoretically that for species whose atmospheric lifetimes are determined by stratospheric sinks, the slope of this linear relationship in the lower stratosphere can be related to the ratio of their atmospheric lifetimes.

It is hoped to makes use of the second of these two theoretical relationships to suggest a means of establishing the lifetime of PV from plots of PV versus $N_2O$. Care must be taken however when applying these relationships to the data obtained from the CCM2. Plumb and Ko made use of theoretical arguments based on a 2-dimensional transport formulation to show how these mixing ratio correlations are maintained in the zonally averaged tracer distributions. Although the NCAR CCM2 is a 3-dimensional general circulation model (GCM) that does not incorporate the 2-dimensional K-theory formalism (section 2.9.4) used by Plumb and Ko, it must be noted that model data are being compared against model data to investigate relationships suggested by a theoretical formalism.

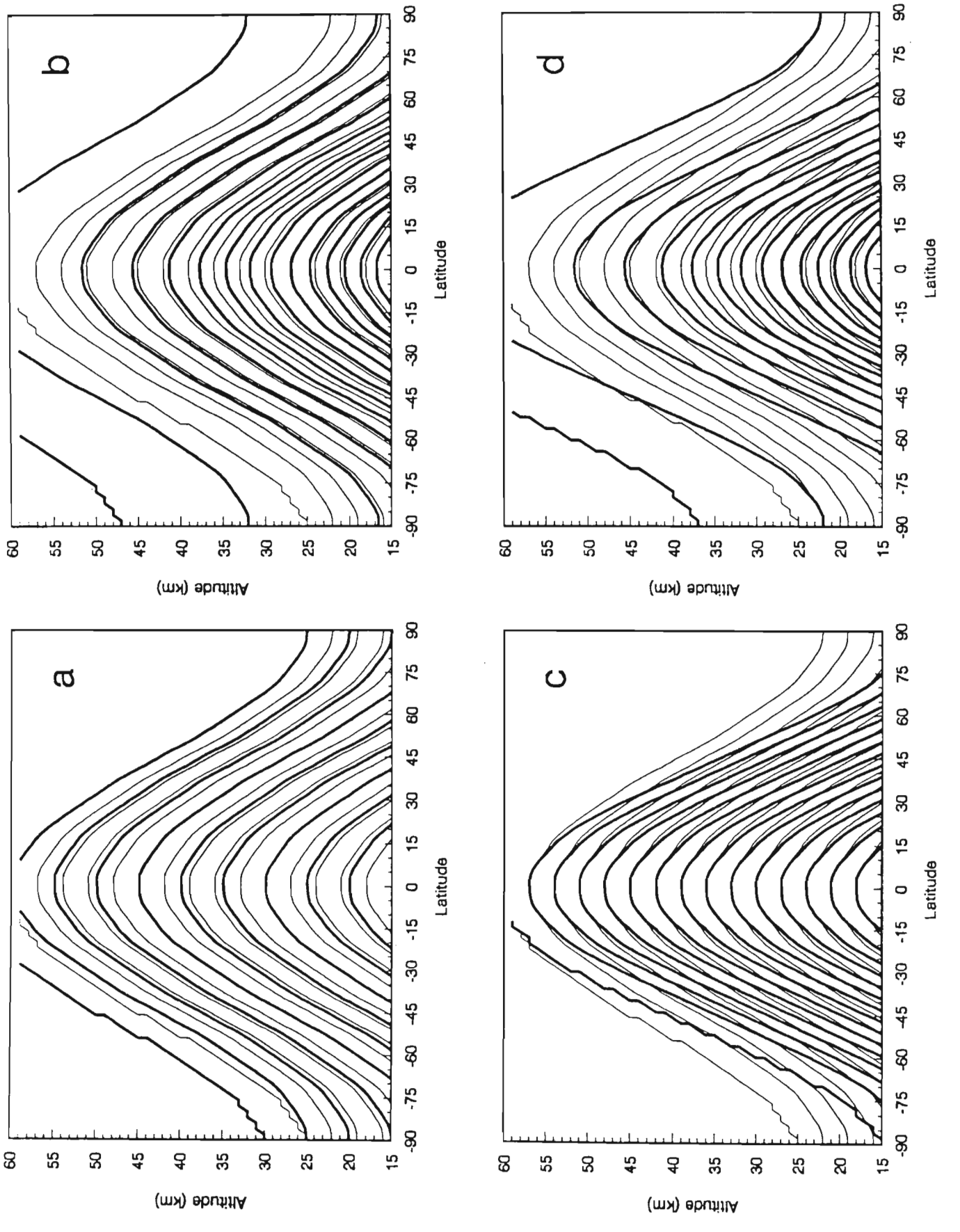


Figure 7.1: Meridional cross-sections of the concentrations of two ideal stratospheric tracers. Species 1 concentrations are shown using thin lines while species 2 concentrations are shown using bold lines. Variations between panels are discussed in the text.

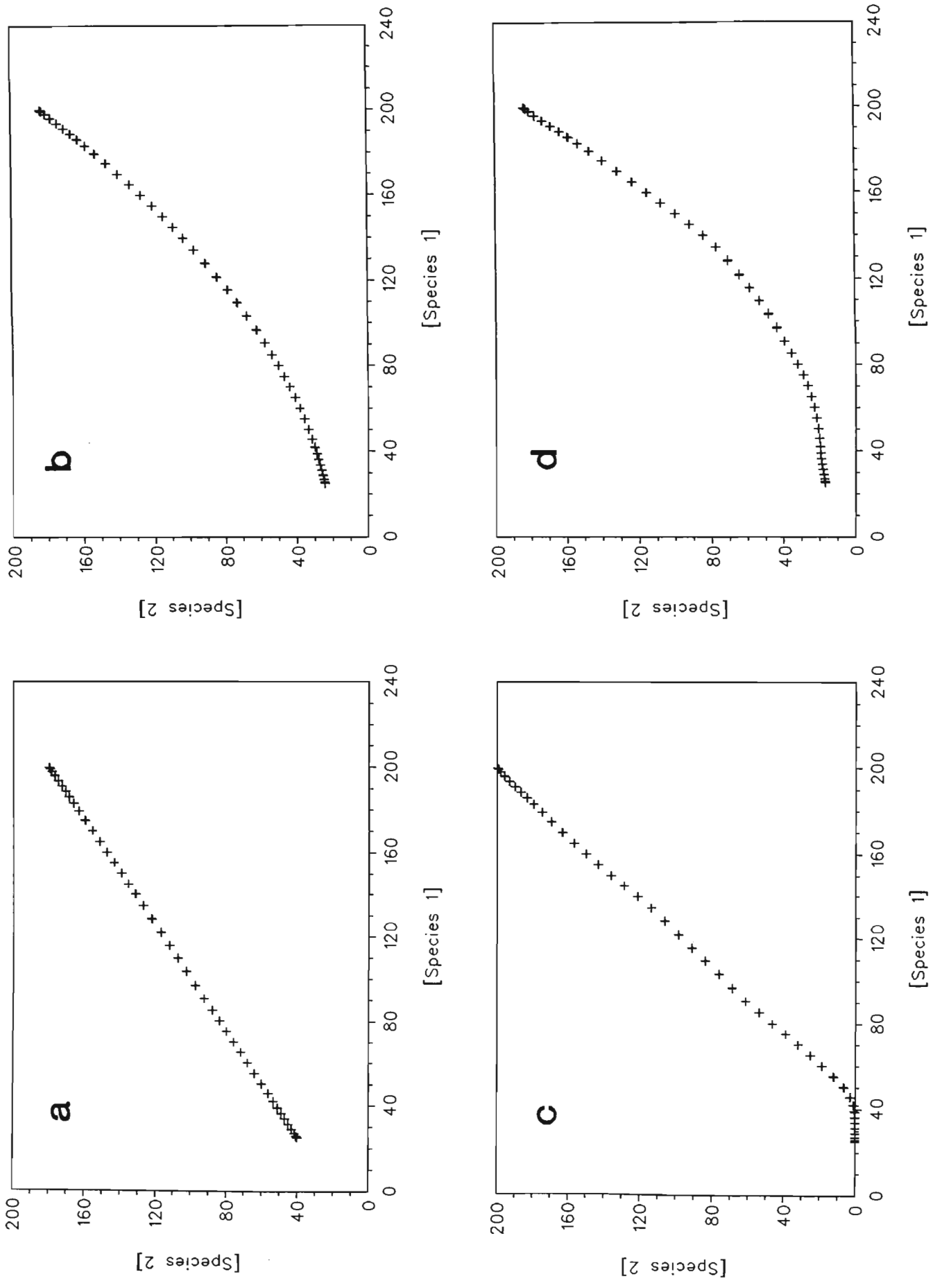


Figure 7.2: Scatter plots of species 1 concentrations against species 2 concentrations at the 20 km level in Figure 7.1.

The concepts discussed above are clarified in Figures 7.1 and 7.2. Four examples of meridional cross-sections of the concentrations of two ideal stratospheric species are shown in Figure 7.1. The species are not associated with real atmospheric constituents and have been generated artificially. Actual mixing ratio values are irrelevant to the discussion and have therefore been omitted from the plots of Figure 7.1 for clarity. These example tracers are assumed to have sources in the troposphere and sinks in the stratosphere which, together with the meridional circulation, accounts for the morphology of the distributions seen in Figure 7.1. Figure 7.3 shows a highly schematic view of the general meridional circulation together with single isopleths for two ideal species.

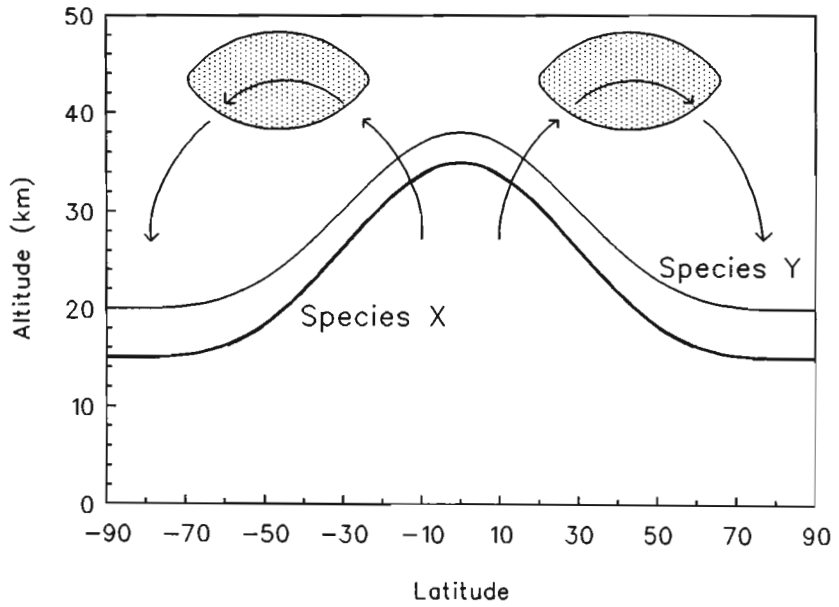


Figure 7.3: A highly schematic presentation of the general circulation showing the processes leading to the assumed morphology of the tracer concentrations observed in Figure 7.1

The stippled regions in Figure 7.3 indicate supposed sinks for both species in the middle stratosphere. Upwelling in the equatorial regions brings tracer rich air from the lower stratosphere into the higher stratosphere resulting in a maximum in the meridional cross-section of the species concentrations. As the air flows poleward it passes through the sink regions and the subsidence over the high latitudes contains tracer poor air. This suppresses the tracer isopleths at high latitudes.

In each panel of Figure 7.1, both species show decreasing concentrations with height as well as characteristic equatorial maxima and polar minima. Species

1 (plotted using thin lines) is kept constant while the distribution of species 2 (plotted using bold lines) is varied in each of the four panels. Figure 7.2 shows the scatter plots of the concentrations of species 1 against species 2 at 20 km corresponding to each of the four panels of Figure 7.1.

Panel **a** of Figure 7.1 shows the meridional cross-section of the concentrations of two species that are in climatological slope equilibrium as evident by the similarity of their isopleths. The concentrations of both species fall linearly with height with the result that they are also in gradient equilibrium. The corresponding scatter plot (Panel **a** of Figure 7.2) shows a linear relationship between the concentrations of the two species as expected from point 2 above.

In panel **b** of Figure 7.1 the concentrations of species 2 (plotted using bold lines) now have a second order dependence on height. This could result from a relative sink of species 2 in the stratosphere with respect to species 1. The tracers are still in climatological slope equilibrium with isopleths of similar shape. The scatter plot shown in panel **b** of Figure 7.2 again shows a compact curve although it is no longer linear. The positive curvature indicates that species 2 experiences a relative sink compared to species 1.

In the third test case (shown in panel **c** of Figure 7.1), although both species concentrations fall linearly with altitude, they are no longer in climatological slope equilibrium. Species 2 concentrations fall more rapidly with increasing latitude compared to species 1. Panel **c** of Figure 7.2 indicates that the relationship between the two tracer concentrations shows small deviations from linearity and at high latitudes, species 2 concentrations fall to zero.

The tracer concentrations shown in panel **d** of Figure 7.1 are neither in climatological slope equilibrium nor in gradient equilibrium. Species 1 concentrations fall linearly with altitude while species 2 concentrations have a second order dependence on height. The scatter plot of the concentrations is still a compact curve similar to that shown in panel **b** of Figure 7.2. Again the positive curvature indicates a relative sink of species 2 with respect to species 1.

The scatter plot of the concentration of one species with respect to the other is therefore related to the sources and sinks of the species (their lifetimes) and the meridional circulation (Figure 7.3). Seasonal changes in these scatter plots may then act as a diagnostic for local dynamics, including planetary wave mixing at midlatitudes, wave erosion of the vortex edge (and associated tracer gradient

steepening) and vortex isolation.

Mixing ratios of $N_2O$, $CH_4$ and PV have been extracted from a ten year model run of the NCAR CCM2. Scatter plots of zonal averages of these constituents show compact curves with strong latitude and altitude dependence, indicative of chemical sources and sinks. Furthermore, in some cases, the shapes of the compact curves observed in the scatter plots appear to be strongly dependent on local dynamics (eg. the formation of a polar vortex).

## 7.3 The CCM2 model

The three-dimensional equations of atmospheric dynamics are usually modelled in one of two ways (Brasseur and Solomon, 1986):

- In the grid point method, the variables are defined at regularly spaced locations in latitude, longitude and altitude. The spatial and temporal resolution of the model determines the stability and accuracy of the solution. Horizontal derivatives are replaced by finite differences.

- In spectral models, the horizontal structure of the meteorological fields is expressed by means of truncated spherical harmonic expansions (which can represent the effects of waves). In the vertical, the variables are determined at discrete grid points. In practice, the horizontal expansion is usually truncated and therefore resolves only large scale motions. This method has the advantage of allowing exact calculation of the horizontal derivatives.

The CCM2 is a three dimensional circulation model that serves as the basis for several studies of stratospheric transport, chemistry and dynamics (Boville, 1993). The model makes use of the spectral method described above, extending the spherical harmonic expansion to 21 waves. In this configuration the model makes use of a horizontal grid of 64 $\times$ 64 cells where the latitudinal spacing of the cells is not constant.

The middle atmosphere version of the CCM2 uses 44 layers in the vertical between the surface and $\sim$0.01 hPa. The vertical coordinate reduces to a pressure coordinate above 100 hPa. In the study presented in this chapter, the 26 levels from 105.4 hPa to 0.175 hPa have been used. The pressure surfaces are separated

by approximately 1.4 km.

The CCM2 model incorporates a full radiation parameterization which includes the effect of clouds. Transport of water and trace species is modelled using the 3-dimensional semi-Lagrangian method of Rasch and Williamson (1991). The initialization for the model run used in this study was taken from an earlier test version of the model rather than from observations (Boville, 1993).

Although the model accurately represents Northern Hemisphere dynamics, there are serious deficiencies in the Southern Hemisphere where the winds within the polar vortex during winter are almost double those observed in the real atmosphere. The source of this discrepancy is an inadequate gravity wave drag parameterization (Boville, 1993). Within the CCM2, gravity wave momentum deposition is parameterized using the technique of McFarlane (1987). This scheme treats only stationary waves generated by flow over orography and does not allow for nonzero phase speeds or generation of waves by shear or convection. This also results in excessively low polar temperatures as a consequence of thermal wind balance and a weak meridional circulation in the lower stratosphere.

The development of the Southern Hemisphere vortex is also too rapid in the upper stratosphere of the model, with the result that zonal wind speeds are about twice as strong as they should be in March. The rapid development of the vortex may be a result of the inadequate gravity wave drag, since the middle mesospheric Rayleigh friction cannot produce significant drag until the wind becomes quite strong (Boville, 1993). The CCM2 model also shows that the differences between the simulations of southern and northern winter are almost entirely due to parameterization of stationary gravity waves generated by topographic forcing and not due to differences in planetary wave activity.

Garcia et al. (1992) developed a severely truncated three dimensional model (allowing for a planetary wave 1 perturbation only) which predicted a $65\,\mathrm{m\,s^{-1}}$ maximum jet velocity which is in good agreement with climatological observations (see Figure 6.7 c). It was stated that the modelled jet strength was substantially weaker than expected under radiative equilibrium conditions as a result of planetary wave dissipation in the winter stratosphere.

## 7.4 A brief climatology of long-lived tracers in the CCM2

A detailed account of the annual cycle and interannual variability in the CCM2 has been given by Boville (1993). A brief synopsis is given here in preparation for the data analysis presented in section 7.5. Monthly average plots of zonal mean $N_2O$ and $CH_4$, used in the discussion of the climatology of CCM2, show the effects of midlatitude planetary wave breaking and tracer gradient steepening at the vortex edge. These processes have been invoked to account for many of the observations in previous chapters. The planetary wave modelling within the CCM2 accurately reproduces the isopleth flattening effects of wave breaking. In addition, examination of zonal mean temperatures shows the marked wintertime cooling of the Antarctic vortex and the weaker cooling of the Arctic vortex. The formation of PSCs within the isolated, cold Antarctic vortex has been discussed extensively in previous chapters. Due to space constraints, only the months of January, April, July and October have been selected to demonstrate the annual cycle in $N_2O$, $CH_4$ and temperature.

Figure 7.4 shows plots of zonal mean stratospheric $N_2O$ (in p.p.b.v.) from 105.4 hPa (15 km) to 0.175 hPa (60 km) for the four months listed above. In all four cases the meridional cross-sections of the $N_2O$ concentration appear similar to the artificial distributions shown in Figure 7.1 a. This cross-section morphology suggests that $N_2O$ has a sink in the stratosphere. Two minor destruction mechanisms for $N_2O$ in the stratosphere are through collisions with excited oxygen atoms in the ($^1$D) state (Brasseur and Solomon, 1986):

$$N_2O + O(^1D) \rightarrow 2NO \tag{7.1}$$

$$N_2O + O(^1D) \rightarrow N_2 + O_2 \tag{7.2}$$

However, the predominant destruction mechanism for $N_2O$ in the stratosphere is through photolysis:

$$N_2O + h\nu \rightarrow N_2 + O(^1D) \tag{7.3}$$

The $N_2O$ concentrations shown in Figure 7.4 fall rapidly with increasing latitude, indicating some impedance to the poleward flow. This phenomenon has been observed previously in UARS data (Randel et al., 1993). However, poleward of 30° to 45° (depending on the hemisphere and season) isopleths flatten, indicating

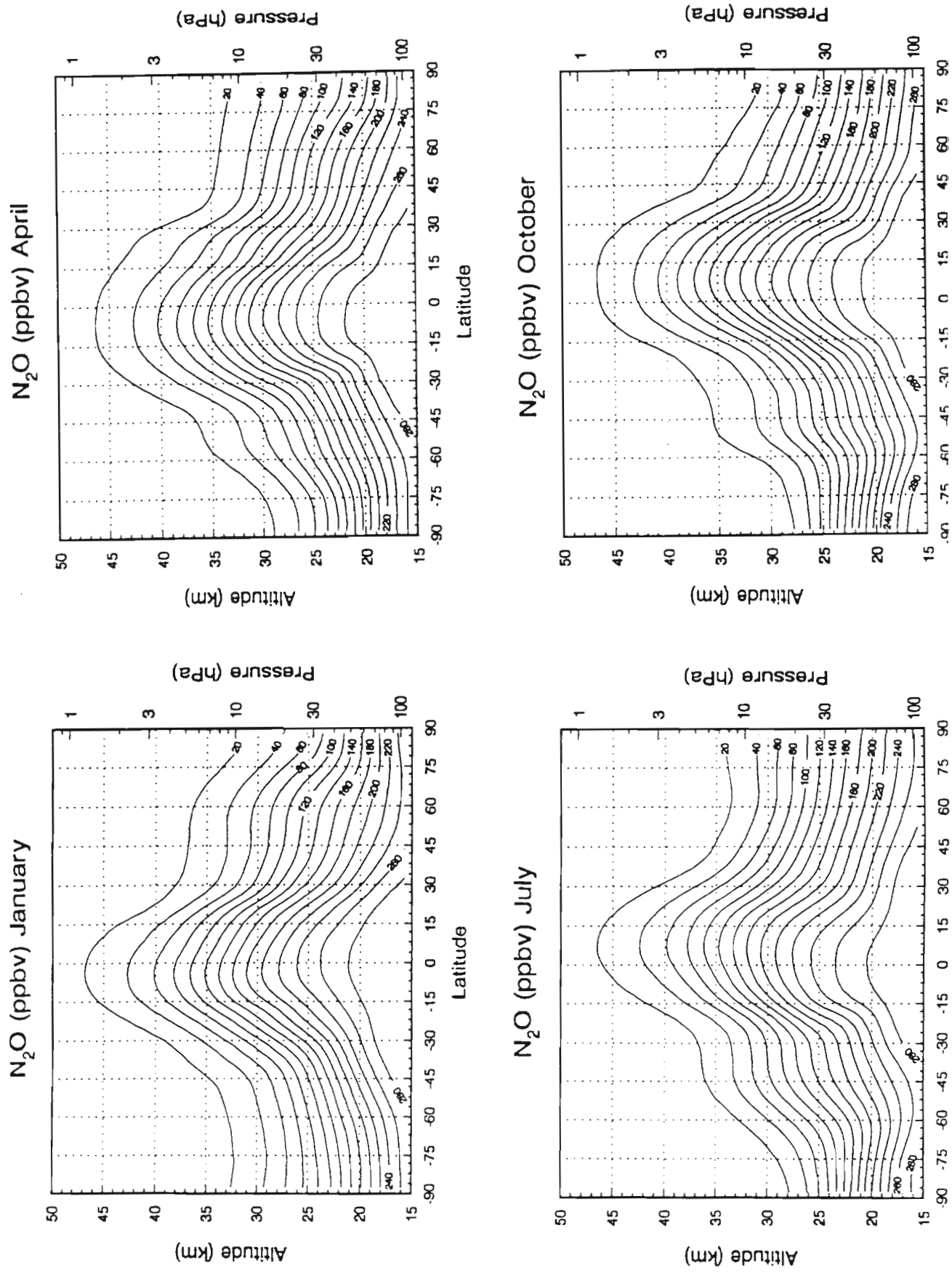


Figure 7.4: Meridional cross-sections of the zonal mean $N_2O$ concentrations (in p.p.b.v.) from the CCM2 for the months of January, April, July and October.

rapid mixing in the meridional direction. This is associated with the planetary wave breaking induced surf zone.

During January, when a weak Arctic vortex forms, another region of steep isopleth gradients is observed poleward of 60°N. The formation of the weak vortex prevents wave breaking induced transport of trace species into the Arctic stratosphere. Likewise, during the Southern Hemisphere winter (Figure 7.4 – July), tracer gradients are steep across the polar vortex above approximately 22 km and poleward of 45°S. Other modelling studies (Garcia et al., 1992) have shown that downwelling at high latitudes during the winter is responsible for the deep high latitude minima seen in the distributions of $N_2O$.

The Antarctic vortex is greater in area than that of the Arctic with the result that the steepening of the tracer gradients occurs closer to the equator (cf. 60°N and 45°S). Similar behaviour has been observed in the model of Garcia et al. (1992). By October, the high latitude tracer gradients in the Southern Hemisphere have become considerably steeper and have moved poleward compared to the distribution in July, as seen in Figure 7.4. This may result from planetary wave erosion of the vortex during the winter which tends to decrease the spatial extent of the vortex and steepens tracer gradients across the vortex edge. The polar minima in $N_2O$ are separated from the midlatitude higher mixing ratios by steep gradients that coincide with the position of the polar night jet.

Garcia et al. (1992) made use of model results to show that planetary wave breaking produces a well-mixed surf zone equatorward of the polar night vortex and drives a meridional circulation with downwelling on the poleward side of the vortex. This combination of mixing and downwelling produces shallow meridional gradients of trace gases in the subtropics and middle latitudes, and very steep gradients at the edges of the polar vortex (Garcia et al., 1992). These planetary wave signatures are clearly visible in the plots of Figure 7.4 as discussed above.

The meridional cross-sections of $CH_4$ concentrations shown in Figure 7.5 show similar morphology to the $N_2O$ distributions of Figure 7.4. There are a number of different destruction processes for $CH_4$ in the stratosphere and mesosphere (Brasseur and Solomon, 1986) including oxidation by the OH radical, reactions with excited oxygen $O(^1D)$, reaction with chlorine and photolysis. As for the $N_2O$ concentrations, the Southern Hemisphere $CH_4$ distributions often show two points of inflection in the isopleths (indicating reduced poleward transport) while it is

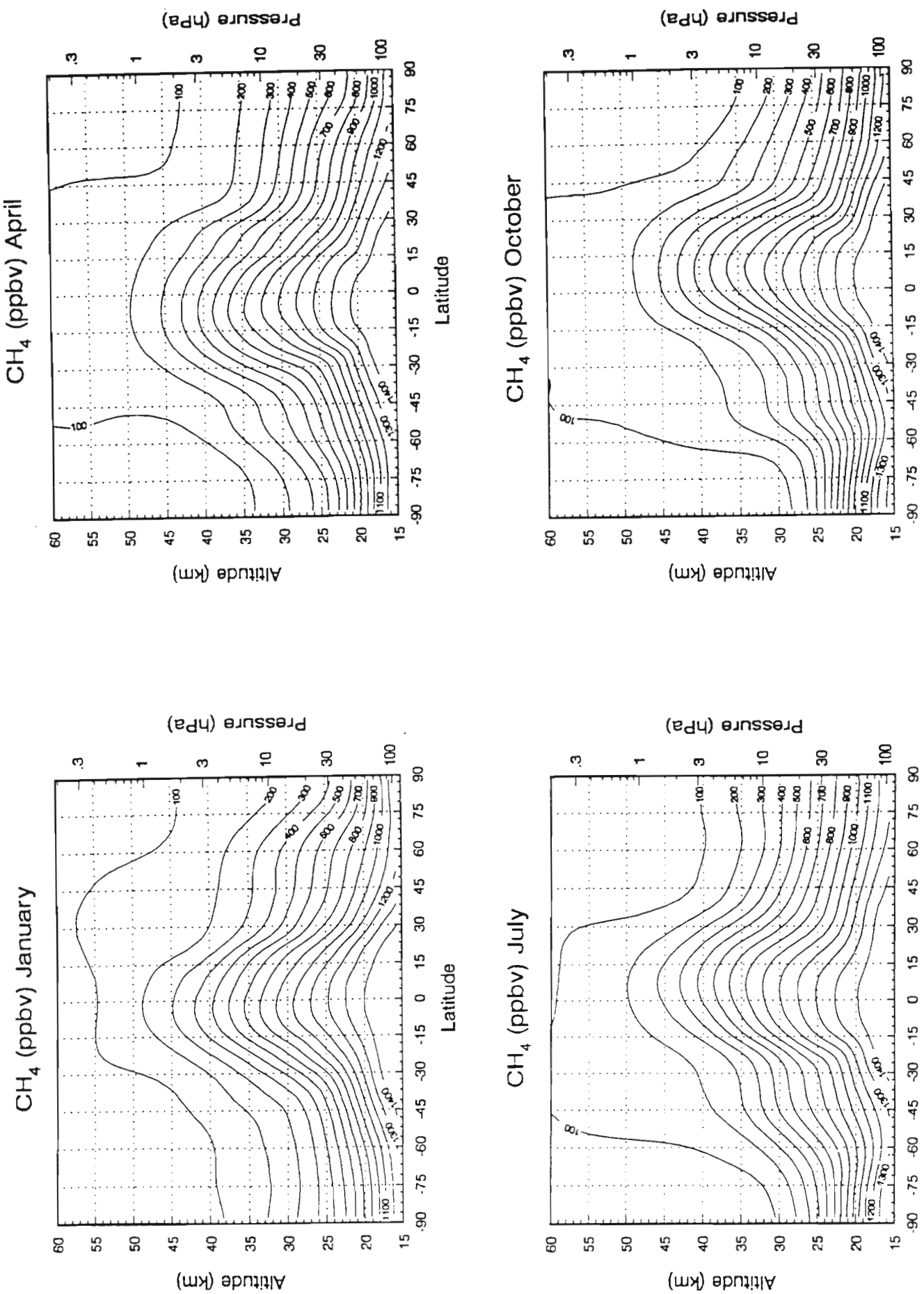


Figure 7.5: Meridional cross-sections of the zonal mean $CH_4$ concentrations (in p.p.b.v.) from the CCM2 for the months of January, April, July and October.

only during January that a high latitude tracer gradient steepening is observed in the Northern Hemisphere. This may be indicative of the stronger and more isolated Antarctic vortex compared to that of the Arctic.

Meridional cross-sections of monthly average zonal mean temperature are plotted in Figure 7.6 for January, April, July and October. Although during January the lower stratosphere temperatures in the Northern Hemisphere fall below 200 K, this is not cold enough for the formation of PSCs. With the onset of the Southern Hemisphere winter (April), high latitude stratospheric temperatures fall rapidly while Arctic stratospheric temperatures rise. By July (Figure 7.6), Antarctic temperatures have fallen below 180 K allowing PSC formation. Note the steep meridional temperature gradient between 45 and 75°S. Application of the geostrophic and thermal wind equations (Chapter 2) indicates that this will form a strong polar jet that increases rapidly in intensity with altitude. In October, although warming has begun to take place at the upper levels of the Antarctic stratosphere, lower stratosphere temperatures are still below 190 K, cold enough for the formation of both type I and II PSCs (Chapter 1).

## 7.5 Correlations between constituent concentrations

### 7.5.1 $CH_4$ versus $N_2O$ scatter plots

To test the theory of Plumb and Ko within the CCM2, scatter diagrams of $CH_4$ concentrations against $N_2O$ concentrations have been plotted in Figure 7.7 for the first 2 days of January, April and July from year 2 of the model run and for the first two days of October from year 1 of the CCM2 model run. Scatter plots at 105.4, 70.7, 31.0, 10.0 and 0.937 hPa are shown for all four months. In each panel of Figure 7.7, $CH_4$ concentrations (in p.p.b.v.) are plotted on the ordinate while $N_2O$ concentrations (in p.p.b.v.) are plotted on the abscissa. Open markers are used to show Southern Hemisphere values while solid fill markers are used to show Northern Hemisphere values. The data sets have been further divided into high latitude (triangle markers), midlatitude (circle markers) and low latitude (square markers). The data plotted in each panel are zonal means.

The purpose of this discussion is to illustrate how the gradients of the scatter

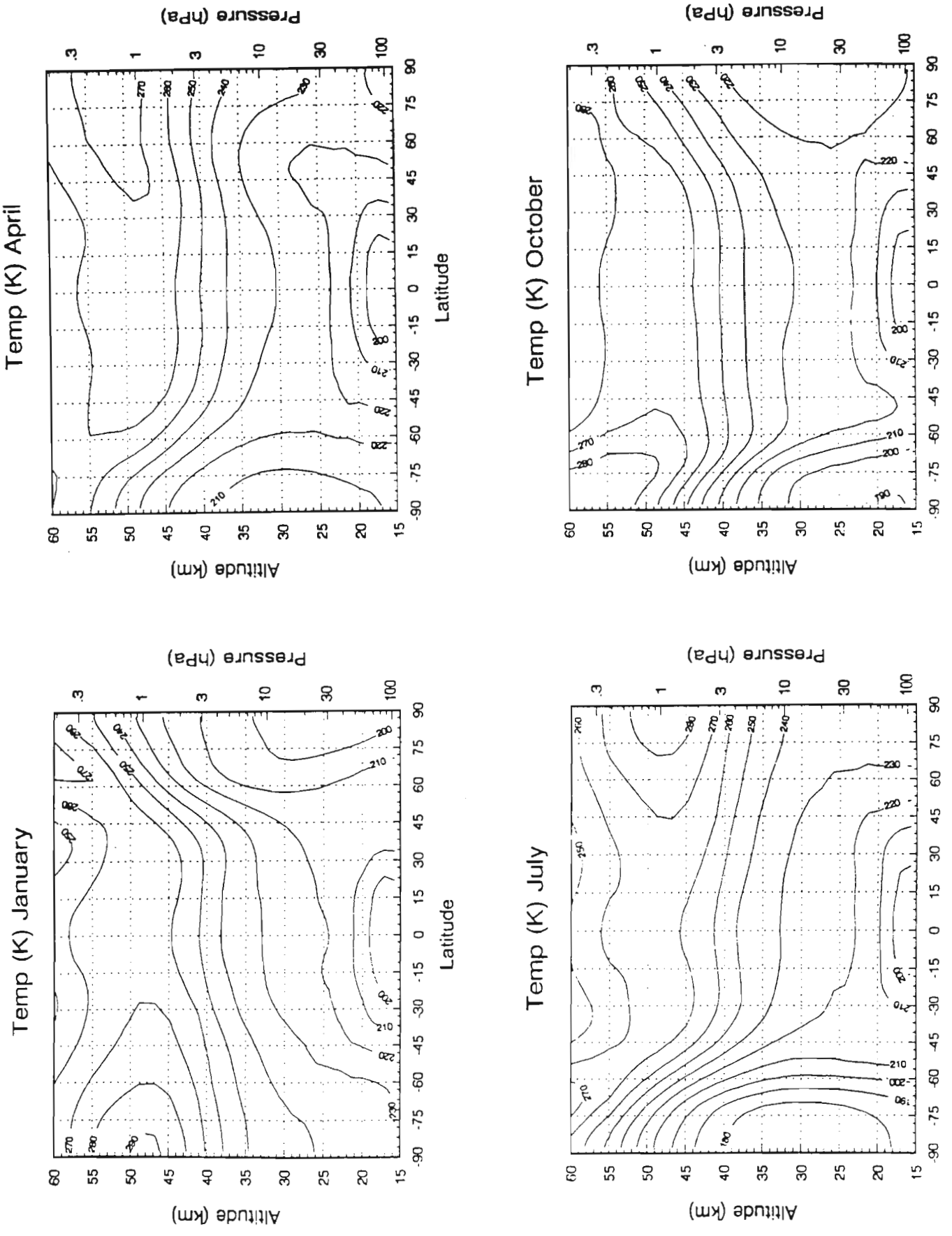


Figure 7.6: Meridional cross-sections of the zonal mean temperature (in K) from the CCM2 for the months of January, April, July and October.

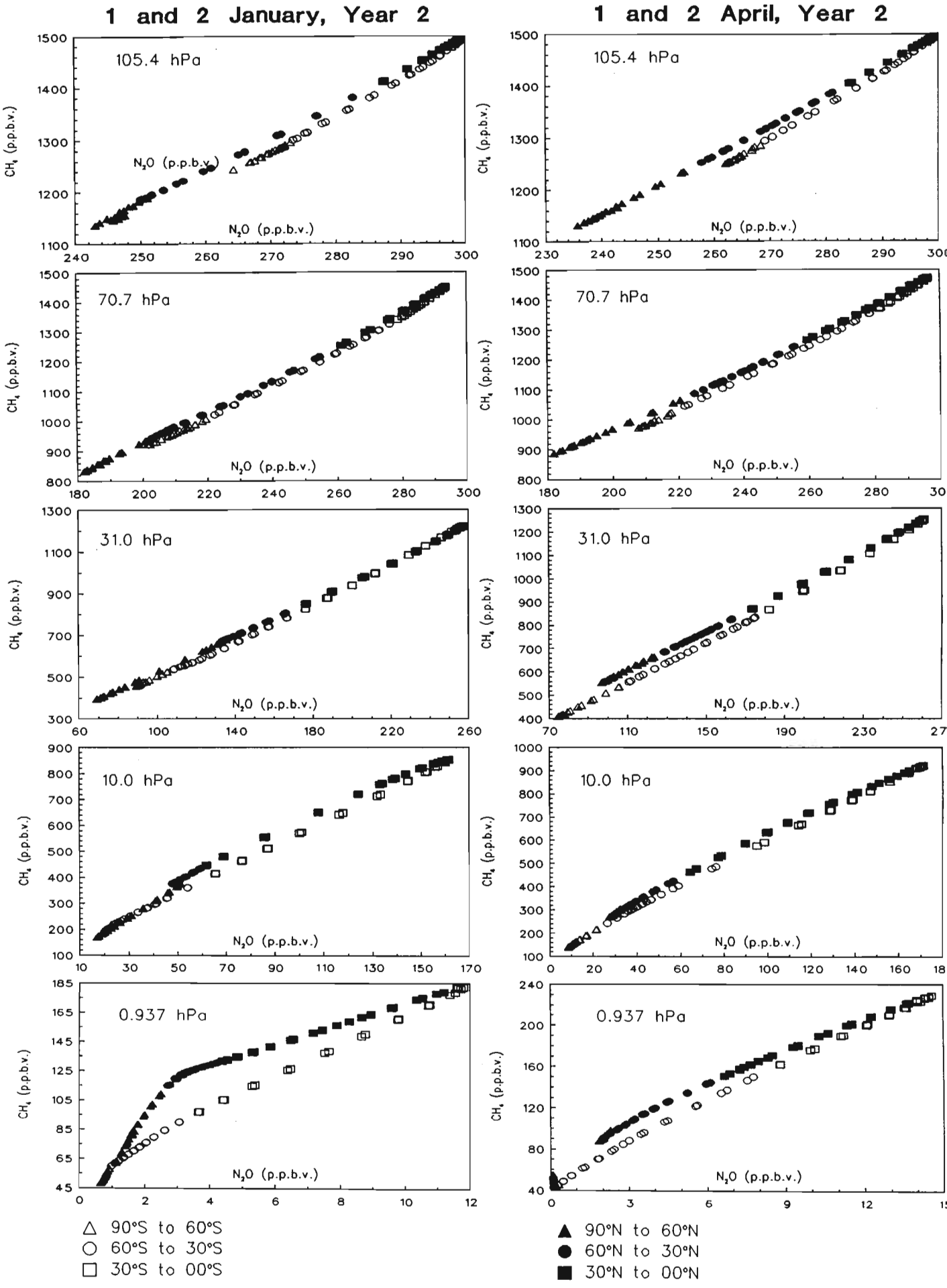


Figure 7.7: Scatter plots of $CH_4$ versus $N_2O$ concentrations (in p.p.b.v.) at 105.4, 70.7, 31.0, 10.0 and 0.937 hPa taken from the CCM2 model. Data for the first two days of January and April are shown.

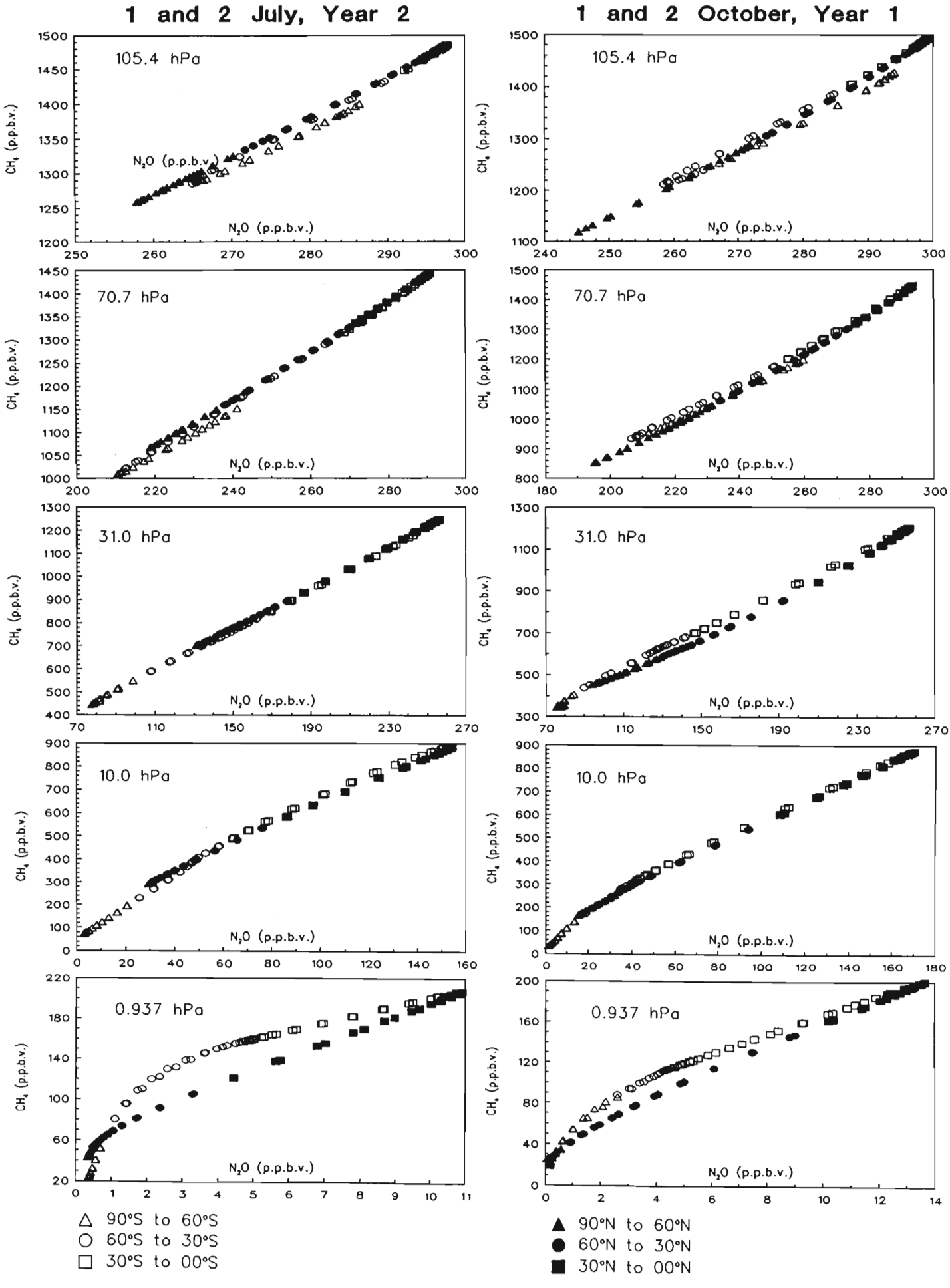


Figure 7.7 cont.: Scatter plots of $CH_4$ versus $N_2O$ concentrations (in p.p.b.v.) at 105.4, 70.7, 31.0, 10.0 and 0.937 hPa taken from the CCM2 model. Data for the first two days of July and October are shown.

plots of Figure 7.7 change with altitude and how they are related to the lifetimes of the two species. To this end meridional cross-sections of $N_2O$ and $CH_4$ lifetimes have been extracted from the CCM2 model for each of the four months of this study. The $N_2O$ lifetimes (in days) for each month are plotted in Figure 7.8. Values are in excess of $10^6$ days at low altitude and in the unilluminated winter high latitudes. Lifetimes fall rapidly with altitude until values of less than 30 days are observed above 45 km in the high latitudes of the summer hemisphere. Similar plots of $CH_4$ lifetimes are given in Figure 7.9. The general morphology of the distributions is similar to that of $N_2O$ except that $CH_4$ lifetimes show a local minimum between 45 and 50 km above which they rise again. This has important consequences for the sign of the curvature in the $CH_4$ versus $N_2O$ scatter plots shown in Figure 7.7. Since it is the ratio of the lifetimes that is instructive in this study, plots of the $CH_4/N_2O$ lifetimes are given in Figure 7.10. For each month, $CH_4$ lifetimes tend to be shorter than those for $N_2O$ in the lower stratosphere but longer in the upper stratosphere. The importance of the change in $CH_4/N_2O$ lifetimes with height will be discussed further below.

Returning to Figure 7.7, it can be seen in each of the panels the scatter plots show compact relationships although these are not always linear. There are a number of interesting features in these plots which are discussed below. In some cases, the discussion of these features requires a knowledge of the meridional circulation. To this end, residual circulations have been extracted from the CCM2 for the months of January and October. These are shown for reference purposes in Figures 7.11 and 7.12.

### Change in curvature with altitude

The curvature of the scatter plots of Figure 7.7 changes from positive at low altitudes to negative at higher altitudes in both hemispheres. This would suggest that at low altitudes, $CH_4$ has a relative sink with respect to $N_2O$ while at high altitudes the opposite is true. These observations are consistent with the $CH_4/N_2O$ lifetime ratios plotted in Figure 7.10 where it is clear that, on the whole, $CH_4$ lifetimes are shorter than those of $N_2O$ at low altitudes and longer than those of $N_2O$ at high altitudes. This shows the increasing importance of the $N_2O$ sink with respect to $CH_4$ with increasing altitude. Analysis of the slopes of the scatter plots indicates that in general, at low (high) altitudes, the upward

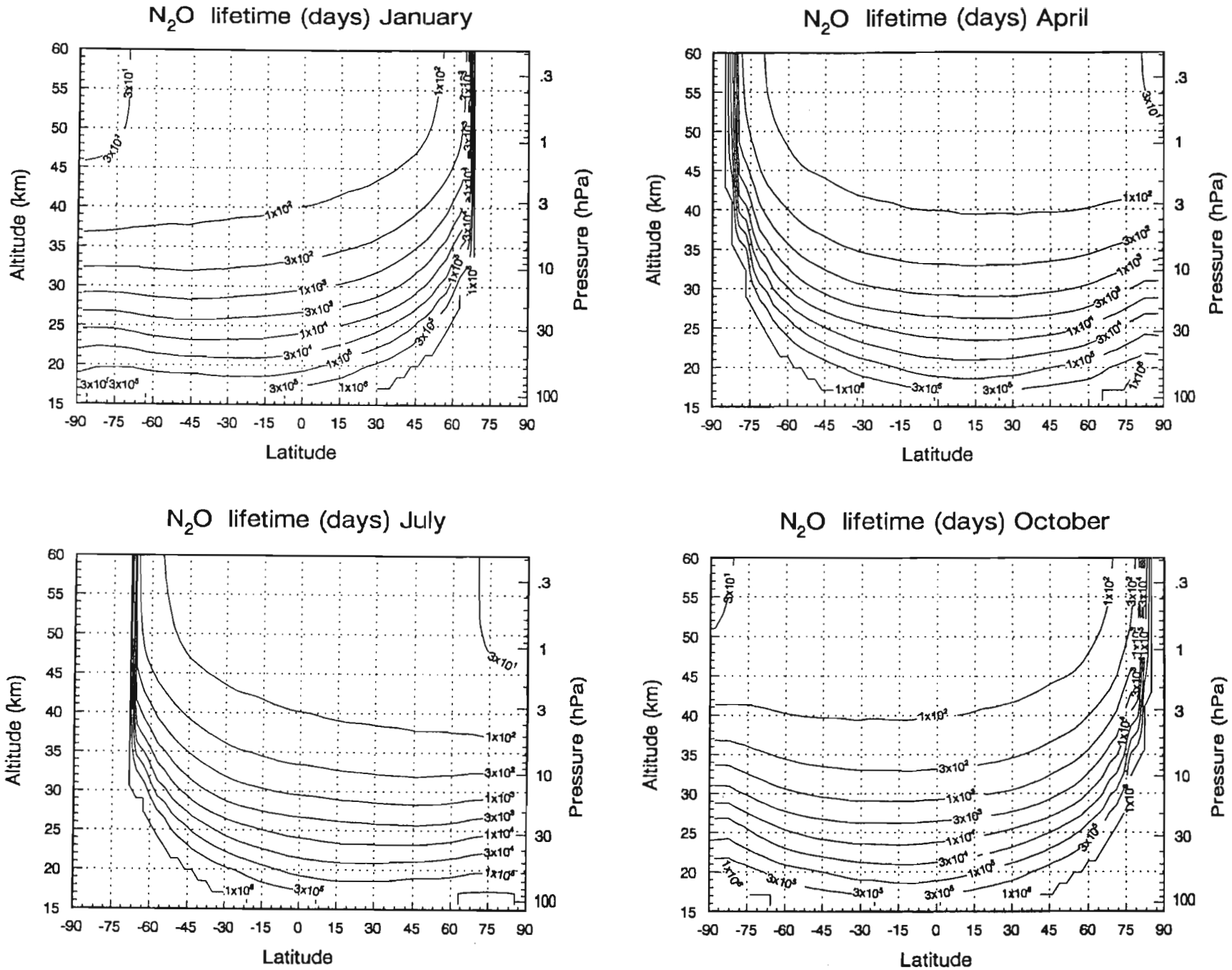


Figure 7.8: Meridional cross-sections of $N_2O$ lifetimes (in days), from the CCM2 model, for January, April, July and October.

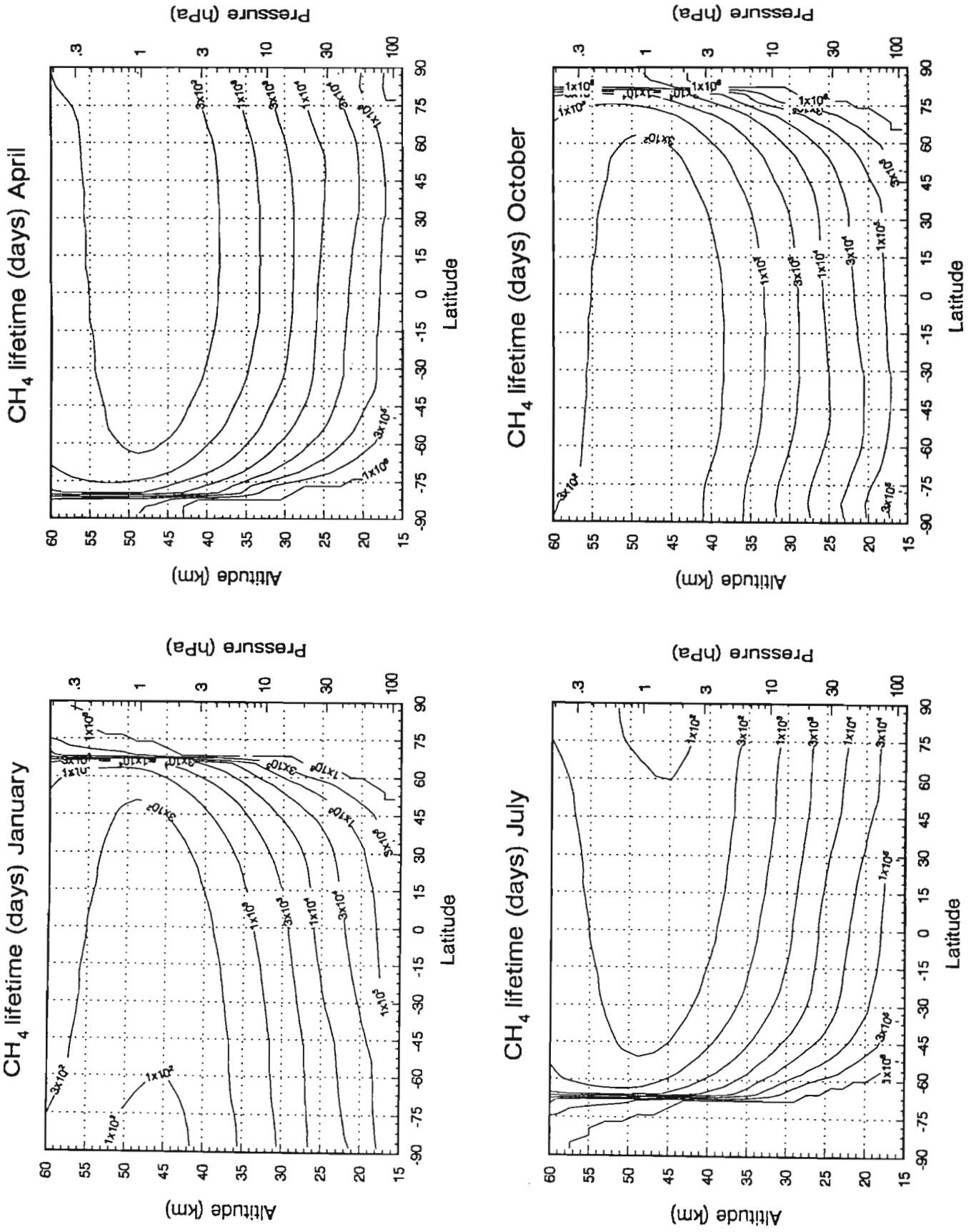


Figure 7.9: Meridional cross-sections of $CH_4$ lifetimes (in days), from the CCM2 model, for January, April, July and October.

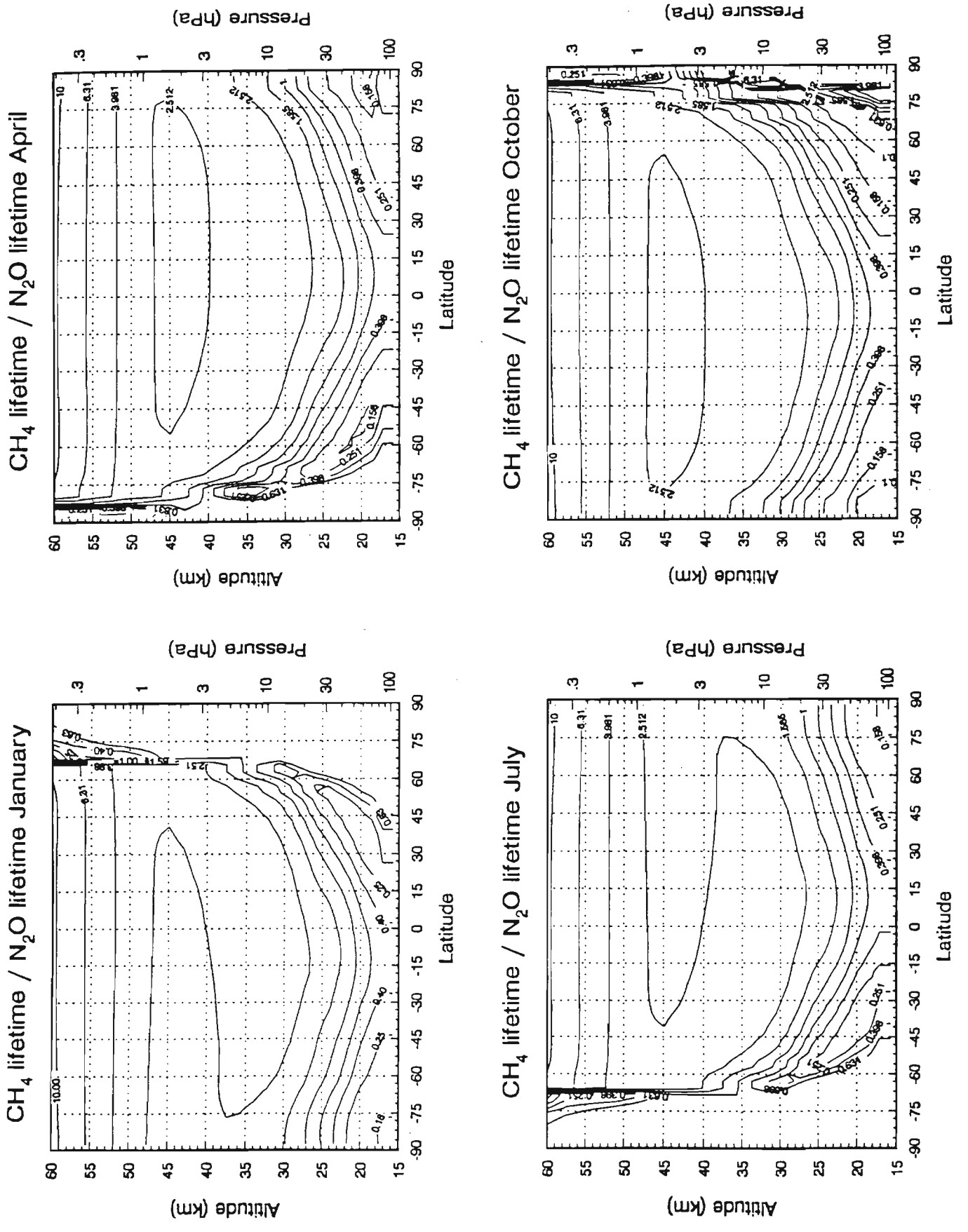


Figure 7.10: Meridional cross-sections of the ratios of the lifetime of $CH_4$ and $N_2O$ plotted in Figures 7.8 and 7.9, for January, April, July and October.

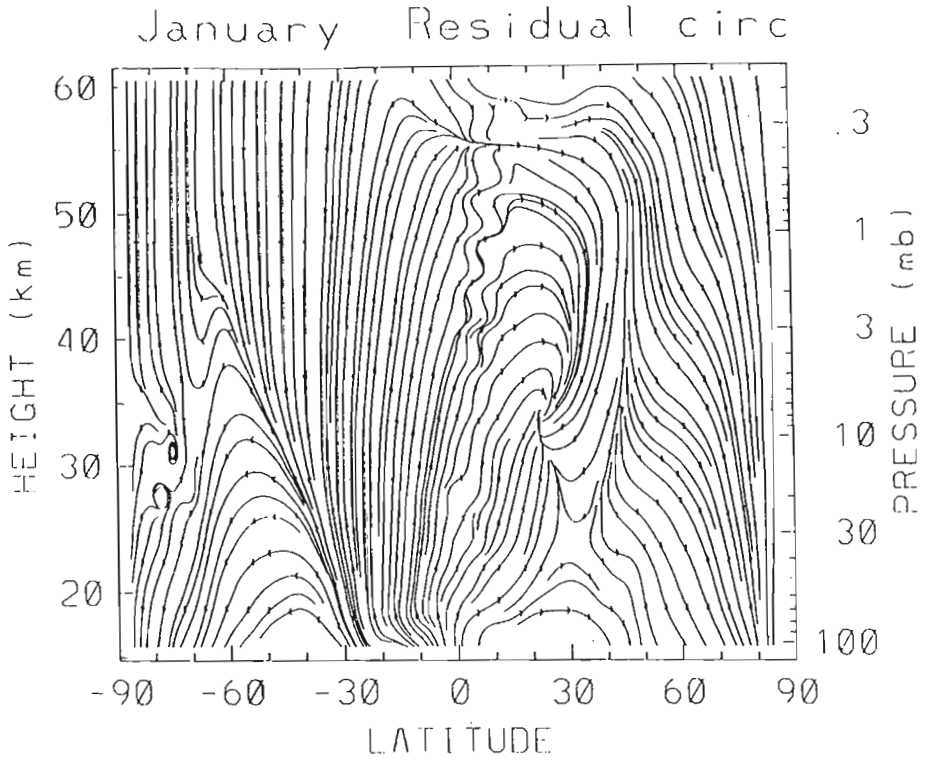


Figure 7.11: The January zonal mean residual circulation obtained from the NCAR CCM2 model.

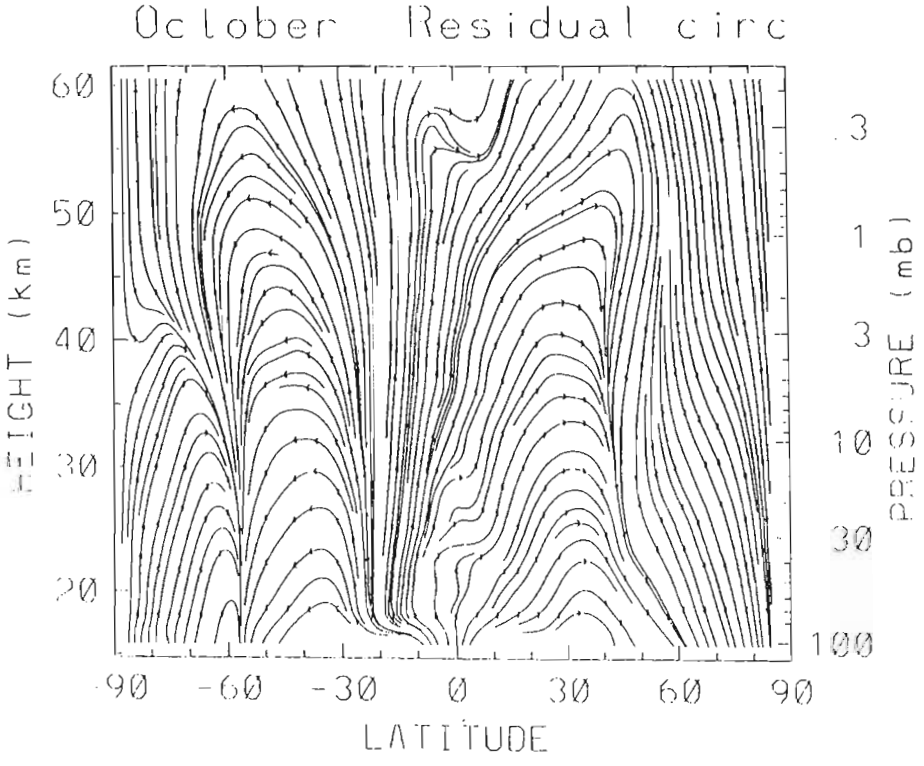


Figure 7.12: The October zonal mean residual circulation obtained from the NCAR CCM2 model.

flux of $CH_4$ with respect to that of $N_2O$ through the rapid exchange surfaces, decreases (increases) with latitude.

### Suppression of isopleths at the vortex edge

Close investigation of the scatter plots at low altitudes during the months of July and October shows that the Southern Hemisphere portion consists of a narrow parabolic shape where polar data points are separated from those at middle and low latitudes. The plot of the October residual circulation (Figure 7.12) shows strong downward motion near 60°S. This would advect $N_2O$ and $CH_4$ poor air from high altitudes to lower altitudes, suppressing the isopleths, as observed in Figures 7.4 and 7.5. The larger and stronger Southern Hemisphere vortex creates a more pronounced signature for the Antarctic winter than for the Arctic winter. These observations suggest that the scatter plots of Figure 7.7 cannot be interpreted in light of the lifetime ratios alone but must incorporate the meridional transport.

### Hemispheric and seasonal differences

During January and April, the Southern Hemisphere sections of the scatter plots are consistently lower than those of the Northern Hemisphere ie. $CH_4$ concentrations are suppressed in the Southern Hemisphere at all altitudes, or equivalently, values of $N_2O$ are suppressed in the Northern Hemisphere at all altitudes. During these months, the summer to winter circulation at high altitudes advects the long-lived $N_2O$ and $CH_4$ tracers from the Southern Hemisphere to the Northern Hemisphere. However, since the lifetime of $CH_4$ is longer than that of $N_2O$ in the transport region, preferential removal of $N_2O$ occurs and relatively smaller amounts of $N_2O$ are transported to lower levels in the Northern Hemisphere, accounting for the observed $N_2O$ suppression. This further stresses the need for an understanding of the meridional transport in the analysis these scatter plots.

### Gradient changes during the winter

At the four lowest altitude levels, the scatter plots are close to linear while above 1 hPa strong seasonally dependent curvature is evident. In particular, at 0.937 hPa during January, the gradient of the $CH_4$ versus $N_2O$ scatter plot is

noticeably steeper than that at middle and low latitudes. Recall from point 1 above that the slope of the scatter plot is equivalent to the ratio of the fluxes of the two species through the surface of constant mixing ratio. This suggests that at 0.937 hPa at high latitudes during the Northern Hemisphere winter, the upward flux of $CH_4$ compared to $N_2O$ is significantly higher than at lower latitudes. Close investigation of the lifetime ratio plot for January in Figure 7.10 shows that above 1 hPa and poleward of approximately 65°N, $CH_4$ lifetimes are lower than those for $N_2O$ while the opposite is true south of 65°N. The relative sink of $CH_4$ with respect to $N_2O$ in the upper levels of the Arctic stratosphere would imply a reduced flux of $CH_4$. Initially this appears to contradict the scatter plots. However, examination of the January residual circulation (Figure 7.11) indicates strong downward motion over the high latitudes of the Northern Hemisphere. Reduced downward flux of $CH_4$ relative to $N_2O$ would then be equivalent to an increase in upward flux of $CH_4$ relative to $N_2O$. The same arguments can be applied to the Southern Hemisphere for the month of July where the Southern Hemisphere scatter plot at 0.937 hPa shows considerable curvature. Note however that the change in slope of the scatter plot occurs between 30° and 60°S whereas that for the Northern Hemisphere occurred poleward of 60°N. This may be indicative of the hemispheric differences in circulation associated with a larger Southern Hemisphere vortex.

In summary, it appears that interpretation of the scatter plots of Figure 7.7, together with the lifetime ratio plots of Figure 7.10 is consistent with the predictions of Plumb and Ko (1992) listed as points 1 and 2 above. In some cases a knowledge of the meridional circulation is required for the interpretation of these plots.

### 7.5.2 PV versus $N_2O$ scatter plots

Prior to generating scatter plots of PV against $N_2O$, 'concentrations' for both tracers have been projected onto surfaces of constant potential temperature (isentropic surfaces) as it is generally isentropic PV that is conserved. A number of standard isentropic surfaces have been established for the analysis of data from the Upper Atmosphere Research Satellite (UARS). Five surfaces at 465, 520, 585, 840 and 1100 K have been selected and measurements of PV and $N_2O$ projected onto them.

Before displaying the scatter plots of PV versus $N_2O$, meridional cross-sections of potential temperature ($\theta$) have been plotted in Figure 7.13 to show the morphology of the isentropic surfaces at different times of the year. During the Northern Hemisphere winter, isentropic surfaces generally slope monotonically downward from the North Pole to the South Pole, although at lower stratospheric levels a local maximum is observed between 45° and 60°N. During April, the onset of radiative cooling above the Antarctic continent results in an increase in the altitude of the $\theta$ surfaces. By July very steep gradients $\theta$ gradients are observed across the Antarctic vortex boundary, Note that these gradients are significantly steeper than those observed across the Arctic vortex boundary during the Northern Hemisphere winter (see January in Figure 7.13). The local maximum at the lower levels occurs around 40°S, closer to the equator than was observed for the Northern Hemisphere. This may be indicative of the larger spatial coverage of the Antarctic vortex. By October, steep $\theta$ gradients are still observed across the vortex edge although these are shallower than those observed in the midwinter.

Figure 7.14 shows plots of zonal mean PV versus $N_2O$ concentrations for January, April, July and October on the $\theta$ surfaces at 465, 520, 585, 840 and 1100 K. The same marker convention as that of Figure 7.7 has been used to denote latitudinal position. Considering each hemisphere separately in Figure 7.14, the curves are compact and close to linear, although often only over restricted latitude ranges. In the analysis presented below, attention is restricted to the Southern Hemisphere portion of each plot since all signatures observed in the Northern Hemisphere signals also appear in the Southern Hemisphere signals and will have similar interpretations.

### Changes in the curvature of the scatter plot

Each plot of Figure 7.14 shows a similar shape with predominantly negative gradients in the Northern Hemisphere and positive gradients in the Southern Hemisphere. This is because PV decreases (becomes more negative) with altitude in the Southern Hemisphere and increases with altitude in the Northern Hemisphere while $N_2O$ concentrations fall with altitude in both hemispheres. PV values are close to zero near the equator due to the dominance of the planetary vorticity term ($f$) which varies with the sine of the latitude.

Investigation of the Southern Hemisphere data for January shows that in general,

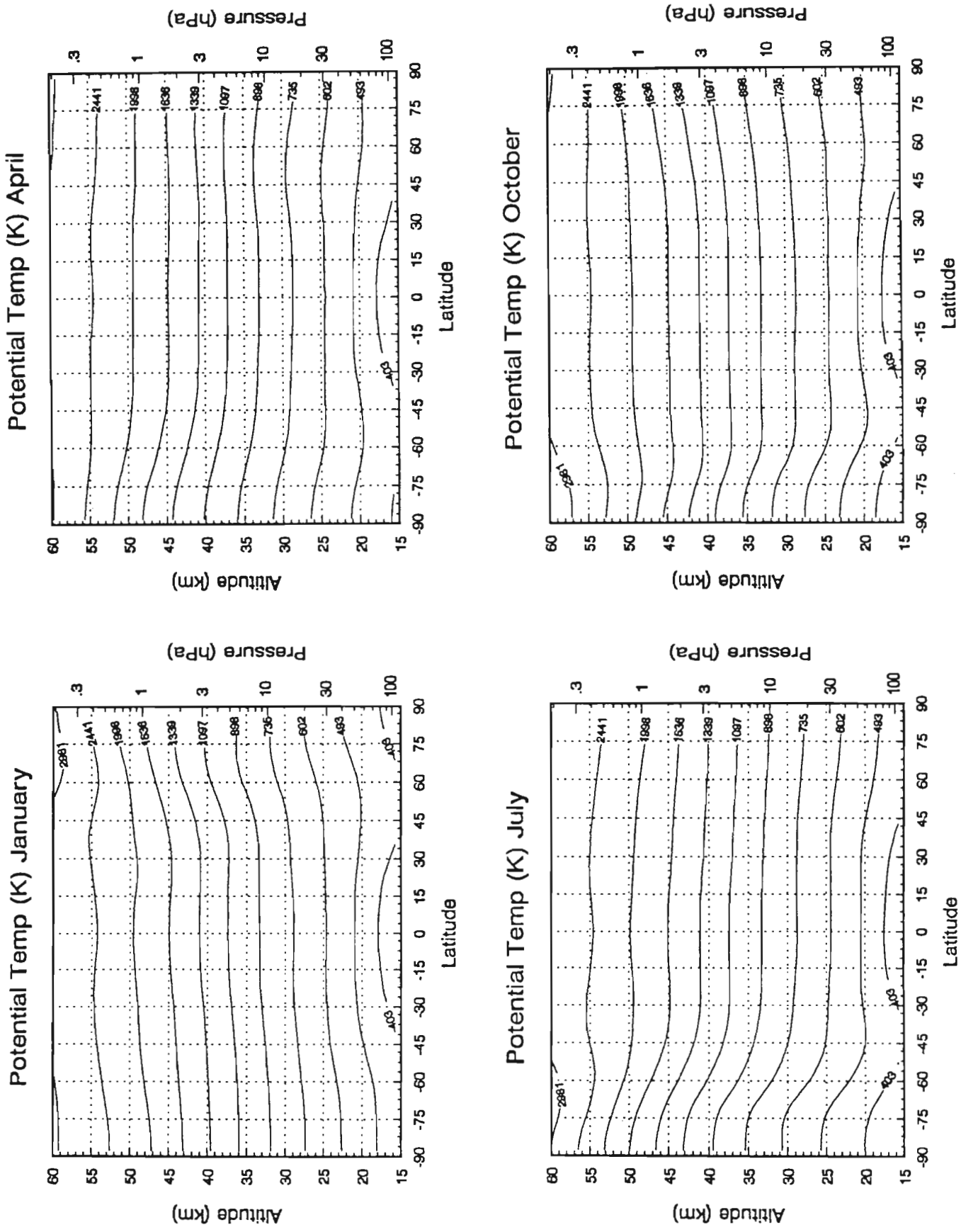


Figure 7.13: Meridional cross-sections of the zonal mean potential temperature averaged for January, April, July and October. Temperature and pressure data used to create these plots were extracted from the CCM2 model run.

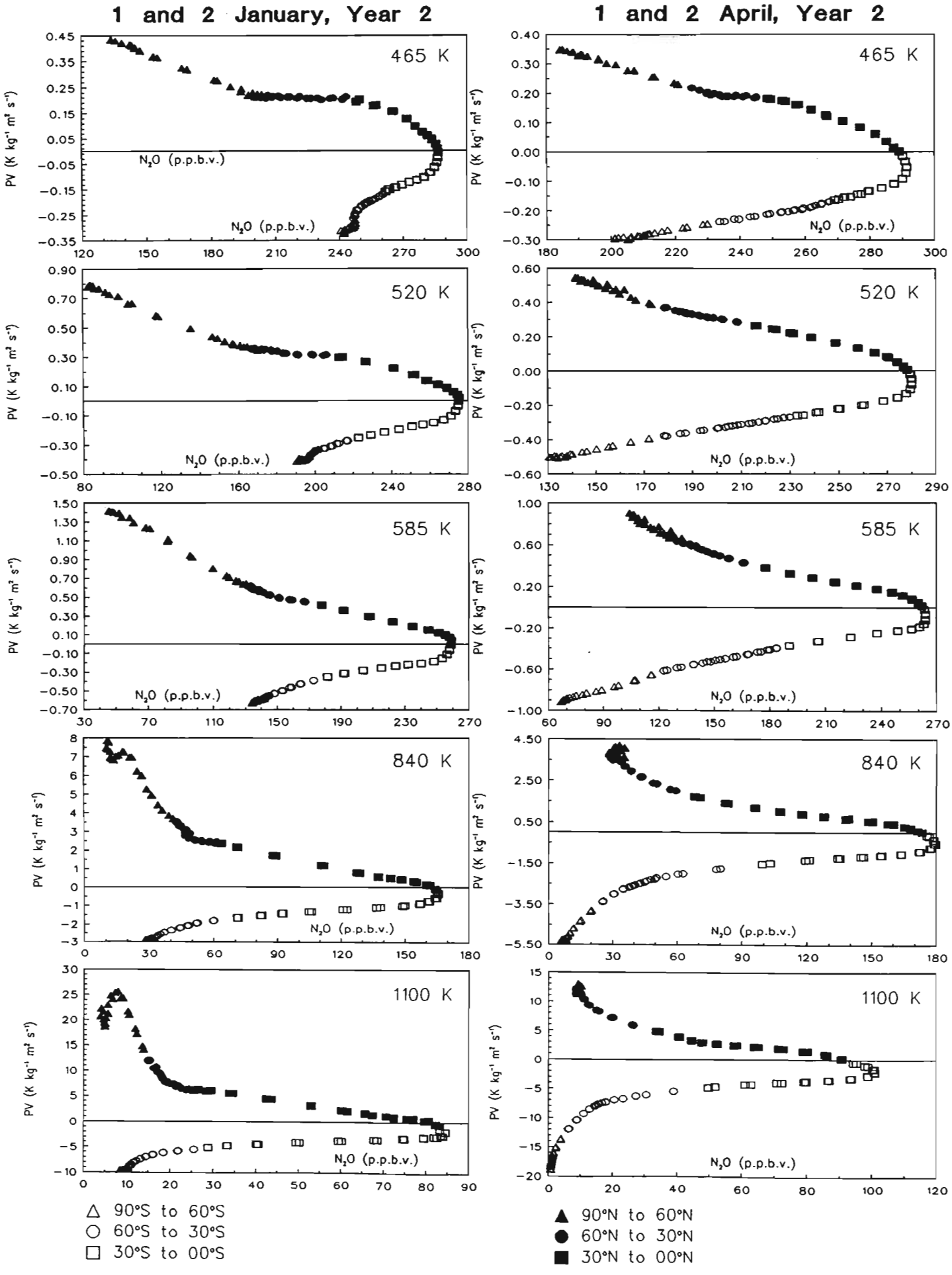


Figure 7.14: Scatter plots of PV (in $K\,kg^{-1}\,m^2\,s^{-1}$) versus $N_2O$ concentrations (in p.p.b.v.) at 465, 520, 585, 840, 1100 K taken from the CCM2 model. Data for the first two days of January and April are shown.

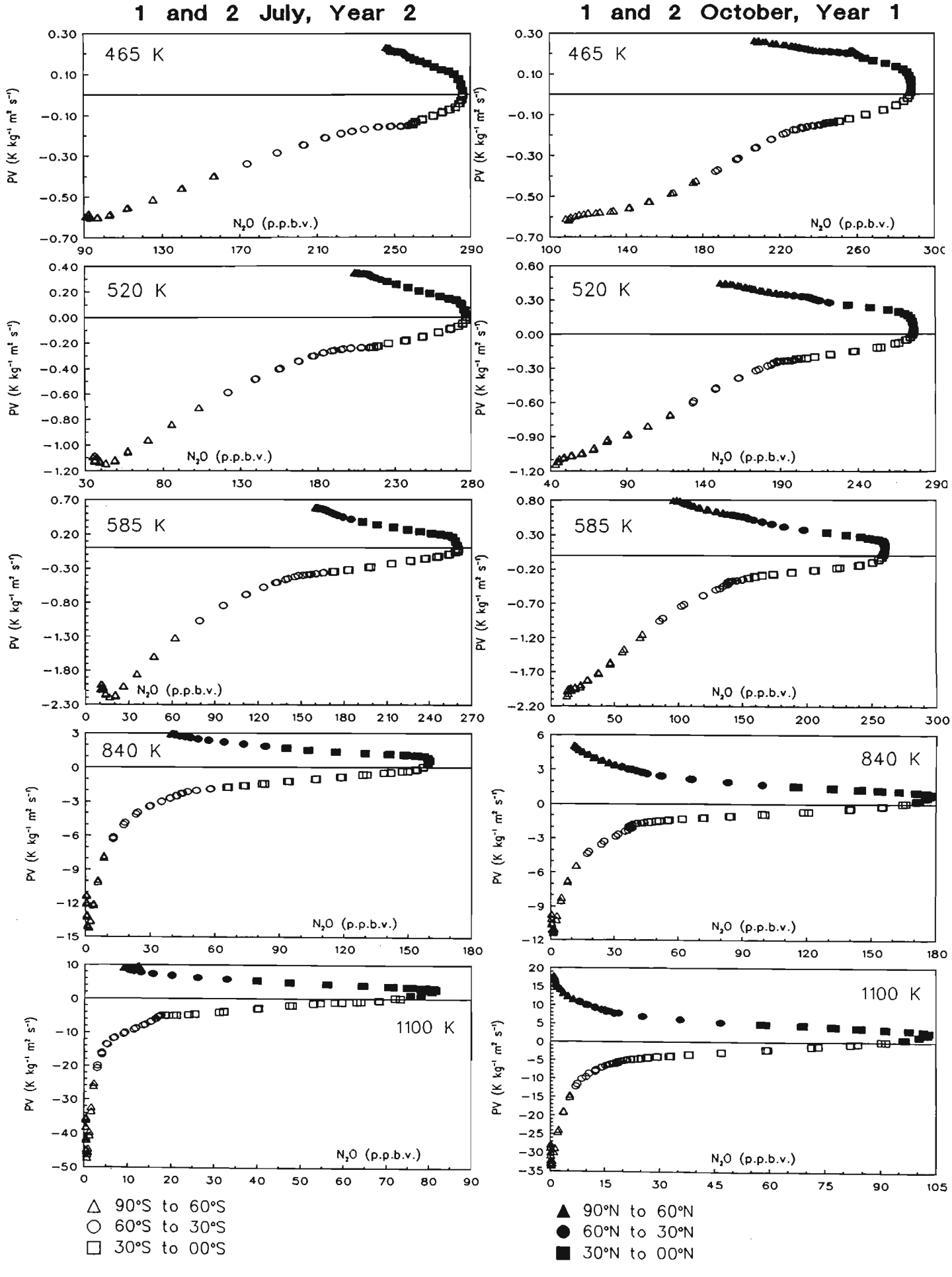


Figure 7.14 cont.: Scatter plots of PV (in $K\,kg^{-1}\,m^2\,s^{-1}$) versus $N_2O$ concentrations (in p.p.b.v.) at 465, 520, 585, 840, 1100 K taken from the CCM2 model. Data for the first two days of July and October are shown.

the scatter plots have negative curvature with PV/$N_2O$ gradients (in the scatter plot) increasing with latitude. This suggests that $N_2O$ has a relative sink with respect to PV at all altitudes. However, this may not be interpreted as PV lifetimes being longer than those of $N_2O$ for two reasons:

- The formalism of Plumb and Ko stresses that the slope of the scatter plot may only be related to the lifetimes of the two species if the scatter plot is linear. The plots of Figure 7.14 are linear only over limited latitude ranges and suggest that it may not be possible to make estimates of PV lifetimes over certain regions using the analysis technique of Plumb and Ko. These difficulties are discussed further below.

- The PV values are dominated by the planetary vorticity ($f$) which has an infinite lifetime. It is suggested that future studies examine only the variation of relative vorticity with respect to $N_2O$ to obtain a measure of the lifetime of PV.

**Behaviour at high latitudes and high altitudes**

A number of small features are evident in the high latitudes of the Southern Hemisphere at high altitude. The course scale on the axes of the plots obscures the morphology of the features. For this reason, plots of PV against $N_2O$ at 1100 K are shown at higher resolution in Figure 7.15. Clear local minima can be seen in the scatter plots for October and even more so for July. These signatures may be indicative of the isolation of the Antarctic vortex. Plumb and Ko (1992) have shown that such signatures correspond to a breakdown in gradient equilibrium and stated that this occurs if the local lifetime of any one tracer is not very long. Since the lifetime of $N_2O$ is very high in the high latitudes of the Southern Hemisphere winter (see Figure 7.8), this can only mean that the PV lifetimes are not sufficiently long for the application of the theory of Plumb and Ko. Previous studies (Newman, 1986) have stated that PV may not be conserved over the entire Southern Hemisphere because of diabatic processes. It is therefore possible that the strong diabatic cooling during the winter months reduces the lifetime of PV.

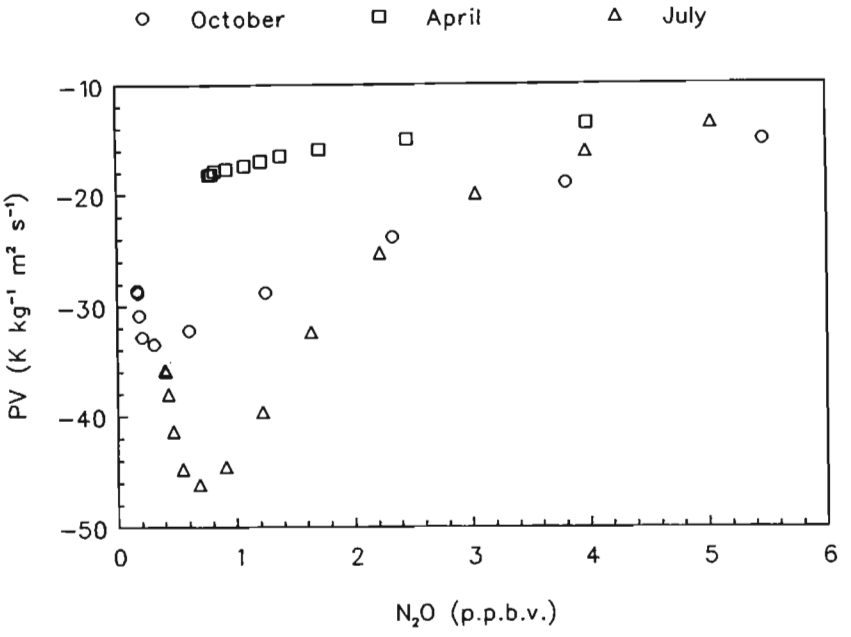


Figure 7.15: Plots of PV versus $N_2O$ at the 1100 K surface for the months of October, April and July.

## 7.6 Discussion

A number of previous studies have made use of models to examine the effects global distributions of long-lived stratospheric tracers (Levy et al., 1979; Garcia et al., 1992). Figure 7.16 shows plots of modelled global distributions of nitrous oxide for January and July obtained form the model of Garcia et al. (1992). The contour superimposed on the $N_2O$ distribution encloses the regions where the diffusion coefficient due to planetary wave breaking exceeds $10^6\,\mathrm{m\,s^{-1}}$. It is clear that planetary wave breaking occurs predominantly in the winter hemisphere and results in flattening of the $N_2O$ isopleths (cf. Figure 7.4). The modelled mean circulation of Garcia et al. (1992) was driven by the dissipation of planetary waves both through breaking and thermal relaxation. It was shown that these process produce the mean circulation through most of the year in the extratropical stratosphere (below approximately 25 km).

The analysis of $CH_4$ versus $N_2O$ scatter plots has shown that the distribution of these two species within the CCM2 supports the predictions of Plumb and Ko (1992). This provides confidence in the analysis of the PV versus $N_2O$ plots. A preliminary investigation of these plots suggests that PV may not be well conserved in the high latitudes of the Southern Hemisphere winter. This has important implications for airborne measurements of trace gases as mentioned

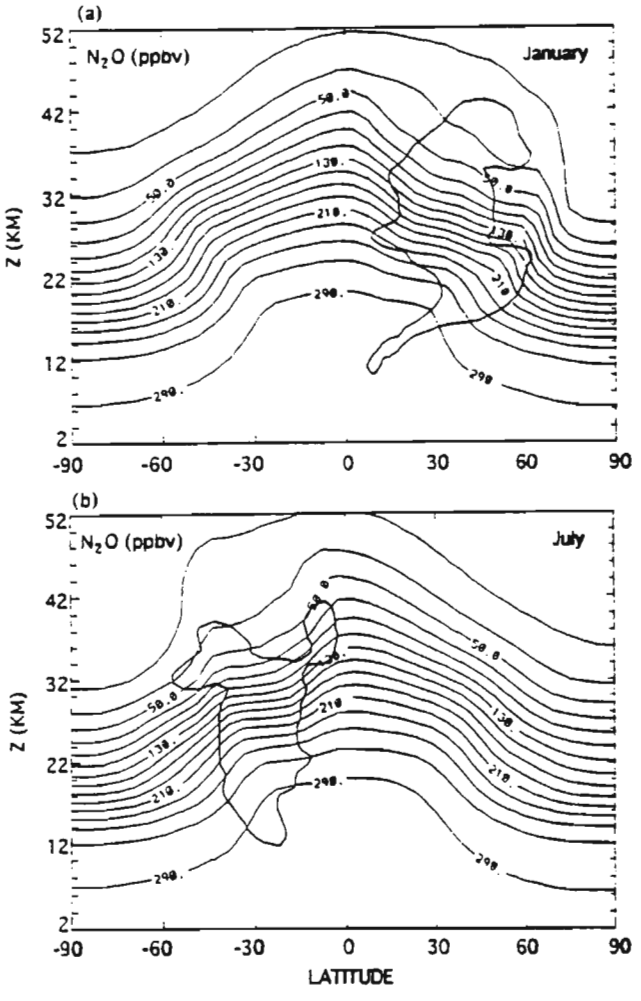


Figure 7.16: Global distributions of $N_2O$ (p.p.b.v.) computed from the model of Garcia et al. (1992).

above.

Scatter plots similar to those of Figure 7.14 require extensive further analysis to obtain a measure of the 'lifetime' of PV at different times of the year and at different altitudes, if at all possible. To use the Plumb and Ko formalism to obtain information on the lifetime of PV from scatter plots such as those of Figure 7.14, it is necessary to find regimes where the scatter plots are linear. Furthermore, the theory of Plumb and Ko requires that the two species be in climatological slope equilibrium. It is currently uncertain as to whether PV meets this requirement. It is necessary that the lifetime of PV must be determined by 'removal' in the stratosphere in order to make use of the slopes of the scatter plots to obtain an estimate of PV lifetime. Although PV values do decrease with altitude in the Southern Hemisphere, this is purely a result of the sign convention

for latitude, and not suggestive of a stratospheric PV sink. Northern Hemisphere PV values increase with altitude, but again this does not imply that PV has a stratospheric source. Finally, it may not be appropriate to make use of zonal means of PV in this analysis, as temporal and spatial averaging will tend to smooth meridional gradients, as discussed above. A preliminary investigation of this problem indicated that the differences between zonal mean and discrete longitude scatter plots were minimal. However, a more complete investigation of this issue should be conducted. These caveats impose a number of restrictions on the use of the Plumb and Ko technique. Future studies must first determine whether PV adheres to the requirements of the Plumb and Ko theory before an attempt is made to derive PV lifetimes from scatter plots of PV against long-lived tracers.

# Chapter 8

# CONCLUSIONS

The work presented in this thesis has shown that planetary waves significantly influence the global distribution of total column ozone and other long-lived trace species. Theoretical developments of Chapter 2 showed how planetary waves may be generated by topographic forcing and land-sea heating contrasts (section 2.6.4). The derivation of Rossby wave propagation equations showed that planetary waves are restricted to regions of westerly flow, the strength of which determines the maximum allowed wavenumber (section 2.8.3).

The studies described in sections 4.3, 4.4 and 4.5 revealed a number of important characteristics of Southern Hemisphere wave behaviour viz.:

- Wave power decreases with increasing wavenumber throughout the year. Power for waves 1 and 2 maximizes during the Antarctic vortex period (AVP) while powers at wavenumber 3 and 4 may be reduced during this period as a result of the critical velocity condition (equation 2.136).
- The AVP depleted ozone mass over the Antarctic appears to be increasing in response to the rise in stratospheric chlorine loading. Over the period 1979 to 1992, stratospheric chlorine levels rose from approximately 2.55 p.p.b.v to 4.02 p.p.b.v. while the Antarctic AVP mean depleted ozone mass increased from close to zero in 1979 to 9.08 Mt in 1992. The high interannual variability in the depleted mass signal was shown to be related to interannual differences in planetary wave activity.
- Differences in AVP mean SANAE temperatures were also shown to be well correlated with midlatitude planetary wave activity (section 4.4.3).

- Manney et al. (1991), referred to as MFM, showed 1987 to be a year of extremely low wave 1 activity over the period 1979 to 1988. Analysis of TOMS data confirmed that this was the year of second lowest total wave power (exceeded only marginally by 1980) over the period 1979 to 1992. It was also shown in section 4.3 that 1987 was the year of most severe Antarctic ozone depletion when normalized to the stratospheric chlorine loading.

- MFM showed that in the 10 years following 1979, 1988 was the only year when the wave 1 amplitudes at 10 hPa exceeded 1400 m, reaching amplitudes greater than 1800 m. Fourier analysis of TOMS data from 1979 to 1992 confirmed 1988 to be the year of highest total wave power and in addition showed this to be the year of smallest Antarctic ozone depletion when normalized to the stratospheric chlorine loading.

- Both the work of MFM and the analysis of wave structures observed in the TOMS data showed that wave 1 is quasi-stationary with occasional wave propagation events (section 4.4.1).

- The 8 most prominent wave 1 propagation events, from July to October of 1979 to 1988, extracted from the TOMS data, were shown to have eastward zonal wave speeds varying between $1.5\,\mathrm{m\,s^{-1}}$ and $14\,\mathrm{m\,s^{-1}}$.

- Both MFM and the analysis of total ozone waves in TOMS distributions showed that eastward wave 2 propagation events occur regularly from July to October over the period 1979 to 1988 (section 4.4.2).

- Analysis of the 8 most prominent wave 2 propagation events observed in the TOMS data indicated eastward wave speeds of between $3\,\mathrm{m\,s^{-1}}$ and $11\,\mathrm{m\,s^{-1}}$.

- The latitudinal positions of the wave maxima for both wave 1 and 2 were seen to increase ie. move poleward over the period from July to October.

- When examining the winter total column ozone and upper air temperature data from an Antarctic site, it is important to consider both the amplitude and phase of planetary wave forcing (section 4.5). The weak results obtained may be indicative of other forcings besides planetary waves.

The analysis of Northern Hemisphere wave powers described in section 4.6 showed that:

- Wave powers in the Northern Hemisphere are greater than those in the Southern Hemisphere for wavenumbers 2 to 4. It appears that the peak powers for wavenumber 1 occur in the Southern Hemisphere.
- Arctic ozone depletion appears to be increasing with the rise in stratospheric chlorine loading. The AVP mean depleted mass rose from 6.9 kt in 1979 to 52.3 kt in 1992.
- The time averaged winter ozone depletion (poleward of 40°) is about 200 times greater over the Antarctic than over the Arctic.
- The AVP mean depleted ozone mass over the Arctic shows little or no correlation with Northern Hemisphere planetary wave activity.

In Chapter 5, the development of a statistical model was discussed to show how planetary waves may affect the error in the prediction of high latitude total column ozone. It was shown that for Antarctic regions (latitude $> 60°S$), if a measure of the early winter wave activity can be derived (either from TOMS ozone distributions, or ECMWF or NMC geopotential height fields), it may be possible to predict the severity of the Antarctic ozone depletion a few months in advance. Derived wave power signals could also be used to partially remove the effects of interannual variability in wave activity when calculating high latitude long-term linear trends from statistical models. A similar technique could not be applied in the high latitudes of the Northern Hemisphere due to the weak dependence of Arctic ozone depletion on wave forcing. Coefficients derived from the statistical model often showed marked zonal asymmetry. These features were usually associated with the formation of a large winter wave 1 in the midlatitudes of the Southern Hemisphere. The high latitudinal resolution of the model also emphasized important fine structure in the meridional distribution of some model coefficients eg. the amplitude of the QBO. Many previous statistical models had smoothed such features by averaging over a fews degrees of latitude.

Analysis of long-term trends derived from the model showed that significant ozone depletion occurred at midlatitudes over the period 1979 to 1992 in addition to polar depletion. These results were confirmed in a comparison with data obtained from Herman et al. (1993). The equatorial QBO in total column ozone was shown

to be half a QBO period out of phase with signals at middle and high latitudes. Over the latitude bands between 10° and 15° , where phase values showed steep gradients, QBO amplitude were close to zero. These observations were supported by comparison with model data obtained from Gray and Pyle (1989).

In Chapter 6 the statistical model was applied to long-term trend determination of total column ozone above 5 South African cities. Model coefficients were used to predict monthly mean total column ozone above each city to the year 2000. When treated as inputs to a global surface irradiance model, it was then possible to predict surface erythemal irradiances to the year 2000. This analysis suggested that the summers of 1997/1998 and 1999/2000 may have above average erythemal irradiance levels for all 5 cities.

The use of CCM2 model data in Chapter 7 showed that long-lived tracer concentrations from the model supported the hypotheses of Plumb and Ko (1992). Scatter plots of potential vorticity against $N_2O$ were used to suggest a technique whereby PV 'lifetimes' could be inferred at different times of the year and at different altitudes and latitudes. Difficulties were experienced with this analysis and suggestions were made for future work in this field. In particular, meridional cross-sections of PV at discrete longitudes must be examined to see if these data conform to the requirements of the Plumb and Ko formalism.

# Appendix A

# CHARACTERIZATION OF WAVE STRUCTURES IN TOMS IMAGES

## A.1 Space-time Fourier Transform methods

The temporal evolution of spatial wave structures observed in a satellite mapped data distribution may be analyzed using a two dimensional space-time Fourier transform (Lait and Stanford, 1988). Data along one line of latitude, or averaged over a number of lines of latitude, are plotted against time to provide the input data field, denoted $\psi(\lambda, t)$ where $\lambda$ is the longitude and $t$ is the time. The discrete zonal space-time Fourier transform of the observed data field, sampled at regular intervals in $\lambda$ and $t$ is then given by:

$$\Psi(m, \sigma) = \sum_{j=0}^{N-1} \sum_{k=0}^{M-1} \psi(\lambda_j, t_k) \exp[-i(m\lambda_j + \sigma t_k)] W \qquad \text{(A.1)}$$

where $\lambda_j$ is the $j$th sampled longitude (N=287 for TOMS plots), $t_k$ is the $k$th sampled time (the day number in this case), $m$ is the zonal wavenumber, $\sigma$ is the angular frequency (in radians per day), and $W$ is a weighting factor dependent on the size, shape and spacing of the data grid. There are two inherent disadvantages in applying this technique:

1) Nonuniform temporal spacing between data points

Consider a polar orbiting satellite making soundings on both the ascending and

descending orbits. Other than at the equator, the time differences between consecutive measurements at the same latitude will differ. In the Southern Hemisphere the time difference between a measurement on the ascending orbit and the preceding measurement will be smaller than the time difference between a measurements on the descending orbit and the preceding measurement since the transit time over the Antarctic will be shorter. This nonuniform temporal spacing between ascending and descending data points results in aliasing of spectral peaks in the output Fourier Transform (Lait and Stanford, 1988). However, since TOMS requires backscattered sunlight to make measurements (see Chapter 3), only data points along the ascending daytime orbit are collected such that the temporal spacing between measurements at the same latitude is uniform. Other than during periods of perpetual polar daylight where TOMS may make measurements on both ascending and descending orbits, this apparent restriction may be ignored.

2) Asynoptic satellite sampling

For computational efficiency it would be advantageous to separate the double sum in equation A.1 into two single summations. However, polar orbiting satellites do not measure global distributions synoptically (as a 'snapshot') but rather asynoptically where the longitude of the measurement ($\lambda$) depends on time ($t$). The time difference between the TOMS measurements at the extreme west (longitude cell number 1) typically occurs 24 hours after the measurement at the extreme east (longitude cell number 288). This dependence of $\lambda$ on $t$ means that the double sum in equation A.1 can no longer be identified with separate zonal and temporal transforms. Lait and Stanford (1988) have shown that this restriction may be avoided by transforming the ($\lambda$, t) coordinate system to a new one whereby the double summation does become separable. It can be shown that the coordinate transform required is simply a rotation about the origin through an angle given by:

$$\tan \alpha = \frac{1}{c_0} \tag{A.2}$$

where $c_0$ is the zonal speed of the satellite, which for sun-synchronous satellites such as TOMS, is $2\pi$ radians per day. Once the coordinate transform has been performed, the Fourier summation is separated and computed. The inverse coordinate transformation is then performed to express the derived Fourier components in ($m$, $\sigma$) coordinates.

Using this technique to process the TOMS data requires an array of 288 × 365 (or 366 for leap years) data points for each year from 1979 to 1992. To process the 14 years of data in this way would be computationally prohibitive. One solution is to reduce the 2 dimensional Fourier transform to a 1 dimensional transform by considering only spatial structure. The salient features of the spatial structure (such as the power at the fundamental wavenumbers) may then be analyzed as a function of time. However, when computing the spatial Fourier transform the problem of asynoptic sampling remains. Recall that the problem of nonuniform temporal spacing between data points does not apply to TOMS since it makes measurements only during dayside passes. The asynoptic sampling problem may be treated in two ways:

- Since the TOMS measurements are made continuously through time, the daily TOMS data plots should not be viewed as discrete data sets. The beginning of a daily total ozone distribution (longitude cell 288) follows smoothly from the end (longitude cell 1) of the previous day's data. If temporal variations over a single ascending orbit are ignored, the temporal and spatial position of a TOMS measurement may be plotted in cylindrical coordinates with time as the vertical coordinate and latitude and longitude as the azimuthal and radial coordinates respectively. In this 3 dimensional representation, a number of consecutive daily hemispheric TOMS plots defines a spiral shown schematically in Figure A.1. Different days are shown in different colours where the end of each day of data tags onto the beginning of the next. The time axis is in the positive z direction. To obtain a synoptic 'snapshot' of the total column ozone distribution, the projection of the spiral onto a plane perpendicular to the time axis is calculated. This requires longitudinally varying interpolation between consecutive data sets.

- Show that the Doppler shifting of waves propagating at typical wave speeds does not significantly affect the wave power computed from the Fourier transform. These tests are shown in detail below.

## A.2 Spatial Fourier Transforms

Daily measures of midlatitude wave power may be obtained by passing a representative zonal total column ozone profile through a Fourier transform and

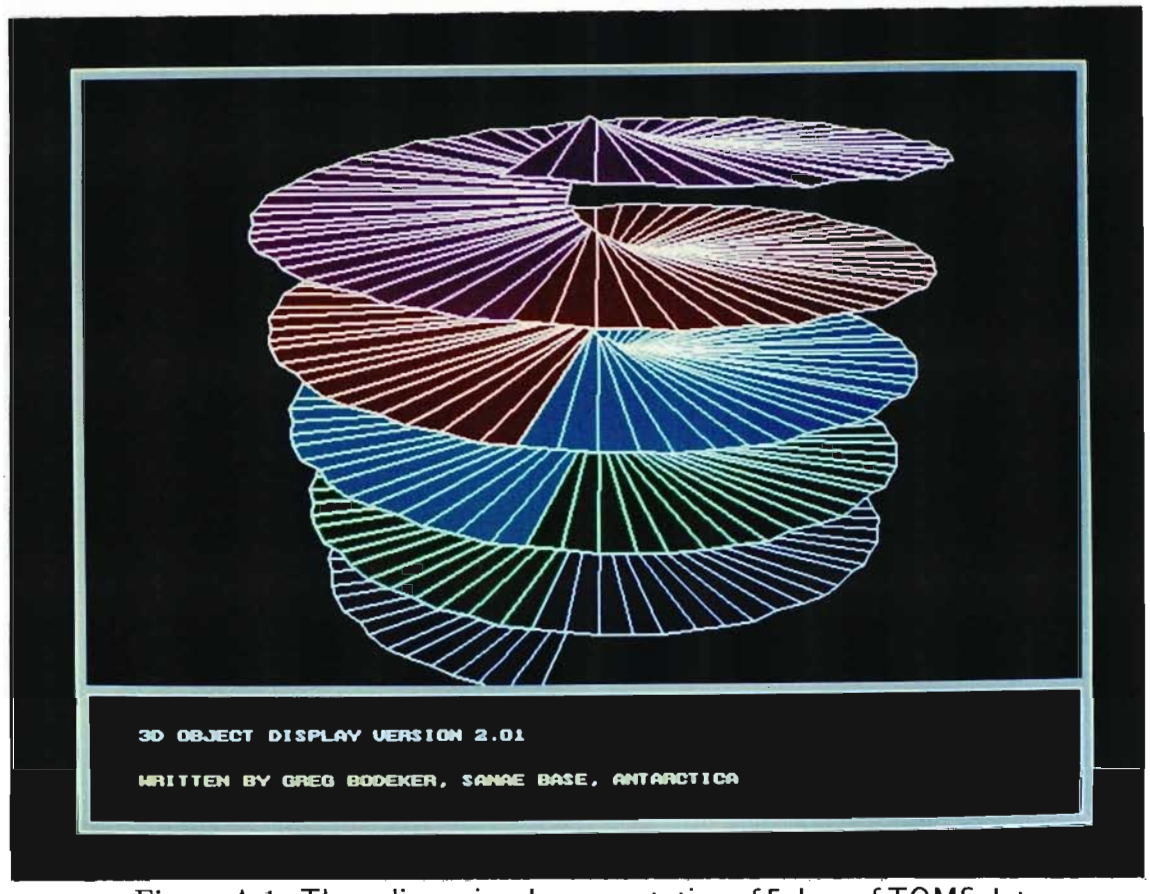


Figure A.1: Three dimensional representation of 5 days of TOMS data. The time axis is vertical, latitude is radial and longitude is azimuthal.

obtaining the power spectrum. The midlatitude representative profile is typically calculated by averaging zonal profiles between 50°and 75° latitude. A number of different Fourier transform techniques are available. A first approach may be to simply apply a discrete Fourier transform to each daily data series. However, the algorithm for this transform is computationally demanding since its speed is of the order $N^2$, where N, the number of data points, is 288 for the TOMS data. In order to reduce the processing time, a Fast Fourier Transform (FFT) algorithm of speed order (N log N) may be used. However, since the FFT algorithm requires that the number of data points be an integral power of 2, the TOMS data series require extensive zero padding to bring the total number of data points to 512 ($2^9$) and make them amenable to FFT analysis. Furthermore Hanning windows must be applied to the edges of the data series to match the data signal with the zero padding. This windowing, together to the zero padding, may lead to loss of resolution within the FFT as discussed later. A final approach may be to

take means over 9 cells within the 288 point data series to obtain a low spatial resolution 32 data point profile. This profile is immediately amenable to FFT analysis, requiring no zero padding or windowing. These three different Fourier transforms,

1. The discrete Fourier transform.
2. The high resolution (512 data points) FFT.
3. The low resolution (32 data point) FFT.

are each tested and discussed in detail below. The algorithms for these Fourier transforms have been adapted from Press et al. (1989) and Brigham (1974).

### A.2.1 Test waveforms

Sixteen artificial waveforms incorporating changes in amplitude, phase and period have been generated to test each of the Fourier algorithms. Furthermore, to investigate the effects of asynoptic sampling of the total column ozone by TOMS, four test waveforms have been generated to show the resultant asynoptically sampled zonal profile of waves travelling eastward at 0°, 10°, 20° and 30° longitude per day. Only eastward propagating waves were considered since preliminary studies using actual TOMS data indicated predominantly eastward wave propagation. The resultant test waveforms are shown in Figure A.2. The four waveforms in panel **a** show variations of amplitude with zero phase and wavenumber 1 period. The four wave 1 signals in panel **b** are of constant 100 DU amplitude and show variations in phase. The four waveforms in panel **c** are of constant amplitude and zero phase, showing variations in wavenumber from 1 to 4. The waveforms of panel **d** show the affect of asynoptic sampling of total column ozone waves propagating at 0°, 10°, 20° and 30° longitude per day.

### A.2.2 The discrete Fourier transform

Consider the discrete Fourier transform of a series of 288 data points. We define the discrete function:

$$h_k \equiv h(\lambda_k) \tag{A.3}$$

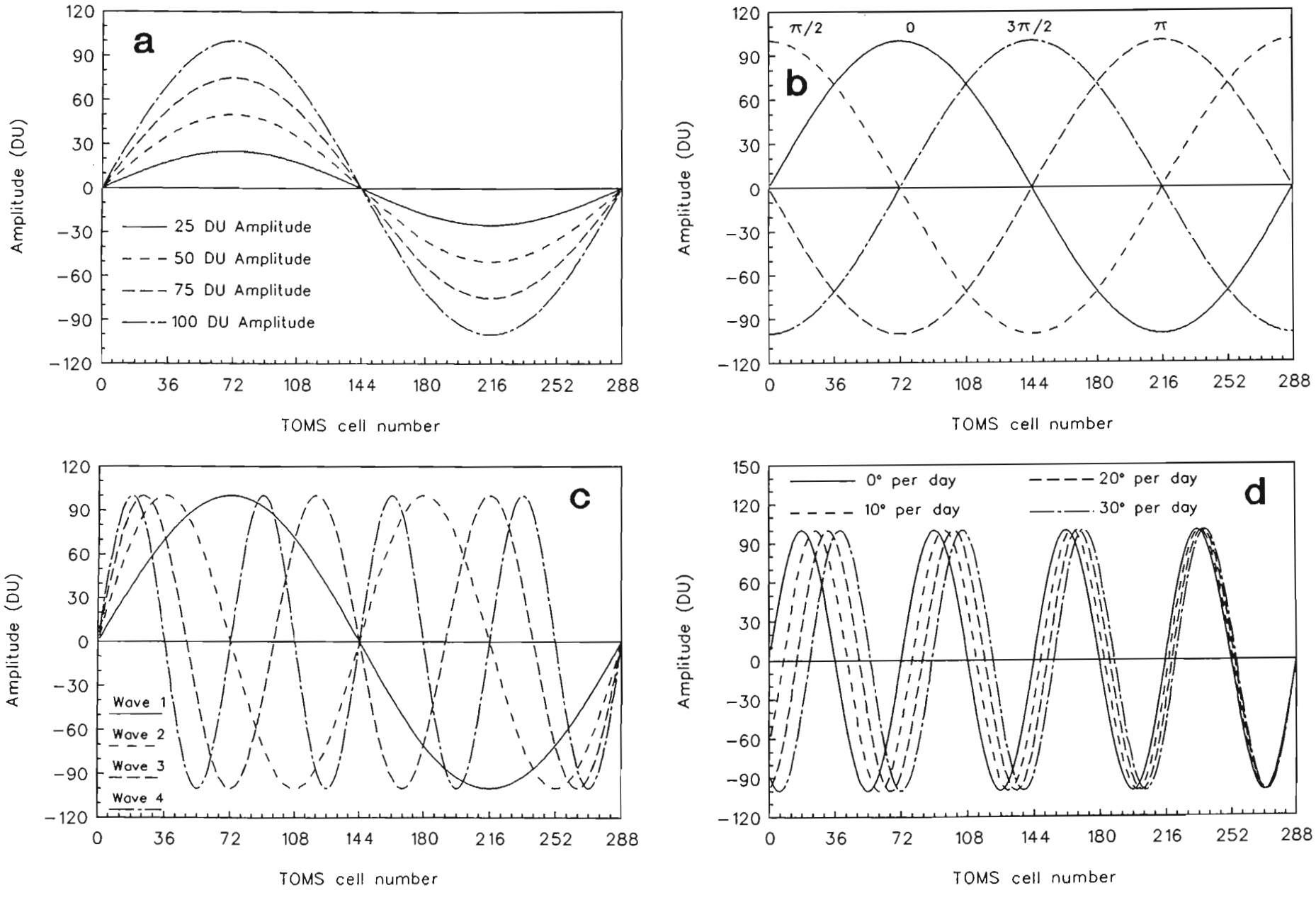


Figure A.2: The sixteen artificial zonal ozone profiles used to test the Fourier transforms.

where $\lambda_k$ is the $k$th longitude defined as $k\Delta$ where k runs from 0 to 287 and $\Delta$, the sampling interval, is 1 TOMS cell. The Fourier integral is approximated as:

$$H(f_n) = \int_{-\infty}^{\infty} h(\lambda_k) e^{2\pi i f_n \lambda} d\lambda \approx \sum_{k=0}^{287} h_k e^{2\pi i f_n \lambda_k} \Delta = \Delta \sum_{k=0}^{287} h_k e^{2\pi i k n/N} \qquad \text{(A.4)}$$

where $f_n$, the discrete spatial frequency (measured in cycles per TOMS cell), is defined as:

$$f_n \equiv \frac{n}{N\Delta}, \qquad n = -\frac{N}{2}, \cdots, \frac{N}{2} \qquad \text{(A.5)}$$

The discrete Fourier transform of the $h_k$ data points is then denoted by:

$$H_n \equiv \sum_{k=0}^{287} h_k \, e^{2\pi i k n/N} \qquad \text{(A.6)}$$

This discrete Fourier transform maps the 288 complex numbers, $h_k$, into the 288 complex numbers, $H_n$. The power in a narrow frequency interval centered at frequency $f_n$ is known as the power spectral density (PSD). When the input data are real as in the case of TOMS data, this power may be defined as:

$$\text{Power}(\text{f}_\text{n}) = \frac{1}{N^2}[\mid H_n \mid^2 + \mid H_{N-n} \mid^2] \qquad k = 1, 2, \cdots, (\frac{N}{2} - 1) \qquad \text{(A.7)}$$

where $n = f_n N\Delta$ and $\mid H_n \mid^2$, the modulus of $H_n$, is defined as:

$$\mid H_n \mid^2 = H_n H_n^* \qquad \text{(A.8)}$$

where $H_n^*$ is the complex conjugate of $H_n$.

The pascal code written to implement this algorithm is listed below.

```
const
  MaxMathArray=1024;

type
  TMathArray=array[1..MaxMathArray] of real;
  PMathArray=^TMathArray;

procedure DFT(var DFTArray:PMathArray;N:integer);

{=============================================================================
{ This procedure performs a Discrete Fourier Transform (DFT) on the data
{ passed via the DFTArray. Note that only the odd numbered positions in this
```

```
{ array contain data since the imaginary parts of the original data points ar
{ zero. The N parameter is the number of original data points which is half
{ number of elements in DFTArray. The DFTArray is replaced by its discrete
{ Fourier transform. In this case N need not be a power of 2 but must be even
{-------------------------------------------------------------------------

var
  I,J:integer;        {Loop control variable.}
  RealComp,ImagComp, {The real and imaginary components within the summation.
  RealTot,ImagTot,    {The totals of the real and imaginary components.}
  T:real;             {The argument for sine and cosine functions.}
  Hold:PMathArray;    {Temporary heap storage array.}
begin
  getmem(Hold,(N+1)*2*sizeof(real));
  for I:=(-1*(N div 2)) to (N div 2) do
  begin                {Primary loop start}
    RealTot:=0.0;
    ImagTot:=0.0;
    for J:=0 to N-1 do
    begin              {Secondary loop start}
      T:=2*pi*J*I/N;
      RealTot:=RealTot+((DFTArray^[J*2+1]*cos(T))-(DFTArray^[J*2+2]*sin(T)));
      ImagTot:=ImagTot+((DFTArray^[J*2+2]*cos(T))+(DFTArray^[J*2+1]*sin(T)));
    end;               {Secondary loop end}
    Hold^[(I+(N div 2))*2+1]:=RealTot;
    Hold^[(I+(N div 2))*2+2]:=ImagTot;
  end;                 {Primary loop end}
  for I:=1 to N do
  begin
    DFTArray^[I*2-1]:=Hold^[I*2-1];
    DFTArray^[I*2]:=Hold^[I*2];
  end;
  freemem(Hold,(N+1)*2*sizeof(real));
end; {End of DFT procedure.}
```

The data are passed to the algorithm via a pointer to an array of reals. The

odd numbered elements of the array contain the real part of $h_n$ while the even numbered elements contain the imaginary part of $h_n$ (in this case zero). The algorithm returns the $H_n$ values via the same array and in the same format. The secondary loop, indicated in the source code listing, is executed $N^2$ times in the computation of the discrete Fourier transform which may be prohibitive when using this algorithm repetitively.

The power spectra obtained from the discrete Fourier transforms of these test waveforms are shown in Figure A.3. Panel **a** shows the power spectra for the four waveforms plotted in panel **a** of Figure A.2. In each case the power spectrum is a delta function positioned at frequency $3.4722 \times 10^{-3}$ cycles per TOMS cell which corresponds to a period of 288 TOMS cells, as expected. Consider the theoretical power of a wavenumber 1 test waveform:

$$\text{Power} = \frac{1}{2\pi} \int_0^{2\pi} A^2 \sin^2 \theta \, d\theta \tag{A.9}$$

where A is the amplitude of the sine wave. Performing the integration gives:

$$\text{Power} = \frac{A^2}{2} \tag{A.10}$$

The wave powers observed in panel **a** agree exactly with the theoretical wave powers. The power spectra shown in panel **b** are all identical indicating that the algorithm performs perfectly for various phase shifts. The discrete power spectra shown in panel **c** indicate that the algorithm provides accurate representation of wave power at different wavenumbers. The power spectra plotted in panel **d** show the effects of the asynoptic sampling by the TOMS satellite. Preliminary studies of this effect indicated that the problem became more acute with increasing wavenumber. For wave 4 it is clear that the asynoptic sampling causes spreading of the power to different wave modes so that the power spectrum no longer appears as a delta function. Table A.1 lists the wave power values at wave 4 for the four different wave speeds considered. For fast moving waves (30°per day) the error in the calculated wave power can be close to 30%.

Although computationally very demanding, and still subject to asynoptic sampling errors, the discrete Fourier transform provides an accurate measure of wave period and power for the test waveforms.

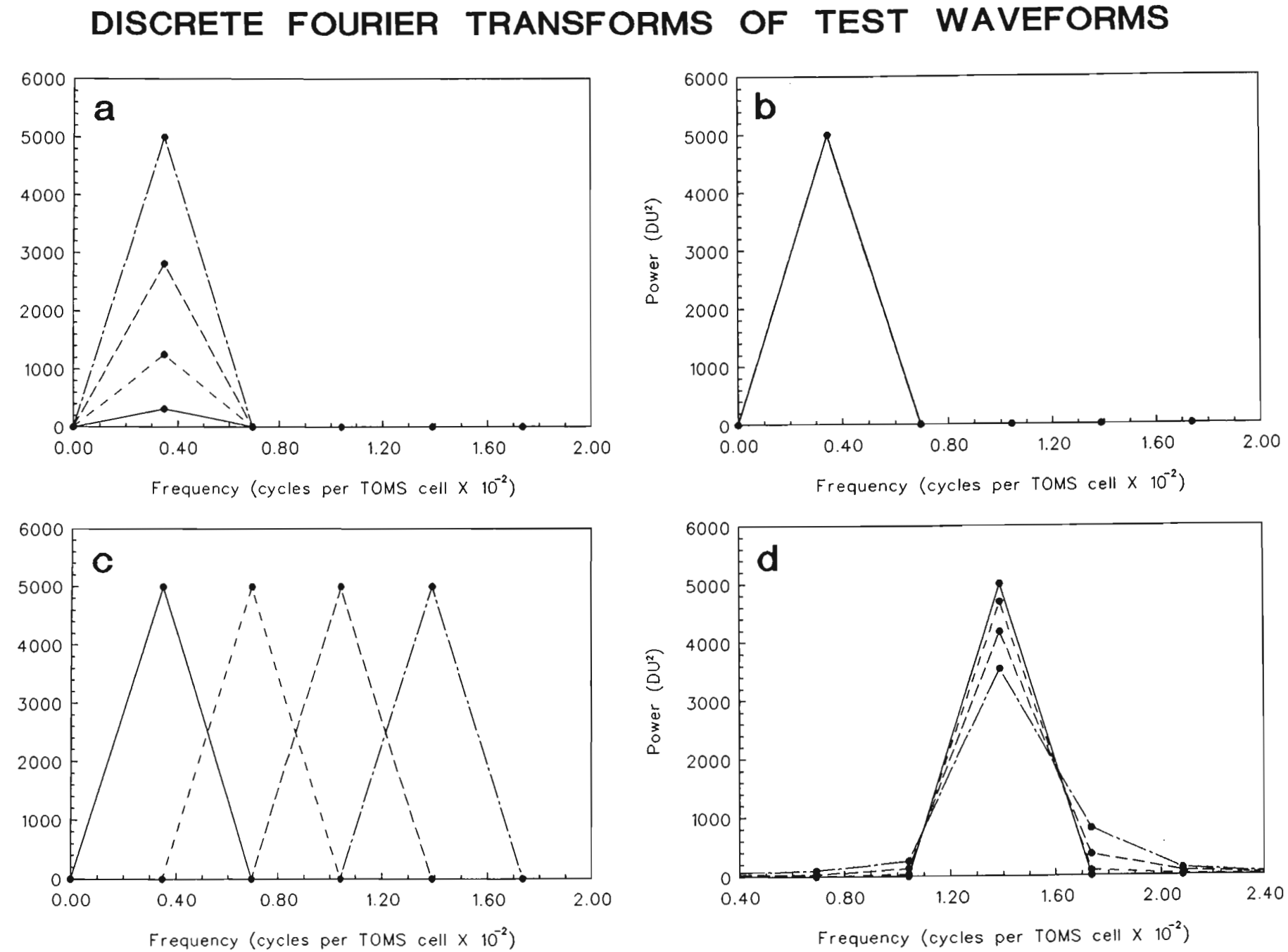


Figure A.3: The power spectra obtained from the discrete Fourier transforms of the test waveforms plotted in Figure A.2.

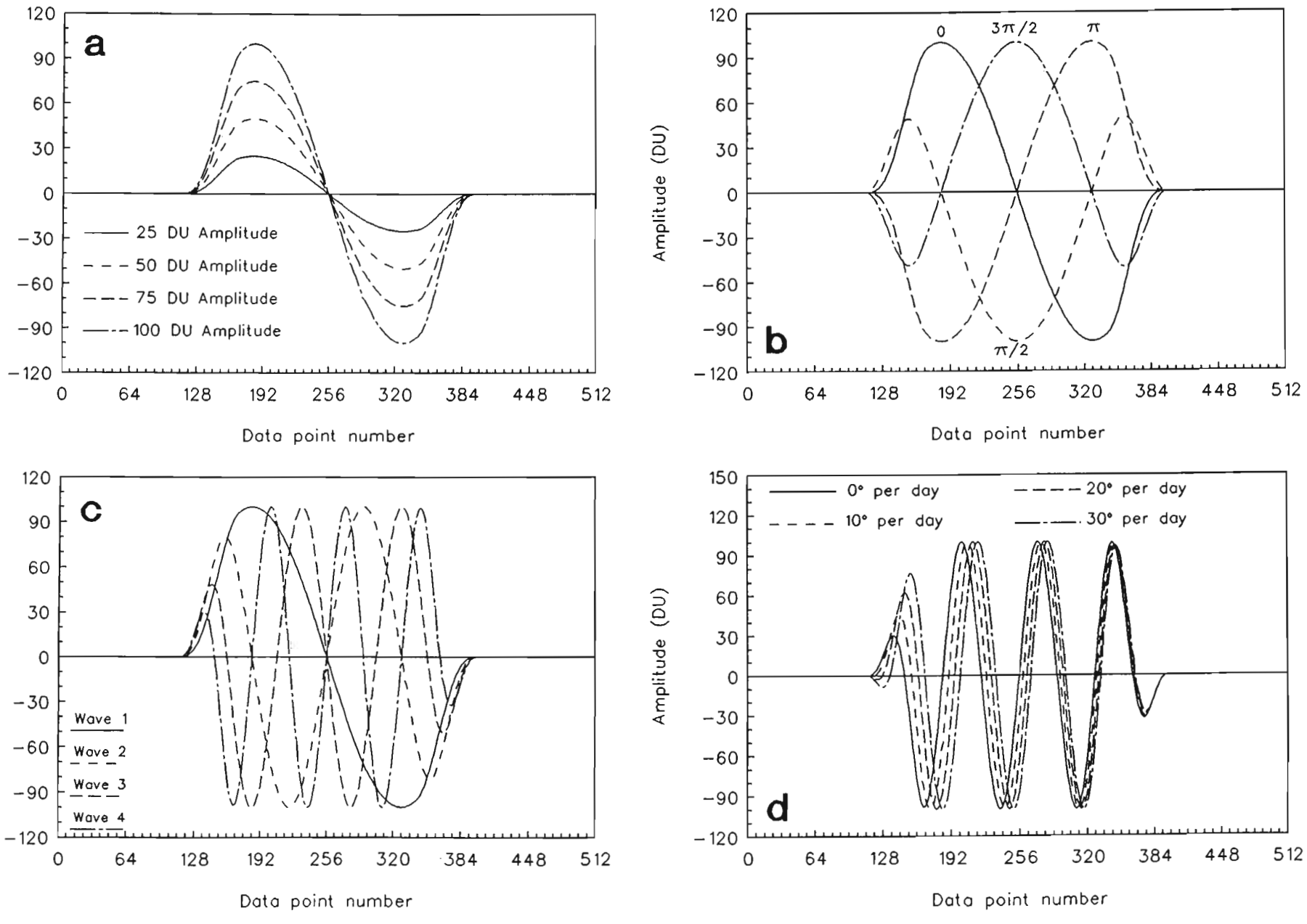


Figure A.4: The sixteen artificial zonal ozone profiles used to test the high resolution FFT.

| Speed (°per day) | Power at wave 4 ($DU^2$) | % error |
|---|---|---|
| 0° | 5000.0 | 0.00% |
| 10° | 4693.4 | 6.13% |
| 20° | 4183.5 | 16.33% |
| 30° | 3542.7 | 29.15% |

Table A.1: Power at wave 4 obtained from a discrete Fourier transform of 4 asynoptically sampled midlatitude zonal total column ozone profiles of waves travelling at speeds of 0°, 10°, 20°, and 30° longitude per day.

### A.2.3 The high resolution Fast Fourier Transform

To reduce the processing time required for the analysis of wave structures in the TOMS data it was decided to test a Fast Fourier Transform (FFT) algorithm. However, as mentioned above, the FFT algorithm requires that the number of data points in the input data array be an integral power of 2. The 288 data points of the test waveforms need to be expanded to 512 ($2^9$) values. This is accomplished by padding the data array with 112 zeros at each end. The process results in sharp steps at the ends of the original data array when the values are not close to zero. It is therefore necessary to impose a cosine Hanning window on the first and last 20% of the data array to smoothly match the original data to the zero padding. The effect of this padding and the windowing on the test data series is shown in Figure A.4 where the modified test waveforms are plotted. The data arrays are distorted at the edges due to the cosine window smoothing.

The FFT algorithm used in this analysis was adapted from Press et al. (1989). The pascal code written to implement the algorithm is listed below.

```
const
  MaxMathArray=1024;

type
  TMathArray=array[1..MaxMathArray] of real;
  PMathArray=^TMathArray;

procedure FFT(var FFTArray:PMathArray;N:integer);
```

```
{=============================================================================
{ This function performs a Fast Fourier Transform (FFT) on the data passed
{ via FFTArray. Note that only the odd numbered positions in this array
{ contain data since the imaginary parts of the original data points are
{ zero. The N parameter is the number of original data points which is half
{ the number of elements in the FFTArray. The FFTArray is replaced by its
{ direct Fourier transform. Note that N must be an integer power of 2.
{-----------------------------------------------------------------------------

var
  Count1,Count2,                    {Loop control variables.}
  NumData,                          {The number of complex data points.}
  MMax,
  istep,
  I,J:integer;                      {Loop control variables.}
  TempReal,TempImag,                {Temporary storage variables.}
  wtemp,wr,wpr,wpi,wi,theta:double;
begin
  J:=1;
  for Count1:=1 to N do
  begin
    I:=2*Count1-1;
    if (J>I) then
    begin
      TempReal:=FFTArray^[J];
      TempImag:=FFTArray^[J+1];
      FFTArray^[J]:=FFTArray^[I];
      FFTArray^[J+1]:=FFTArray^[I+1];
      FFTArray^[I]:=TempReal;
      FFTArray^[I+1]:=TempImag;
    end;
    NumData:=N;
    while((NumData>=2) and (J>NumData)) do
    begin
      J:=J-NumData;
      NumData:=NumData div 2;
```

```
    end;
    J:=J+NumData;
  end;
  MMax:=2;
  while (N*2>MMax) do
  begin
    istep:=2*MMax;
    theta:=6.28318530717959/MMax;
    wpr:=-2.0*sqr(sin(0.5*theta));
    wpi:=sin(theta);
    wr:=1.0;
    wi:=0.0;
    for Count1:=1 to (MMax div 2) do
    begin
      NumData:=2*Count1-1;
      for Count2:=0 to ((N-NumData) div istep) do
      begin
        I:=NumData+Count2*istep;
        J:=I+MMax;
        TempReal:=(wr*FFTArray^[J])-(wi*FFTArray^[J+1]);
        TempImag:=(wr*FFTArray^[J+1])+(wi*FFTArray^[J]);
        FFTArray^[J]:=FFTArray^[I]-TempReal;
        FFTArray^[J+1]:=FFTArray^[I+1]-TempImag;
        FFTArray^[I]:=FFTArray^[I]+TempReal;
        FFTArray^[I+1]:=FFTArray^[I+1]+TempImag;
      end;
      wtemp:=wr;
      wr:=wr*wpr-wi*wpi+wr;
      wi:=wi*wpr+wtemp*wpi+wi;
    end;
    MMax:=istep;
  end;
end; {End of FFT procedure.}
```

The data are passed to the procedure and returned from it in the same way as for the discrete Fourier transform. However, the frequencies associated with each

returned complex $H_n$ value are in a different order, although it is not necessary to discuss this in detail here. This algorithm, running at speed order (N log N), is at least 100 times faster than the discrete Fourier transform when processing the 288 data point TOMS arrays.

The power spectra obtained by processing the test waveforms plotted in Figure A.4 through the high resolution FFT are plotted in Figure A.5. There are a number of large differences between these results and those plotted in Figure A.3. The first anomaly arises from the fact that the power data points do not fall exactly on the frequencies for wavenumber 1, 2 $\cdots$ N. This means that it is necessary to interpolate between data points to obtain the power at a specific integral wavenumber. Furthermore, because the power is not concentrated at one frequency value, the total power is significantly suppressed. The ratios of the power maxima shown in panel **a** are the same as those for the power maxima shown in panel **a** of Figure A.3. The spread in power to different frequencies also arises from the distortion by the cosine window smoothing at the ends of the data arrays. In panel **b**, the FFT outputs for the waves of phase 0 and $\pi$ are identical while the power spectra for the waves of phase $\pi/2$ and $3\pi/2$ are also the same, although significantly lower than for the first two waves. This again results from the cosine window smoothing which has a significantly larger effect for the second two waves which are more like cosine and negative cosine functions. Panel **c** shows the effects of changes in wavenumber. Although the wave amplitudes were kept constant, the resultant maximum powers are not constant. This occurs both as a result of the cosine window smoothing and the distribution of the wave power amongst different frequencies. Panel **d** shows the effects of the asynoptic sampling. For increasing wave speed, the position of the power maximum appears to move to higher frequency. These are extremely undesirable effects.

Although the FFT algorithm is significantly faster than the discrete fourier transform, it performance is poor enough to discount its use in the processing of the TOMS data.

### A.2.4 The low resolution Fast Fourier Transform

Since the problems of poor performance with the high resolution FFT arise in part from zero padding and cosine window smoothing, it was decided to test a low resolution FFT where the data are averaged in sets of 9 TOMS cells to produce

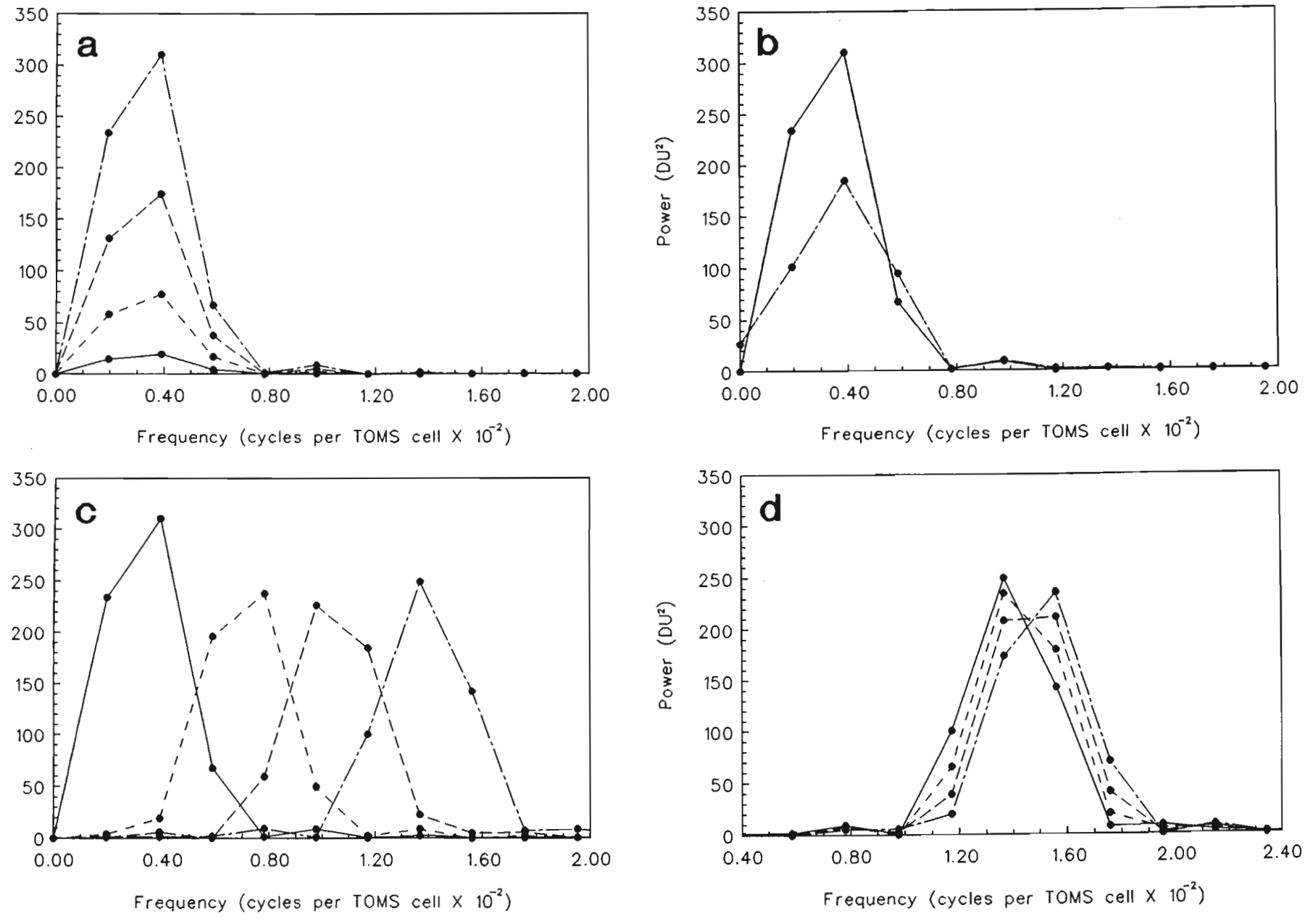


Figure A.5: The power spectra obtained from the high resolution FFTs of the test waveforms plotted in Figure A.4.

an array of 32 data points. Since these signals require no zero padding, they also require no cosine window smoothing. The modified test waveforms as a result of the averaging are shown in Figure A.6. The low resolution becomes apparent at wavenumbers of 3 and higher. These test waveforms were then passed through the FFT algorithm described in section A.2.3 to obtain the power spectra shown in Figure A.7. These show a marked improvement over the power spectra displayed in Figure A.5. The data points now fall on frequencies corresponding to integral wavenumbers with the result that the peaks in the power spectrum occur as delta functions. Note that the amplitudes of the powers are 4 times smaller than those of the discrete Fourier transforms (cf. Figure A.3). This does not present a problem since it is the day-to-day variation in power and the ratios between wave powers that are of interest. In panel **b** of Figure A.7 the power for the waves with phase 0 and $\pi$ is slightly lower than that for waves with phase $(\pi/2)$ and $(3\pi/2)$. This is a result of the 9 cell averaging which is also responsible for the differences in wave powers observed in panel **c**. These powers should be identical since only wave period varies between the 4 waveforms. The power spectra shown in panel **d** show similar behavior to those of panel **d** of Figure A.3 as a result of the asynoptic sampling by TOMS.

Although the low resolution FFT shows a considerable improvement over the high resolution FFT, there are still a number of important drawbacks in using this algorithm. Variations in wavenumber and, less importantly, variations in phase, result in differences in power amplitude where there should be no differences. The problem of asynoptic sampling remains.

## A.3 Least squares fit methods

One final algorithm was developed to derive wave amplitude and phase for a fixed wavenumber from a midlatitude representative ozone profile. Zonal profiles are extracted from the daily TOMS distributions and the mean is removed. For wavenumber N, a sine curve of the form

$$y_i = \mathrm{A}\ \sin((\mathrm{i} + \phi) \times 2\pi)/(288/N) \qquad \text{(A.11)}$$

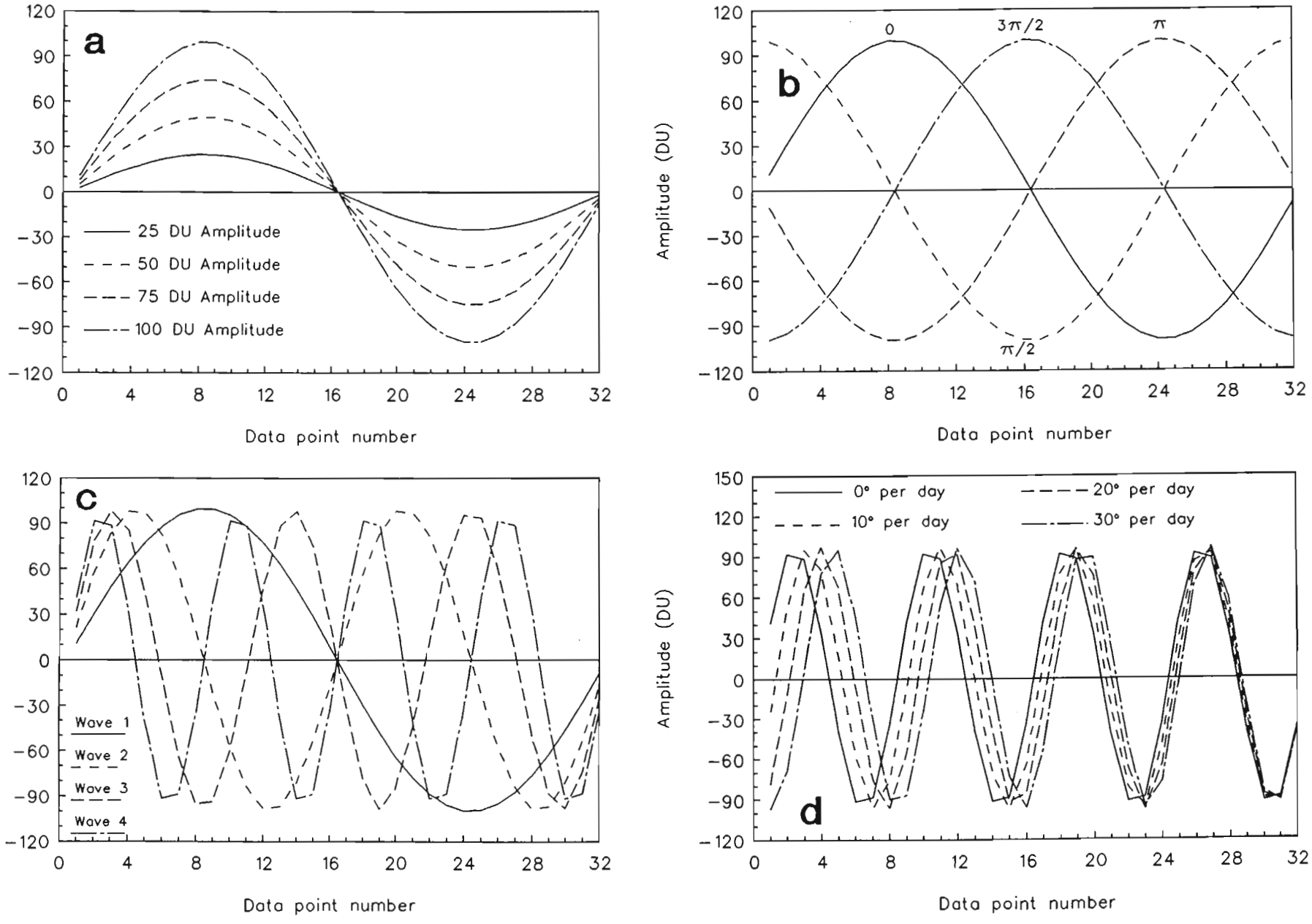


Figure A.6: The sixteen artificial zonal ozone profiles used to test the low resolution FFT. The data signals plotted in Figure A.2. have been reduced to 32 data point resolution.

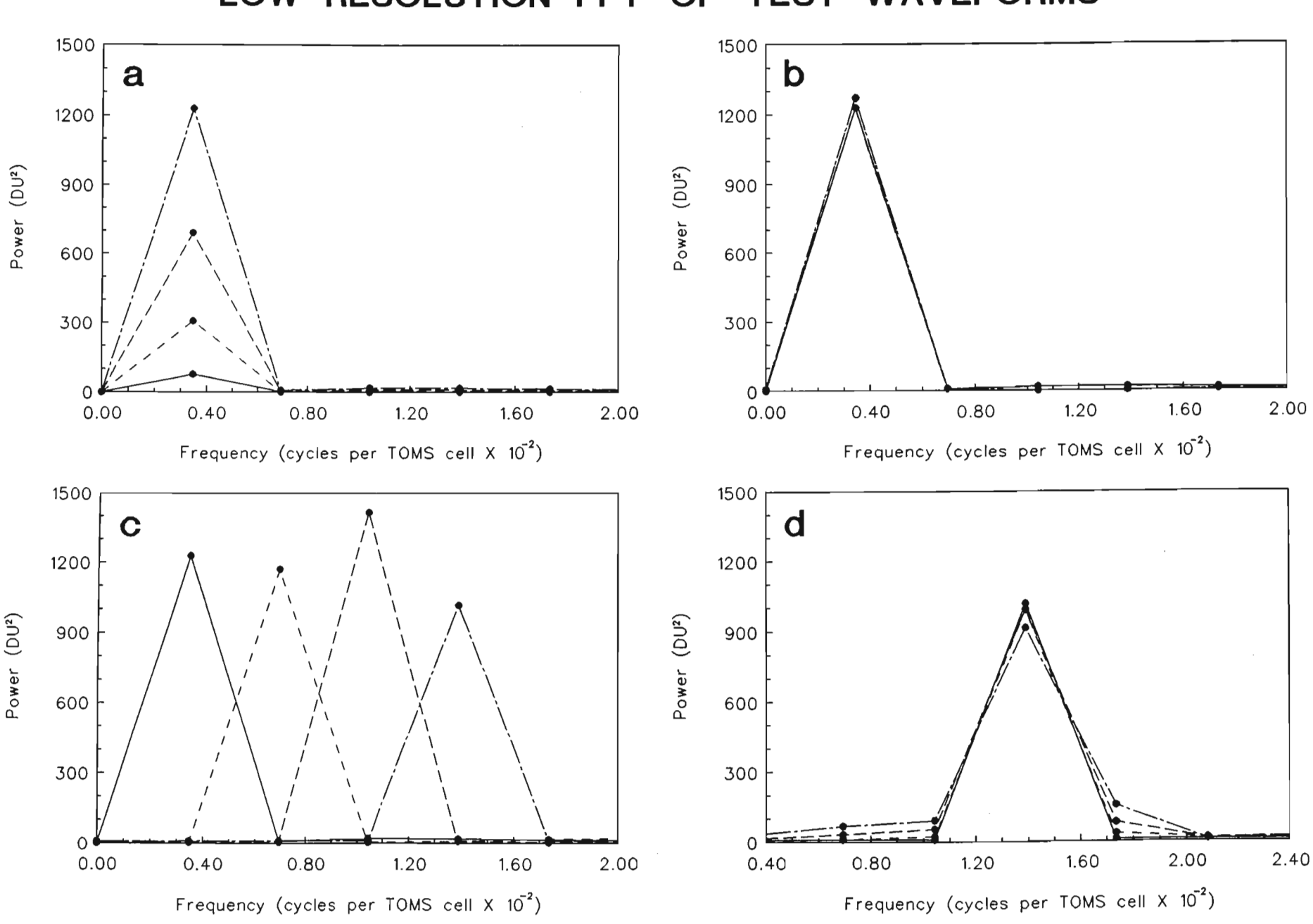


Figure A.7: The power spectra obtained from the low resolution FFTs of the test waveforms plotted in Figure A.6.

where A is the amplitude, i is the TOMS cell number (1 to 288) and $\phi$ is the phase, is fitted to the zonal profile by minimizing the quantity

$$D = \sum_{i=1}^{288} (x_i - y_i)^2 \qquad (A.12)$$

where $x_i$ are the measured data and $y_i$ are the values from the sine function (equation A.11). An iterative process is used to find the absolute minimum of D in the two dimensional amplitude-phase space associated with equation A.11. An example of this curve fitting process is shown in Figure A.8 where the zero mean 65°S profile for day 272 (28 September) of 1988 is shown together with the best fit sine curve. This particular profile was selected since it occurs in one

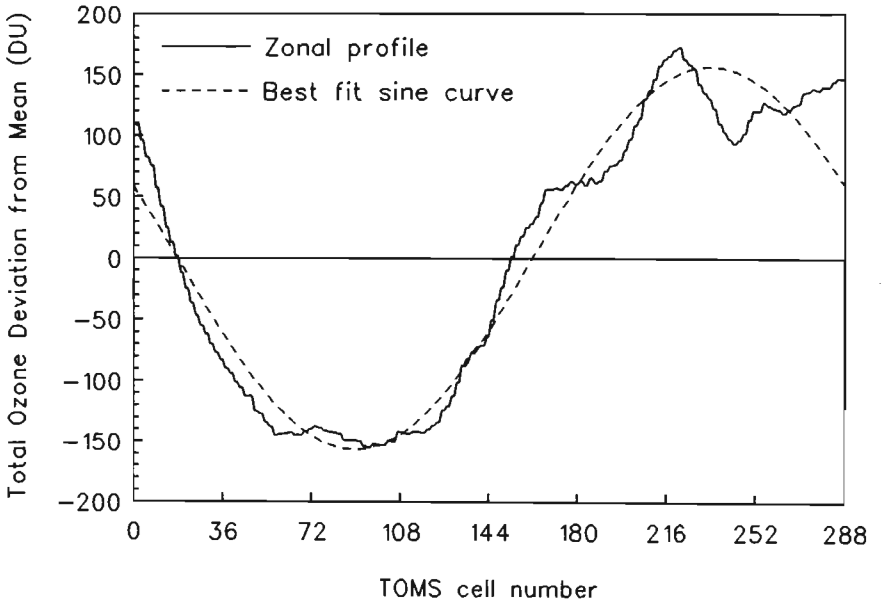


Figure A.8: Zonal total column ozone profile along 65°S for day 272 (28 September) 1988 together with its best fit wave 1 period sine function.

of the largest wave 1 events over the 1979 to 1988 period. The abscissa labels are TOMS cell number going from 1 to 288, encircling the globe. The fitted sine curve (dashed line) with an amplitude of 156.4 DU and a phase of 126.2 TOMS cells accurately tracks the zonal profile (solid line). The phase in units of TOMS cells may be converted to units of longitude using

$$\text{Longitude} = (\text{Phase} \times 1.25) - 180.625 \qquad (A.13)$$

In this case a phase of 22.9°W is obtained. This technique may be used to obtain an accurate measure of wave amplitude and phase at any wavenumber.

# Appendix B

# THE SURFACE UV IRRADIANCE MODEL: VERSION 1.00

## B.1 Description of the model

A single layer model has been developed to map the extraterrestrial irradiance to the clear-sky global UV irradiance incident on a horizontal surface. This model is based on that of McKenzie (1991) with additions from Iqbal (1983). The global irradiance (irradiance integrated over a zenith pointing hemisphere) on a horizontal surface is given by the sum of the direct component and the diffuse component. Direct spectral irradiance at any wavelength on a surface normal to the incoming radiation is given by (Iqbal, equation 6.14.3):

$$I_{n\lambda} = I_{0n\lambda}\tau_{\lambda} \tag{B.1}$$

and on a horizontal surface (Iqbal, equation 6.14.4):

$$I_{b\lambda} = I_{0n\lambda}\cos(\theta_z)\tau_{\lambda} \tag{B.2}$$

where $I_{0n\lambda}$ is the extraterrestrial irradiance, the $n$ subscript denoting direct irradiance and $\lambda$ indicates the wavelength dependence. The extraterrestrial irradiance data used in this model were extracted from a data base on the NASA/GSFC TOMS CD-ROM. $\theta_z$ is the solar zenith angle, the angle from the vertical (zenith) to the sun. $\tau_{\lambda}$ is the transmittance which is expanded further below. The diffuse

spectral irradiance on a horizontal surface is given by (Iqbal, equation 6.15.8):

$$I_{d\lambda} = I_{dr\lambda} + I_{da\lambda} + I_{dm\lambda} \tag{B.3}$$

where:

$I_{dr\lambda}$ is the diffuse spectral irradiance produced by Rayleigh scattered light incident on the ground after the first pass through the atmosphere.

$I_{da\lambda}$ is the diffuse spectral irradiance produced by aerosol scatter incident on the ground after the first pass through the atmosphere.

$I_{dm\lambda}$ is the diffuse spectral irradiance produced by multiple reflections.

The diffuse component can be expressed as (Iqbal equations 6.19.1 and 6.19.2):

$$I_{dm\lambda} = \frac{(I_{dr\lambda} + I_{da\lambda} + I_{n\lambda}\cos\theta_z)(\rho_{g\lambda}\rho_{a\lambda})}{1 - \rho_{g\lambda}\rho_{a\lambda}} \tag{B.4}$$

where $\rho_{g\lambda}$ is the ground reflectance and $\rho_{a\lambda}$ is the effective sky reflectance. The ground reflectance is small ($\sim$0.1) in the absence of snow cover (Iqbal, pg 285) and although the effective sky reflectance is solar zenith angle dependent, its value may be kept constant at 0.1 for irradiances in the UV (McKenzie, 1991). The diffuse component of the surface irradiance is therefore given by (Iqbal, equation 6.20.1):

$$I_{d\lambda} = \frac{I_{n\lambda}\cos\theta_z\rho_{g\lambda}\rho_{a\lambda} + I_{dr\lambda} + I_{da\lambda}}{1 - \rho_{g\lambda}\rho_{a\lambda}} \tag{B.5}$$

The total global irradiance on a horizontal surface is then given by:

$$I_{\lambda} = \frac{I_{n\lambda}\cos\theta_z + I_{dr\lambda} + I_{da\lambda}}{1 - \rho_{g\lambda}\rho_{a\lambda}} \tag{B.6}$$

### B.1.1 The direct beam irradiance

Expanding the $\tau_\lambda$ term in equation B.2, the direct beam irradiance may be expressed as:

$$I_{n\lambda} = \epsilon I_{0n\lambda}\tau_{r\lambda}\tau_{a\lambda}\tau_{o\lambda}\tau_{g\lambda}\tau_{wa\lambda} \tag{B.7}$$

where $\epsilon$ is a factor accounting for the ellipticity of the Earth's orbit around the Sun, $I_{0n\lambda}$ is the extraterrestrial irradiance and the transmission coefficients are as follows (Iqbal, pg 133):

$\tau_{r\lambda}$ is the spectral transmittance of the direct beam due to molecular scattering.

$\tau_{a\lambda}$ is the spectral transmittance due to scattering and absorption by aerosols.

$\tau_{o\lambda}$ is the spectral transmittance due to absorption by the ozone layer.

$\tau_{g\lambda}$ is the spectral transmittance of the direct beam due to absorption by uniformly mixed gases such as $CO_2$ and $O_2$.

$\tau_{wa\lambda}$ is the spectral transmittance due to absorption by water vapour.

The only gas considered in this model is ozone since between 280 and 340 nm, for solar zenith angles less that 85°, absorption by other gases is small (McKenzie, 1991). Therefore only the first three transmittance terms are retained in equation B.7. The equation of the ellipticity of the Earth's orbit is given by (Iqbal, equation 1.2.1):

$$\epsilon = 1.000110 + 0.034221 \cos\Gamma + 0.001280 \sin\Gamma + 0.000719 \cos 2\Gamma + 0.000077 \sin 2\Gamma \tag{B.8}$$

where $\Gamma$ is the day angle given by:

$$\Gamma = 2\pi(d_n - 1)/365 \tag{B.9}$$

and $d_n$ is the day of the year.

### Transmittance due to molecular (Rayleigh) scattering

Rayleigh scattering by air molecules is based on the assumption that the scattering particles are less than 0.2 $\lambda$ in diameter and that the particles scatter independently of one another. The Rayleigh transmission is then given by:

$$\tau_{r\lambda} = \exp(-k_{r\lambda} m_a) \tag{B.10}$$

where $k_{r\lambda}$ is the Rayleigh extinction cross-section per unit mass at wavelength $\lambda$ measured in $\mu$ m and $m_a$ is the relative airmass. The scattering coefficient of dry air is given by (Iqbal, equation 6.4.1):

$$k_{r\lambda} = 0.008735\, \lambda^{-4.08} \tag{B.11}$$

Note that the wavelength coefficient of -4.08 is close to the expected theoretical value of 4.00 for Rayleigh scattering. The relative optical airmass is given by (McKenzie, equation 7):

$$m_a = \frac{p_s}{p_0} \frac{35}{\sqrt{1224 \cos^2\theta_z + 1}} \tag{B.12}$$

where $p_s$ is the surface pressure and $p_0$ is the standard atmospheric pressure (101325 Pa).

### Transmittance due to scattering by aerosols and suspended water droplets

Using Ångström's turbidity formula (Iqbal, equation 6.6.1) the aerosol transmittance is given by:

$$\tau_{a\lambda} = \exp(-\beta\lambda^{-\alpha} m_a) \tag{B.13}$$

where $\beta$, the Ångström turbidity coefficient, varies from less than 0.05 for clean air (visibility $>$ 90 km) to greater than 0.5 for turbid conditions (visibility $<$ 5 km), $\alpha$ is the wavelength exponent for aerosol scatter with a fixed value of 1.3 (Iqbal, 1983) and $\lambda$ is the wavelength measured in $\mu$m.

### Transmittance due to absorption by the ozone layer

Ozone absorbs in the ultraviolet, visible and infrared, although in this model only absorption in the ultraviolet is of interest. Rewriting the Bouguer law (Iqbal, equation 6.2.1) for ozone we obtain:

$$\tau_{o\lambda} = \exp(-k_{o\lambda} l m_o) \tag{B.14}$$

where $k_{o\lambda}$ is the absorption cross-section for ozone, $l$ is the total column ozone in DU and $m_o$ is the relative airmass factor for ozone. The ozone absorption cross-sections used in this model were obtained from Bass and Paur (1981). The airmass factor for ozone may be expressed as (Iqbal, equation 5.9.1):

$$m_o = \frac{1 + z/R}{\sqrt{\cos^2\theta_z + 2(z/R)}} \tag{B.15}$$

where it is assumed that all the ozone is concentrated in a thin layer centered at height $z$ and $R$ is the radius of the Earth. The errors introduced by this assumption are discussed in McKenzie (1991). The assumption is necessary since only total column ozone data are available.

## B.1.2 The Rayleigh scattered diffuse irradiance

The diffuse spectral irradiance produced by Rayleigh scattering that is incident on the ground after the first pass through the atmosphere is given by (Iqbal, equation 6.16.2):

$$I_{dr\lambda} = \epsilon I_{0n\lambda} \cos(\theta_z) \tau_{ma\lambda} [0.5(1 - \tau_{r\lambda})\tau_{a\lambda}] \tag{B.16}$$

### B.1.3 The aerosol scattered spectral diffuse irradiance

The diffuse spectral irradiance produced by aerosol scattering that is incident on the surface after the first pass through the atmosphere is given by (Iqbal, equation 6.17.2):

$$I_{da\lambda} = I_{0n\lambda}\cos(\theta_z)\tau_{ma\lambda}[F_c\omega_0(1 - \tau_{a\lambda})\tau_{r\lambda}] \tag{B.17}$$

where $F_c$ is the forward scatterance, the ratio of energy scattered in the forward direction to the total energy scattered, and $\omega_0$ is the single scattering albedo, which is approximately 0.95 for clean air, but reduces to approximately 0.7 for absorbent aerosols (McKenzie, 1991). A function was fitted to the forward scatterances listed in Iqbal (Table 6.17.1) to obtain $F_c$ as a function of solar zenith angle:

$$F_c = -9.04914 \times 10^{-5}\theta_z^2 + 3.251221 \times 10^{-3}\theta_z + 0.900303156 \tag{B.18}$$

## B.2 Software for running the model

Software has been developed to encode this model. Use was made of the Object Orientated features of Borland Pascal 7 to write the object code listed at the end of this Appendix. Descriptions of the object methods are given below:

### B.2.1 How to use this object

Any pascal program making use of this model must incorporate the **uvmodel** unit under the **uses** unit list. This unit defines a UV model object of type TUVModel. An instance of the model object is created as follows:

```
var
  UVM:TUVModel;
```

There are then three steps that must be followed when running the model, viz:

1. Initialize the model
2. Run the model
3. Close the model

### Initializing the model

The model is initialized by calling the model constructor method. Model parameters are passed to the object through a *Transfer Record* of type TUVTransRec. An instance of this record must therefore be created at the beginning of the program as follows:

```
var
  TransRec:TUVTransrec;
```

The different fields of this record are:

1) Time: This itself is a record of type TTime defined in the source code. It has three fields viz. hour, minute and second.

2) SolarTime: This is a boolean variable and must be set to true if the time passed to the object is solar time eg. to obtain model data for when the sun is highest in the sky, the time is set as 12:00:00 and this field is set to true. Usually local time is used and this variable is set to false.

3) TimeZone: This is an integer variable specifying the time zone of the location. For South Africa this is +2. The value of this field is ignored if SolarTime is set to true.

4) Day: This is the day number of the year eg. 160 for 9 June of a non leap year.

5) Latitude: This is the latitude of the location in decimal degrees, positive for the Northern Hemisphere and Negative for the Southern Hemisphere.

6) Longitude: This is the longitude of the location in decimal degrees, positive for east and negative for west. The value of this field is ignored if SolarTime is set to true.

7) SurfacePressure: This value is currently unused in version 1.00 of the model and may be set to 101325 Pa.

8) UpperAirTemp: This is the upper air temperature (in K) used to determine the ozone absorption cross-sections which are temperature dependent. The value must be taken at the assumed altitude of the ozone layer. For most applications this value may be set to 213.0

9) Alpha: This is wavelength component for aerosol scatter as described in section B.1. This may be set to 1.3.

10) Beta: This is the Ångström turbidity coefficient as described in section B.1. This may be set to 0.05.

11) Ozone: This is the total column ozone in DU.

Values may then be assigned to this record as follows:

```
with TransRec do
begin
  Time.Hour:=12;
  Time.Minute:=10;
  Time.Second:=13;
  SolarTime:=false;
  TimeZone:=2;
  Day:=160;
  Latitude:=-30.673;
  Longitude:=29.087;
  SurfacePressure:=101325;
  UpperAirTemp:=213.0;
  Alpha:=1.3;
  Beta:=0.05;
  Ozone:=287.0;
end;
```

Once values have been assigned to the transfer record, the model is initialized as follows:

```
UVM.Init(TransRec);
```

A complete description of this routine is given in section B.2.2. If at a later stage in the program new values need to be assigned to the model parameters, UVM.Init need not be called again. In fact UVM.Init should be called only once ever in the program. New values may be assigned to the model after calling UVM.Init by using:

```
UVM.SetNewValues(TransRec);
```

where the new values have first been assigned to TransRec. Note that it is not necessary to reassign all the values. Only those which have changed need be reset.

In addition to initializing the model variables, the solar spectrum data and ozone absorption cross-section data must also be loaded. If it is necessary to multiply the surface spectrum by an action spectrum, then these data must also be loaded from a file. These three spectra must all be of 0.05 nm wavelength resolution, ranging from 280 nm to 340 nm. A detailed description of the format of these three files is given in section B.2.3. The files may be loaded into the model as follows:

```
UVM.LoadIrrad('uvm_dat1.dat',Ecode);
UVM.LoadAbsorb('uvm_dat2.dat',Ecode);
UVM.LoadAction('uvm_dat3.dat',Ecode);
```

A complete description of these methods is given in section B.2.2. The methods can be called at any time during the program to reset the relevant spectra.

It may be that at times during the model run it is necessary to remove effects of the action spectrum. Rather than loading a file full of 1's, the action spectrum may be set to 1 by calling:

```
UVM.SetActionOne;
```

There may also be occasions where it is required that surface irradiances be calculated at fixed solar zenith angles rather than at fixed times. The solar zenith angle is calculated when UVM.Init or UVM.SetNewValues is called. The calculated value may then be overwritten by calling:

```
UVM.SetSZA(83.0)
```

where 83.0 is the new value for the solar zenith angle. A complete description of this routine is given in section B.2.2.

### Running the model

To run the model, the CalcSpec routine is called as follows:

```
UVB:=UVM.CalcSpec(280,340,1);
```

The routine returns the integrated irradiance under the surface spectrum after it has been multiplied by the action spectrum. The first two numbers passed to the routine specify the start and end wavelengths over which this integration is performed. The third parameter indicates whether the global, direct or diffuse spectrum must be calculated. A complete description of this routine is given in section B.2.2. If the calculated solar zenith angle is required after calling UVM.Init or UVM.SetNewValues then this may be retrieved by calling the function:

```
SZA:=UVM.GetSza;
```

Since the UVM.CalcSpec routine returns only a single number (the integrated irradiance), the program may also require the complete surface irradiance spectrum. To do this, a pointer to the spectrum (of type PSpectrum) must first be defined as follows:

```
var
  OutSpec:PSpectrum;
```

Then to retrieve the surface irradiance spectrum, call:

```
UVM.GetOutSpec(OutSpec);
```

The values of OutSpec may then be listed as follows:

```
for I:=1 to 1201 do writeln(OutSpec^[I]);
```

### Closing the model

The model must only be closed once at the end of the program as follows:

```
UVM.Done;
```

### B.2.2 Object methods

```
constructor TUVModel.Init(TransRec:TUVTransRec)
```

This initialization routine sets the values of all of the object variables and gets memory for all of the object spectrum arrays. The model settings are passed via the TransRec array. The value of the solar zenith angle is calculated based on the day of the year and the time of day. A call is made to the GetZenithAngle function defined in the astro unit. The procedure also sets the values of the action spectrum to 1 in case no action spectrum file is loaded. The values of the irradiance and absorption cross-section spectrum are set to zero to allow for error checking. Note that the routine assigns memory from the heap for the spectrum arrays. This procedure need be called only once in every program.

```
procedure TUVModel.LoadIrrad(Filename:string;var ErrorCode:byte)
```

This routine loads the irradiance data for extraterrestrial irradiance spectrum. The complete file name for the data file is passed to the routine via the Filename parameter. The variable ErrorCode parameter is used to return an error code. The values of ErrorCode are as follows:

0: Load was successful.

1: File could not be found.

2: File didn't have enough data.

3: X values in file were wrong.

```
procedure TUVModel.LoadAbsorb(Filename:string;var ErrorCode:byte)
```

This routine reads in the data from the ozone absorption cross-section file. Since this file has data for a number of different temperatures, interpolation must be performed to get the values for the specified upper air temperature. The ErrorCode values are as follows:

0: Load was successful.

1: File could not be found.

2: File didn't have enough data.

3: X values in file were wrong.

4: The upper air temperature was outside the range of the data in the file.

```
procedure TUVModel.LoadAction(Filename:string;var ErrorCode:byte)
```

This routine may be used to load an action spectrum which allows for the calculation of erythemal irradiances etc. The name of the action spectrum data file is passed via Filename. The returned ErrorCodes are as follows:
0: Load was successful.
1: File could not be found.
2: File didn't have enough data.
3: X values in file were wrong.

```
procedure TUVModel.SetNewValues(TransRec:TUVTransRec)
```

This routine sets new values for all of the model parameters. It is the responsibility of the calling program to reload absorption cross-section data etc. if necessary. The procedure also sets the values of the output spectrum to zero. The solar zenith angle is recalculated based on the new values of day number and time.

```
procedure TUVModel.SetActionOne
```

This routine sets all of the action spectrum values to 1. This may be called if no action spectrum is required after having used an action spectrum earlier in the program.

```
procedure TUVModel.SetSZA(NewSZA:double)
```

This routine assigns a new value to the SZA variable. This may be used when the time of the day is not known and the surface spectrum must be calculated for some specified solar zenith angle. The new solar zenith angle is passed via the NewSZA parameter.

```
function TUVModel.CalcSpec(StartWave,EndWave:double;SType:byte):double
```

This procedure calculates the surface irradiance spectrum. The StartWave and EndWave parameters specify the wavelength range over which the calculation of the spectrum is performed. The values of SType are:
1 = Calculate Global Irradiance.
2 = Calculate only Direct Irradiance.
3 = Calculate only Diffuse Irradiance.

```
unit uvmodel;

{*************************************************************************}
{*                                                                       *}
{*                          THE  UV  MODEL  UNIT                         *}
{*                              VERSION 1.00                             *}
{*                                                                       *}
{* This unit defines an object which encodes the surface UV model of Iqbal.  *}
{* The source equations for this model were obtained from:               *}
{* 1) McKenzie R.L., Application of a simple model to calculate latitudinal  *}
{*    and hemispheric differences in ultraviolet radiation, Weather and      *}
{*    Climate, volume 11, pages 3 to 14, 1991.                           *}
{* 2) Iqbal M., An introduction to Solar Radiation, Academic Press, 1983.    *}
{*                                                                       *}
{* The purpose of this model is to take a solar irradiance spectrum at the   *}
{* top of the atmosphere and map it to a surface irradiance spectrum as a    *}
{* function of ozone, time of the day, day number of the year, time zone of  *}
{* the location, latitude, longitude, surface pressure, the 50 hPa temper-   *}
{* ature, the wavelength exponent for scatter, and the Angstrom turbidity    *}
{* coefficient. If an action spectrum is available, the model may also be    *}
{* be used to calculate UV doses for various biological systems. There are   *}
{* two input data files which the model MUST have in order to work:          *}
{* A 0.05 nm resolution file of irradiances at the top of the atmsosphere.   *}
{* A 0.05 nm resolution file of ozone absorption cross-sections.             *}
{* These two files are supplied together with the model. If any changes are  *}
{* made to the files, please ensure that they are kept in the same format.   *}
{* The wavelength operating range for this model is 280 to 340 nm.           *}
{*                                                                       *}
{* Written by: Greg Bodeker.                                             *}
{* Date: 17 November 1993.                                               *}
{* Last updated: 2 December 1993.                                        *}
{* History: Version 1.00 completed - 17 November 1993.                   *}
{*************************************************************************}

interface

const
  MaxSpec=1201;          {Number of values in each spectrum.}
  StartSpec=280;         {The starting point for each spectrum.}
  SpecInc=0.05;          {The step size in nm.}

type
  TSpectrum=array[1..MaxSpec] of double; {The spectrum type.}
  PSpectrum=^TSpectrum;                  {Spectrum pointer type.}
  TTime=record                           {The time type.}
          Hour,                          {The hour.}
          Minute,                        {The minutes.}
          Second:byte;                   {The seconds.}
        end;
  TUVTransRec=record                     {To pass object parameters.}
            Time:TTime;                  {The time of the day.}
```

```
{=============================================================================}
{ This routine loads the irradiance data for the solar spectrum at the top of }
{ the atmosphere. The complete file name for the data file is passed to the   }
{ routine via the Filename parameter. The variable ErrorCode parameter is     }
{ used to return an error code. The values of ErrorCode are as follows:       }
{ 0: Load was successful.                                                     }
{ 1: File could not be found.                                                 }
{ 2: File didn't have enough data.                                            }
{ 3: X values in file were wrong.                                             }
{-----------------------------------------------------------------------------}

var
  Source:text;           {The file handle for the text data file.}
  WaveLengthIn,          {The wavelength read in from the file.}
  WaveLengthTest:double; {A test wavelength.}
  I:integer;             {Loop control variable.}
  BadXData:boolean;      {Set to true if X data were bad.}
begin
  if not CheckFile(Filename) then            {CheckFile is in the files unit.}
  begin
    ErrorCode:=1;
    exit;
  end;
  assign(Source,Filename);
  reset(Source);
  DeleteHeader(Source);                   {DeleteHeader is in the files unit.}
  I:=0;
  repeat
    inc(I);
    readln(Source);
  until eof(Source);
  if I<MaxSpec then
  begin
    ErrorCode:=2;
    exit;
  end;
  DeleteHeader(Source);
  BadXData:=false;
  for I:=1 to MaxSpec do
  begin
    readln(Source,WaveLengthIn,Irrad^[I]);
    WaveLengthTest:=StartSpec+(I-1)*SpecInc;
    if WaveLengthTest<>WaveLengthIn then BadXData:=true;
  end;
  close(Source);
  if BadXData then ErrorCode:=3 else ErrorCode:=0;
end; {End of TUVModel.LoadIrrad.}


procedure TUVModel.LoadAbsorb(Filename:string;var ErrorCode:byte);
```

```
{=============================================================================}
{ This routine reads in the data from the ozone absorption cross-section      }
{ file. Since this file has data for a number of different temperatures,      }
{ interpolation must be performed to get the values for the specified upper   }
{ air temperature. The ErrorCode values are as follows:                       }
{ 0: Load was successful.                                                     }
{ 1: File could not be found.                                                 }
{ 2: File didn't have enough data.                                            }
{ 3: X values in file were wrong.                                             }
{ 4: The upper air temperature was outside the range of the data in the file. }
{-----------------------------------------------------------------------------}

var
  Source:text;            {The file handle for the text data file.}
  WaveLengthIn,           {The wavelength read in from the file.}
  WaveLengthTest:double; {A test wavelength.}
  Col,                    {Col and Col+1 are the two columns to be read in.}
  I,J:integer;            {Loop control variables.}
  BadXData:boolean;       {Set to true if X data were bad.}
  DataIn:array[1..5] of double; {The five cross-section values read in.}
const
  {These are the tempatures for the data listed in the file.}
  Temp:array[1..5] of double=(300.0,275.0,250.0,220.0,195.0);
begin
  Col:=0;
  for I:=1 to 4 do
  if ((UpperAirTemp<Temp[I]) and (UpperAirTemp>=Temp[I+1])) then Col:=I;
  if Col=0 then
  begin
    ErrorCode:=4;
    exit;
  end;
  if not CheckFile(Filename) then
  begin
    ErrorCode:=1;
    exit;
  end;
  assign(Source,Filename);
  reset(Source);
  DeleteHeader(Source);                    {DeleteHeader is in the files unit.}
  I:=0;
  repeat
    inc(I);
    readln(Source);
  until eof(Source);
  if I<MaxSpec then
  begin
    ErrorCode:=2;
    exit;
  end;
  DeleteHeader(Source);
```

```
  BadXData:=false;
  for I:=1 to MaxSpec do
  begin
    read(Source,WaveLengthIn);
    for J:=1 to 5 do read(Source,DataIn[J]);
    readln(Source);
    WaveLengthTest:=StartSpec+(I-1)*SpecInc;
    if (WaveLengthTest-WaveLengthIn>0.0) then BadXData:=true;
    Absorb^[I]:=DataIn[Col]+((UpperAirTemp-Temp[Col])*(Datain[Col+1]-
               DataIn[Col])/(Temp[Col+1]-Temp[Col]));
  end;
  close(Source);
  if BadXData then ErrorCode:=3 else ErrorCode:=0;
end; {End of TUVModel.LoadAbsorb.}


procedure TUVModel.LoadAction(Filename:string;var ErrorCode:byte);

{=============================================================================}
{ This routine may be used to load an action spectrum which allows for the    }
{ calculation of erythemal irradiances etc. The name of the action spectrum   }
{ data file is passed via Filename. The returned ErrorCodes are as follows:   }
{ 0: Load was successful.                                                     }
{ 1: File could not be found.                                                 }
{ 2: File didn't have enough data.                                            }
{ 3: X values in file were wrong.                                             }
{-----------------------------------------------------------------------------}

var
  Source:text;           {The file handle for the text data file.}
  WaveLengthIn,          {The wavelength read in from the file.}
  WaveLengthTest:double; {A test wavelength.}
  I:integer;             {Loop control variable.}
  BadXData:boolean;      {Set to true if X data were bad.}
begin
  if not CheckFile(Filename) then
  begin
    ErrorCode:=1;
    exit;
  end;
  assign(Source,Filename);
  reset(Source);
  DeleteHeader(Source);
  I:=0;
  repeat
    inc(I);
    readln(Source);
  until eof(Source);
  if I<MaxSpec then
  begin
    ErrorCode:=2;
```

```
    exit;
  end;
  DeleteHeader(Source);              {This routine is in the files unit.}
  BadXData:=false;
  for I:=1 to MaxSpec do
  begin
    readln(Source,WaveLengthIn,Action^[I]);
    WaveLengthTest:=StartSpec+(I-1)*SpecInc;
    if WaveLengthTest<>WaveLengthIn then BadXData:=true;
  end;
  close(Source);
  if BadXData then ErrorCode:=3 else ErrorCode:=0;
end; {End of TUVModel.LoadAction.}


procedure TUVModel.SetNewValues(TransRec:TUVTransRec);

{=============================================================================}
{ This routine sets new values for all of the model parameters. It is the     }
{ responsibility of the calling program to reload absorption cross-section    }
{ data etc. if necessary. The procedure also sets the values of the output    }
{ spectrum to zero. The solar zenith angle is recalculated based on the new   }
{ values of day number and time.                                              }
{-----------------------------------------------------------------------------}

var
  Hour:double; {The time in decimal hours.}
begin
  TimeZone:=TransRec.TimeZone;
  Time.hour:=TransRec.Time.hour;
  Time.minute:=TransRec.Time.minute;
  Time.second:=TransRec.Time.second;
  SolarTime:=TransRec.SolarTime;
  Day:=TransRec.Day;
  Latitude:=TransRec.Latitude;
  Longitude:=TransRec.Longitude;
  SurfacePressure:=TransRec.SurfacePressure;
  UpperAirTemp:=TransRec.UpperAirTemp;
  Alpha:=TransRec.Alpha;
  Beta:=TransRec.Beta;
  Ozone:=TransRec.Ozone;
  FillChar(OutSpec^,MaxSpec*sizeof(double),0);
  Hour:=Time.hour+(Time.minute/60.0)+(Time.second/3600.0);
  SZA:=GetZenithAngle(Day,TimeZone,Latitude,Longitude,Hour,SolarTime);
end; {End of TUVModel.SetNewValues.}


procedure TUVModel.SetActionOne;

{=============================================================================}
{ This routine sets all of the action spectrum values to 1. This may be       }
```

```
{ called if no action spectrum is required after having used an action        }
{ spectrum earlier in the program.                                            }
{-----------------------------------------------------------------------------}

var
  I:integer; {Loop control variable.}
begin
  for I:=1 to MaxSpec do Action^[I]:=1.0;
end; {End of TUVModel.SetActionOne.}


procedure TUVModel.SetSZA(NewSZA:double);

{=============================================================================}
{ This routine assigns a new value to the SZA variable. This may be used when }
{ the time of the day is not known and the surface spectrum must be calcul-   }
{ ated for some specified solar zenith angle. The new solar zenith angle is   }
{ passed via the NewSZA parameter.                                            }
{-----------------------------------------------------------------------------}

begin
  SZA:=NewSZA;
end; {End of TUVModel.SetSZA.}


function TUVModel.CalcSpec(StartWave,EndWave:double;SType:byte):double;

{=============================================================================}
{ This procedure calculates the surface spectrum. The StartWave and EndWave   }
{ parameters specify the wavelength range over which the calculation of the   }
{ spectrum occurs. The values of SType are:                                   }
{ 1 = Calculate Global Irradiance.                                            }
{ 2 = Calculate only Direct Irradiance.                                       }
{ 3 = Calculate only Diffuse Irradiance.                                      }
{-----------------------------------------------------------------------------}

var
  Total,       {Used for integration.}
  Eccent,      {To account for the eccentricity of the Earth's orbit.}
  Lambda,      {The day angle.}
  AirMass,     {The air mass factor.}
  WaveLength,  {The wave length.}
  k_r,         {The Rayleigh extinction cross section per unit airmass.}
  T_r,         {The Rayleigh transmission.}
  T_a,         {The aerosol transmittance.}
  T_m,         {The ozone transmittance.}
  k_O,         {The ozone absorption cross section for ozone.}
  u_O,         {The ozone concentration in molecules/cm^2.}
  m_O,         {The relative airmass for ozone.}
  E_n,         {The direct beam irradiance.}
  E_dr,        {The rayleigh scattered diffuse irradiance.}
```

```
  E_da,        {The aerosol scattered diffuse irradiance.}
  F_c,         {Energy scattered in forward direction/total energy scattered.}
  omega_0,     {The single scattering albedo.}
  r_g,         {The ground albedo.}
  r_a:double;  {The effect sky reflectance.}
  StartIndex,EndIndex, {Index equivalents for wavelength limits.}
  I:integer;           {Loop control variable.}
const
  EarthRadius=6.38E6; {Mean radius of the Earth in meters.}
  OzoneHeight=20E3;   {The assumed height of the ozone layer in meters.}
begin
  Total:=0.0;
  Lambda:=2*pi*(Day-1)/365.0;
  Eccent:=1.000110+0.034221*cos(Lambda)+0.001280*sin(Lambda)+
          0.000719*cos(2*Lambda)+0.000077*sin(2*Lambda);
  AirMass:=35.0/sqrt(1224.0*cos(SZA)*cos(SZA)+1);          {McKenzie eqn 7.}
  u_0:=Ozone*2.69E16;
  m_0:=(1+OzoneHeight/EarthRadius)/
       sqrt(cos(SZA)*cos(SZA)+2*(OzoneHeight/EarthRadius));
  F_c:=-9.04914E-5*sqr(SZA*180.0/pi)+3.251221E-3*(SZA*180.0/pi)+0.900303156;
  omega_0:=0.9;
  r_g:=0.1;
  r_a:=0.1;
  StartIndex:=1+round((StartWave-StartSpec)/SpecInc);
  EndIndex:=1+round((EndWave-StartSpec)/SpecInc);
  for I:=StartIndex to EndIndex do
  begin
    WaveLength:=StartSpec+(I-1)*SpecInc;
    k_r:=0.008735*power(WaveLength*0.001,-4.08);     {Iqbal eqn 6.4.1 pg 115.}
    T_r:=exp(-k_r*AirMass);                          {Iqbal eqn 6.4.2 pg 116.}
    T_a:=exp(-1.0*beta*power(WaveLength,-1.0*alpha)*AirMass);
    k_0:=Absorb^[I];
    T_m:=exp(-1.0*k_0*u_0*m_0);
    E_n:=Eccent*Irrad^[I]*T_r*T_a*T_m;
    E_dr:=Eccent*Irrad^[I]*cos(SZA)*T_m*(0.5*(1.0-T_r)*T_a); {Iqbal 6.16.2}
    E_da:=Eccent*Irrad^[I]*cos(SZA)*T_m*(F_c*omega_0*(1-T_a)*T_r);
    case SType of
      1:OutSpec^[I]:=((E_n*cos(SZA)+E_dr+E_da)/(1-(r_g*r_a)))*Action^[I];
      2:OutSpec^[I]:=(E_n*cos(SZA)/(1-(r_g*r_a)))*Action^[I];
      3:OutSpec^[I]:=((E_dr+E_da)/(1-(r_g*r_a)))*Action^[I];
    end;
    Total:=Total+SpecInc*OutSpec^[I];
  end;
  CalcSpec:=Total;
end; {End of TUVModel.CalcSpec.}


procedure TUVModel.GetOutSpec(var ReturnSpec:PSpectrum);

{=============================================================================}
{ This routine passes a pointer to the output spectrum back to the calling    }
```

```
{ program via the variable ReturnSpec parameter.                              }
{-----------------------------------------------------------------------------}

begin
  ReturnSpec:=OutSpec;
end; {End of TUVModel.GetOutSpec.}


function TUVModel.GetSZA:double;

{=============================================================================}
{ This function returns the current solar zenith angle value.                }
{-----------------------------------------------------------------------------}

begin
  GetSZA:=SZA;
end; {End of TUVModel.GetSZA.}


destructor TUVModel.Done;

{=============================================================================}
{ This destructor frees the memory that was assigned to the 4 model spectra.  }
{-----------------------------------------------------------------------------}

begin
  FreeMem(Irrad,MaxSpec*sizeof(double));
  FreeMem(Absorb,MaxSpec*sizeof(double));
  FreeMem(Action,MaxSpec*sizeof(double));
  FreeMem(OutSpec,MaxSpec*sizeof(double));
end; {End of TUVModel.Done.}


end. {End of uvmodel unit.}
```

# REFERENCES


Andrews D.G. and M.E. McIntyre (1976), *Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the zonal mean acceleration.*, Journal of the Atmospheric Sciences, **33**, 2031.

Andrews D.G., J.R. Holton and C.B. Leovy (1987), *Middle Atmosphere Dynamics.*, Academic Press, San Diego, California.

Angell J.K. and J. Korshover (1962), *The biennial wind and temperature oscillations of the equatorial stratosphere and their possible extension to higher latitudes.*, Monthly Weather Review, **90**, 127 - 132.

Angell J.K. and J. Korshover (1964), *Quasi-biennial variations in temperature, total ozone and tropopause height.*, Journal of the Atmospheric Sciences, **21**, 479 - 492.

Angell J.K. and J. Korshover (1967), *Biennial variation in springtime temperature and total ozone in extratropical latitudes.*, Monthly Weather Review, **95**, 757 - 762.

Angell J.K. and J. Korshover (1973), *Quasi-biennial and long-term fluctuations in total ozone.*, Monthly Weather Review, **101**, 426 - 443.

Angell J.K. and J. Korshover (1976), *Global analysis of recent total ozone fluctuations.*, Journal of the Atmospheric Sciences, **104**, 63.

Angell J.K. and J. Korshover (1978), *Global ozone variations: An update into 1976.*, Monthly Weather Review, **106**, 725.

Austin J. and N. Butchart (1992), *A three-dimensional modelling study of the influence of planetary wave dynamics on polar ozone photochemistry.*, Journal of Geophysical Research, **97**, no. D9, 10165 - 10186.

Bass A.M. and R.J. Paur (1981), *UV absorption cross sections for ozone: The temperature dependence.*, Journal of Photochemistry, **17**, 141.

Berggren R., and K. Labitzke (1968), *The distribution of ozone on pressure surfaces.*, Tellus, **20**, 88.

Bhartia P.K., K.F. Klenk, C.K. Wong, and D. Gordon (1984), *Intercomparison of the Nimbus 7 SBUV/TOMS total ozone sets with Dobson and M83 results*,

Journal of Geophysical Research, **89**, 5239 – 5247.

Blackshear W.T. and R.H. Tolson (1978), *High correlation between variations in monthly averages of solar activity and total atmospheric ozone.*, Geophysical Research Letters, **5**, 921 – 924.

Bodeker G.E. and R.L. McKenzie (1993), *Erythemal UV at 45°S: Longitudinal and secular variability*, Weather and Climate, submitted.

Bodeker G.E. and M.W.J. Scourfield (1993), *Long-term Ozone and Temperature Correlations Above SANAE, Antarctica.*, 'Proceedings of the Quadrennial Ozone Symposium', Charlottesville, U.S.A., in press.

Bojkov R., L. Bishop, W.J. Hill, G.C. Reinsel and G.C. Tiao (1990), *A statistical trend analysis of revised Dobson total ozone data over the Northern Hemisphere.*, Journal of Geophysical Research, **95**, no. D7, 9785 – 9807.

Boville B.A. (1993), *Middle atmosphere version of CCM2: Annual cycle and interannual variability.*, Journal of Geophysical Research, submitted.

Bowman K.P. and A.J. Krueger (1985), *A global climatology of total ozone from the Nimbus 7 total ozone mapping spectrometer*, Journal of Geophysical Research, **90**, 7967 – 7976.

Bowman K.P. (1986), *Interannual variability of total ozone during the breakdown of the Antarctic circumpolar vortex.*, Geophysical Research Letters, **13**, no. 12, 1193 – 1196.

Brasseur and Solomon (1986), *Aeronomy of the Middle Atmosphere*, D. Reidel Publishing Company, 2nd Edition.

Brigham O.E. (1974), *The fast fourier transform.*, Prentice-Hall.

Brühl C. and P.J. Crutzen (1989), *On the disproportionate role of tropospheric ozone as a filter against solar UV-B radiation.*, Geophysical Research Letters, **16**, no. 7, 703 – 706.

Brusa R.W. and C. Frohlich (1982), *Recent solar constant determination from high altitude balloons.*, Paper presented at the symposium on the solar constant and the spectral distribution of solar irradiance, IAMAP third scientific Assembly, Published by the Radiation Commission, Boulder, CO, USA.

Callis L.B., R.E. Boughner, M. Natarajan, J.D. Lambeth, D.N. Baker and J.B.

Blake (1991), *Ozone depletion in the high latitude lower stratosphere: 1979 - 1990.*, Journal of Geophysical Research, **96**, no. D2, 2921 - 2937.

Cebula R.P., H. Park and D.F. Heath (1988), *Characterization of the Nimbus-7 SBUV radiometer for the long-term monitoring of stratospheric ozone.*, Journal of Atmosphere and Ocean Technology, **5**, 215 - 227.

Chandra S. and R.D. McPeters (1986), *Some observations on the role of planetary waves in determining the spring time ozone distribution in the Antarctic.*, Geophysical Research Letters, **13**, no. 12, 1224 - 1227.

Chandra S., R.D. McPeters, R.D. Hudson and W. Planet (1990), *Ozone measurements from the NOAA-9 and the Nimbus-7 satellites: implications of short and long term variabilities.*, Geophysical Research Letters, **17**, no. 10, 1573 - 1576.

Charney J.G. and P.G. Drazin (1961), *Propagation of planetary-scale disturbances from the lower into the upper atmosphere.*, Journal of Geophysical Research, **66**, no. 1, 83 - 109.

Chen P. and W.A. Robinson (1992), *Propagation of planetary waves between the troposphere and stratosphere.*, Journal of the Atmospheric Sciences, **49**, no. 24, 2533 - 2545.

Crutzen P.J. (1988), *Tropospheric ozone: an overview.*, In 'Tropospheric ozone', edited by I.S.A. Isaksen, 3 - 32, Reidel, Dordrecht.

Dave J.V. and C.L. Mateer (1967), *A preliminary study on the possibility of estimating total atmospheric ozone from satellite.*, Journal of the Atmospheric Sciences, **24**, 414.

Dave J.V. (1978), *Effect of aerosols on the estimation of total ozone in an atmospheric column from the measurements if its ultraviolet irradiance.*, Journal of the Atmospheric Sciences, **35**, 899 - 911.

Deshler T., D.J. Hofmann, J.V. Hereford and C.B. Sutter (1990), *Ozone and temperature profiles over McMurdo station Antarctica in the spring of 1989.*, Geophysical Research Letters, **17**, no. 2, 151 - 154.

Diab R., J. Barsby, G. Bodeker, M.W.J. Scourfield and L. Salter (1992), *Satellite observations of total ozone above South Africa.*, South African Geographical Journal, **74**, 13 - 18.

Diffey B.L. (1991), *Solar ultraviolet radiation effects on biological systems.*, Phys. Med. Biol., **36**, no. 3, 299 – 328.

Dobson G.M.B., D.N. Harrison and J. Lawrence (1929), *Measurements of the amount of ozone in the earth's atmosphere and its relation to other geophysical conditions.*, Proceedings of the Royal Society of London, Sec. A, **122**, 456.

Dunkerton T.J. (1991), *LIMS (Limb Infrared Monitor of the Stratosphere) observation of traveling planetary waves and potential vorticity advection in the stratosphere and mesosphere.*, Journal of Geophysical Research, **96**, no. D2, 2813 – 2834.

Evans W.F.J. (1990), *Ozone depletion in the Arctic vortex at Alert during February 1989.*, Geophysical Research Letters, **17**, no. 2, 167 – 170.

Farman J.C., B.G. Gardiner and J.D. Shanklin (1985), *Large losses of total ozone in Antarctica reveal seasonal $ClO_x/NO_x$ interaction.*, Nature, **315**, 207 – 210.

Fishman J., C.E. Watson, J.C. Larsen and J.A. Logan (1990), *Distribution of tropospheric ozone determined from satellite data.*, Journal of Geophysical Research, **95**, no. D4, 3599 – 3617.

Fleig A.J., P.K. Bhartia, C.G. Wellemeyer and D.S. Silberstein (1986 a), *Seven years of total ozone from the TOMS instrument - A report on data quality.*, Geophysical Research Letters, **13**, no. 12, 1355 – 1358.

Fleig A.J., P.K. Bhartia and D.S. Silberstein (1986 b), *An assessment of the long-term drift in SBUV total ozone data, based on comparison with the dobson network.*, Geophysical Research Letters, **13**, no. 12, 1359 – 1362.

Garcia R.R. and S. Solomon (1983), *A numerical model of the zonally averaged dynamical and chemical structure of the middle atmosphere.*, Journal of Geophysical Research, **88**, no. C2, 1379 – 1400.

Garcia R.R. and S. Solomon (1987), *A possible relationship between interannual variability in Antarctic ozone and the quasi-biennial oscillation.*, Geophysical Research Letters, **14**, 848.

Garcia R.R., F. Stordal, S. Solomon and J.T. Kiehl (1992), *A new numerical model of the middle atmosphere 1. Dynamics and transport of tropospheric source gases.*, Journal of Geophysical Research, **97**, no. D12, 12967 – 12991.

Ghazi A., A. Ebel and D.F. Heath (1976) *A study of satellite observations of ozone and stratospheric temperatures during 1970-1971.*, Journal of Geophysical Research, **81**, no. 30, 5365 - 5373.

Gleason J.F., P.K. Bhartia, J.R. Herman, R. McPeters, P. Newman, R.S. Stolarski, L. Flynn, G. Labow, D. Larko, C. Seftor, C. Wellemeyer, W.D. Komhyr, A.J. Miller and W. Planet (1993), *Record low global ozone in 1992.*, Science, **260**, 523 - 526.

Gray L.J. and J.A. Pyle (1989), *A two-dimensional model of the quasi-biennial oscillation of ozone.*, Journal of the Atmospheric Sciences, **46**, no. 2, 203 - 220.

Gray L.J. and S. Ruth (1993), *The modeled latitudinal distribution of the ozone quasi-biennial oscillation using observed equatorial winds.*, Journal of the Atmospheric Sciences, **50**, no. 8, 1033 - 1046.

Hare F.K. and B.W. Boville (1965), *The polar circulation.*, Technical note 70, World Meteorological organization, Geneva, Switzerland.

Hartmann D.L. and R.R. Garcia (1979), *A mechanistic model of ozone transport by planetary waves in the stratosphere.*, Journal of the Atmospheric Sciences, **36**, 350 - 364.

Hartmann D.L., L.E. Heidt, M. Loewenstein, J.R. Podoloske, J. Vedder, W.L. Starr and S.E. Strahan (1989), *Transport into the polar vortex in early spring.*, Journal of Geophysical Research, **94**, no. D14, 16779 - 16795.

Hayashi Y. and D.G. Golder (1993), *Tropical 40-50- and 25-30-day oscillations appearing in realistic and idealized GFDL climate models and the ECMWF dataset.*, Journal of the Atmospheric Sciences, **50**, no. 3, 464 - 494.

Heath D.F., A.J. Krueger, H.A. Roeder and B.D. Henderson (1975), *The Solar Backscatter and Total Ozone Mapping Spectrometer (SBUV/ TOMS) for NIMBUS G.*, Optical Engineering, **14**, 323 - 331.

Heath D.F. (1988), *Non-seasonal changes in total column ozone from satellite observations, 1970 - 1986.*, Nature, **332**, 219 - 227.

Henderson-Sellers A. and P.J. Robinson (1987), *Contemporary climatology*, Longman Scientific & Technical, 220 - 222.

Herman J.R., R.D. Hudson and G. Serafino (1990), *Analysis of the eight-year*

*trend in ozone depletion from empirical models of solar backscattered ultraviolet instrument degradation.* Journal of Geophysical Research, **95**, no. D6, 7403 – 7416.

Herman J.R., R.D. Hudson, R. McPeters, R. Stolarski, Z. Ahmad, X.-Y. Gu, S. Taylor and C. Wellemeyer (1991), *A new self-calibration method applied to TOMS and SBUV backscattered ultraviolet radiation data to determine long-term global ozone change.*, Journal of Geophysical Research, **96**, no. D4, 7531 – 7545.

Herman J.R., R.D. McPeters and D. Larko (1993), *Ozone depletion at Northern and Southern latitudes derived from January 1979 to December 1991 Total Ozone Mapping Spectrometer Data.*, Journal of Geophysical Research, **98**, no. D7, 12783 – 12793.

Hofmann D.J. and S. Solomon (1989), *Ozone destruction through heterogeneous chemistry following the eruption of El Chichón.*, Journla of Geophysical Research, **94**, no. D4, 5029 – 5041.

Hofmann D.J. and T. Deshler (1991), *Stratospheric cloud observations during formation of the Antarctic ozone hole in 1989.*, Journal of Geophysical Research, **96**, no. D2, 2897 – 2912.

Holton J.R. and R.S. Lindzen (1972), *An updated theory for the quasi-biennial cycle of the tropical stratosphere.*, Journal of the Atmospheric Sciences, **29**, 1076 – 1080.

Holton J.R. and H.-C. Tan (1980), *The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb.*, Journal of the Atmospheric Sciences, **37**, 2200 – 2208.

Holton J.R. and H.-C. Tan (1982) *The quasi-biennial oscillation in the Northern Hemisphere lower stratosphere.*, Journal of the Meteorological Society of Japan, **60**, 140 – 148.

Holton J.R. (1992), *An Introduction to Dynamic Meteorology.*, 3 rd edition, Academic Press.

Hood L.L. and J.P. McCormack (1992), *Components of interannual ozone change based on Nimbus 7 TOMS data.*, Geophysical Research Letters, **19**, no. 23, 2309 – 2312.

Houghton J.T. (1989), *The physics of atmospheres.*, 2nd edition, Cambridge Uni-

versity Press.

Hudson R.D. and J. Kim (1993), *Direct measurements of tropospheric ozone using TOMS data.*, 'Proceedings Quadrennial Ozone Symposium, 1992', in press, 4 – 13.

Iqbal M. (1983), *An Introduction to Solar Radiation*, Academic Press, Toronto, Canada.

Juckes M.N. and M.E. McIntyre (1987), *A high-resolution one-layer model of breaking planetary waves in the stratosphere.*, Nature, **328**, 590 – 596.

Kiehl J.T., B.A Boville and B.P. Briegleb (1988), *Response of a general circulation model to a prescribed Antarctic ozone hole.*, Nature, **332**, 501 – 504.

Klenk K.F., P.K. Bhartia, A.J. Fleig, V.G. Kaveeshwar, R.D. McPeters and P.M. Smith (1982), *Total ozone determination from the backscattered ultraviolet (BUV) experiment.*, Journal of Applied Meteorology, **21**, 1672 – 1684.

Klenk K.F., P.K. Bhartia, A.J. Fleig and C.L. Mateer (1983), *Vertical ozone profile determination from Nimbus-7 SBUV measurements*, paper presented at the Fifth Conference on Atmospheric Radiation, Baltimore MD, Oct 31 - Nov 4.

Kohno J. (1984), *Stratospheric ozone transport due to transient large amplitude planetary waves.*, Journal of the Meteorological Society of Japan, **62**, no. 3, 413 – 439.

Kondo Y., P. Aimedieu, M. Koike, Y. Iwasaka, P.A. Newman, U. Schmidt, W.A. Matthews, M. Hayashi and W.R. Sheldon (1992), *Reactive nitrogen, ozone and nitrate aerosols observed in the Arctic stratosphere in January 1990.*, Journal of Geophysical Research, **97**, no. D12, 13025 – 13038.

Krueger A.J., M.R. Schoeberl, R.S. Stolarski and F.S. Sechrist (1988), *The 1987 Antarctic ozone hole: A new record low.*, Geophysical Research Letters, **15**, no. 12, 1365 – 1368.

Krueger A.J., R.S. Stolarski and M.R. Schoeberl (1989), *Formation of the 1988 Antarctic ozone hole.*, Geophysical Research Letters, **16**, no. 5, 381 – 384.

Krueger A.J., M.R. Schoeberl, P. Newman, R.S. Stolarski (1992), *The 1991 Antarctic ozone hole: TOMS observations.*, Geophysical Research Letters, **19**, no. 12, 1215 – 1218.

Labitzke K. (1982), *On the interannual variability of the middle stratosphere during the northern winters.*, Journal of the Meteorological Society of Japan, **60**, 124 – 139.

Lait L.R. and J.L. Stanford (1988), *Applications of asynoptic Space-time Fourier Transform methods to scanning satellite measurements.*, Journal of the Atmospheric Sciences, **45**, no. 24, 3784 – 3799.

Leovy C.B., C-R. Sun, M. Hitchman, E.E. Remsberg, J.M. Russell III, L.L Gordley, J.C. Gille and L.V. Lyjak (1985), *Transport of ozone in the middle stratosphere: evidence for planetary wave breaking.*, Journal of the Atmospheric Sciences, **42**, no. 3, 230 – 244.

Levy H., J.D. Mahlman and W.J. Moxim (1979), *A preliminary report on the numerical simulation of the three-dimensional structure and variability of atmospheric $N_2O$.*, Geophysical Research Letters, **6**, 155.

Lighthill J. (1980), *Waves in fluids.*, Cambridge University Press.

Lindzen R.S. and J.R. Holton (1968), *A theory of the quasi-biennial oscillation.*, Journal of the Atmospheric Sciences, **25**, 1095 – 1107.

Mahlman J.D. and S.B. Fels (1986), *Antarctic ozone decrease: A dynamical cause?*, Geophysical Research Letters, **13**, no. 12, 1316 – 1319.

Mancini E., G. Visconti, G. Pitari and M. Verdecchia (1991), *An estimate of the Antarctic ozone modulation by the QBO.*, Geophysical Research Letters, **18**, no. 2, 175 – 178.

Manney G.L., J.D. Farrara and C.R. Mechoso (1991), *The behaviour of wave 2 in the southern hemisphere stratosphere during late winter and early spring.*, Journal of the Atmospheric Sciences, **48**, no. 7, 976 – 998.

Matsuno T. (1980), *Lagrangian motion of air parcels in the stratosphere in the presence of planetary waves.*, Pure and Applied Geophysics, **118**, 189 – 216.

McElroy M.B., R.J. Salawitch, S.C. Wofsky and J.A. Logan (1986), *Reductions of Antarctic ozone due to synergistic interactions of chlorine and bromine.*, Nature, **321**, 759 – 762.

McFarlane N.A. (1987), *The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere.*, Journal of the

Atmospheric Sciences, **44**, 1775 – 1800.

McIntyre M.E. and T.N. Palmer (1984), *The 'surf zone' in the stratosphere.*, Journal of Atmospheric and Terrestrial Physics, **46**, no. 9, 825 – 849.

McKenzie R.L. (1991), *Application of a simple model to calculate latitudinal and hemispheric differences in ultraviolet radiation.*, Weather and Climate, **11**, 3 – 14.

McPeters R.D., R.D. Hudson, P.K. Bhartia and S.L. Taylor (1986), *The vertical ozone distribution in the Antarctica ozone minimum by SBUV.*, Geophysical Research Letters, **13**, no. 12, 1213 – 1216.

McPeters R.D. and the NASA Ozone Processing Team (1990), *Results of recalibration of Nimbus 7 TOMS ozone data using the pair justification technique*, Digest of Topical Meeting on Optical Remote Sensing of the Atmosphere, **4**, 413 – 416.

McPeters R.D. and W.D. Komhyr (1991), *Long-term changes in SBUV/TOMS relative to world primary standard dobson spectrometer 83*, Journal of Geophysical Research, in press.

Newman P.A. (1986), *The final warming and polar vortex disappearance during the southern hemisphere spring.*, Geophysical Research Letters, **13**, no. 12, 1228 – 1231.

Newman P.A. and W.J. Randel (1988), *Coherent ozone-dynamical changes during the southern hemisphere spring, 1979-1986.*, Journal of Geophysical Research, **93**, no. D10, 12585 – 12606.

Newman P.A., Stolarski R., Schoeberl M.R., McPeters R.D., and A.J. Krueger (1991), *The 1990 Antarctic ozone hole as seen by TOMS.*, Geophysical Research Letters, **18**, no. 4, 661 – 664.

Niu X., J.E. Frederick, M.L. Stein and G.C. Tiao (1992), *Trends in column ozone based on TOMS Data: Dependence on Month, Latitude and Longitude.*, Journal of Geophysical Research, **97**, no. D13, 14661 – 14669.

Oltmans S.J. (1985), *Tropospheric ozone at four remote observatories.*, In 'Atmospheric Ozone', edited by C.S. Zerefos and A. Ghazi, 769 – 802, Reidel, Dordrecht.

Pèdlosky (1979), *Geophysical Fluid Dynamics.*, Springer-Verlag, 2nd edition.

Penner J.E. and Chang J.S. (1978), *Possible variations in atmospheric ozone related to the eleven-year solar cycle.*, Geophysical Research Letters, **5**, 817 - 820.

Perliski L.M., S Solomon and J. London (1989), *On the interpretation of seasonal variations of stratospheric ozone.*, Planetary Space Science, **37**, no. 12, 1527 - 1538.

Pitari G., S. Palermi, G. Visconti and R.G. Prinn (1992), *Ozone response to a $CO_2$ doubling: results from a stratospheric circulation model with heterogeneous chemistry.*, Journal of Geophysical Research, **97**, no. D5, 5953 - 5962.

Plumb R. and Ko M. (1992), *Interrelationships between mixing ratios of long-lived stratospheric constituents.*, Journal of Geophysical Research, **97**, no. D9, 10145 - 10156.

Podolske J.R., M. Loewenstein, S.E. Strahan and K.R. Chan (1989), *Stratospheric nitrous oxide distribution in the southern hemisphere.*, Journal of Geophysical Research, **94**, no. D14, 16767 - 16772.

Prather M. (1992), *Catastrophic loss of stratospheric ozone in dense volcanic clouds.*, Journal of Geophysical Research, **97**, no. D9, 10187 - 10191.

Press W.H., B.R. Flannery, S.A. Teukolsky and W.T. Vettering (1989), *Numerical recipes in Pascal.*, Cambridge University Press.

Pyle J.A. and C.F. Rogers (1980), *Stratospheric transport by stationary planetary waves - the importance of chemical processes.*, Quarterly Journal of the Royal Meteorological Society, **106**, 421 - 446.

Randel W.J. (1987), *A study of planetary waves in the southern winter troposphere and stratosphere. Part I: Wave structure and vertical propagation.*, Journal of the Atmospheric Sciences, **44**, no. 6, 917 - 935.

Randel W.J., D.E. Stevens and J.L. Stanford (1987), *A study of planetary waves in the southern winter troposphere and stratosphere. Part II: Life cycles.*, Journal of the Atmospheric Sciences, **44**, no. 6, 936 - 949.

Randel W.J. (1992), *Global atmospheric circulation statistics 1000 - 1 mb.*, NCAR technical note NCAR/TN-366 + STR.

Randel W.J. (1993 a), *Ideas flow on Antarctic vortex.*, Nature, **364**, 105 - 106.

Randel W.J. (1993 b), *Global Normal-Mode Rossby Waves Observed in Stratospheric Ozone Data.*, Journal of the Atmospheric Sciences, **50**, no. 3, 406 – 420.

Randel W.J., J.C. Gille, A.E. Roche, J.B. Kumer, J.L. Mergenthaler, J.W. Waters, E.F. Fishbein and W.A. Lahoz (1993), *Stratospheric transport from the tropics to middle latitudes by planetary-wave mixing.*, Nature, **365**, 533 – 535.

Rasch P.J. and D.L Williamson (1991), *Sensitivity of a general circulation model climate to the moisture transport formulation.*, Journal of Geophysical Research, **96**, 13123 – 13137.

Reed R.J. (1950), *The role of vertical motions in ozone weather relationships.*, Journal of Meteorology, **7**, 263 – 267.

Reed R.J. (1964), *A tentative model of the 26-month oscillation in tropical latitudes.*, Quarterly Journal of the Royal Meteorological Society, **90**, 441 – 466.

Reed R.J. and K.E. German (1965), *A contribution to the problem of stratospheric diffusion by large scale mixing.*, Monthly Weather Review, **93**, 313.

Reinsel G.C., G.C. Tiao, S.K. An, M. Pugh, S. Basu, J.J. DeLuisi, C.L. Mateer, A.J. Miller, P.S. Connell and D.J. Wuebbles (1988), *An analysis of the 7-year record of SBUV satellite ozone data: Global profile features and trends in total ozone.*, Journal of Geophysical Research, **93**, 1689 – 1703.

Rodgers C.D. (1976), *Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation.*, Reviews of Geophysics and Space Physics, **14**, 609 – 624.

Rood R.B. (1986), *Global ozone minima in the historical record.*, Geophysical Research Letters., **13**, 1244 – 1247.

Rowland F.S. and M.J. Molina (1975), *Chlorofluoromethanes in the Environment.*, Reviews of Geophysics and Space Science, **13**.

Salby M.L. and R.R. Garcia (1990), *Dynamical perturbations to the ozone layer.*, Physics Today, **43**, no. 3, 38 – 46.

Salby M.L. and P.F. Callaghan (1993), *Fluctuations of total ozone and their relationship to stratospheric air motions.*, Journal of Geophysical Research, **98**, no. D2, 2715 – 2727.

Schmidt M. (1982), *The influence of large scale advection on the vertical distribu-*

*tion of stratospheric source gases in 44° and 41° North.*, Journal of Geophysical Research, **87**, 11239.

Schoeberl M.R. and M.A. Geller (1976), *The structure of stationary planetary waves in winter in relation to the polar night jet intensity.*, Geophysical Research Letters, **3**, no. 3, 177 – 180.

Schoeberl M.R. and A.J. Krueger (1983), *Medium scale disturbances in total ozone during southern hemisphere summer.*, Bulletin of the American Meteorological Society, **64**, 1358.

Schoeberl M.R., A.J. Krueger and P.A. Newman (1986), *The morphology of Antarctic total ozone as seen by TOMS.*, Geophysical Research Letters, **13**, no. 12, 1217 – 1220.

Schoeberl M.R., R.S. Stolarski and A.J. Krueger (1989), *The 1988 Antarctic ozone depletion: comparison with previous year depletions.*, Geophysical Research Letters, **16**, no. 15, 377 – 380.

Schoeberl M.R. and D.L. Hartmann (1991), *The dynamics of the stratospheric polar vortex and its relation to springtime ozone depletions.*, Science, **251**, 46 – 52.

Schoeberl M.R., L.R. Lait, P.A. Newman and J.E. Rosenfield (1992) *The structure of the polar vortex.*, Journal of Geophysical Research, **97**, no. D8, 7859 – 7882.

Shiotani M. and J.C. Gille (1987), *Dynamical factors affecting ozone mixing ratios in the Antarctic lower stratosphere.*, Journal of Geophysical Research, **92**, no. D8, 9811 – 9824.

Shiotani M. (1992), *Annual, quasi-biennial, and El Niño-southern oscillation (ENSO) time-scale variations in equatorial ozone.*, Journal of Geophysical Research, **97**, no. D7, 7625 – 7633.

Solomon S. and R.R. Garcia (1984), *On the distribution of long-lived tracers and chlorine species in the middle atmosphere.*, Journal of Geophysical Research, **89**, no. D7, 11633 – 11644.

Solomon S., R.R. Garcia and F. Stordal (1985), *Transport Processes and Ozone Perturbations.*, Journal of Geophysical Research, **90**, no. D7, 12981 – 12989.

Solomon S., R.R. Garcia, F.S. Rowland and D.J. Wuebbles (1986), *On the deple-*

*tion of Antarctic ozone.*, Nature, **321**, 755 - 758.

Solomon S. (1990), *Progress towards a quantitative understanding of Antarctic ozone depletion.*, Nature, **347**, 347 - 354.

Solomon S., J.P. Smith, R.W. Sanders, L. Perliski, H.L Miller, G.H Mount, J.G. Keys and A.L. Schmeltkopf (1993), *Visible and Near-Ultraviolet Spectroscopy at McMurdo Station, Antarctica 8. Observations of Nighttime $NO_2$ and $NO_3$ From April to October 1991.*, Journal of Geophysical Research, **98**, no. D1, 993 - 1000.

Stephenson J.A.E. and M.W.J. Scourfield (1991), *Importance of energetic solar protons in ozone depletion.*, Nature, **352**, 137 - 139.

Stolarski R.S. (1988), *The Antarctic ozone hole.*, Scientific American, **258**, 20 - 26.

Stolarski R.S., M.R Schoeberl, P.A. Newman, R.D. McPeters and A.J. Krueger (1990), *The 1989 Antarctic ozone hole as observed by TOMS.*, Geophysical Research Letters, **17**, no. 9, 1267 - 1270.

Stolarski R.S., P. Bloomfield, R.D. McPeters and J.R. Herman (1991), *Total ozone trends deduced from Nimbus 7 TOMS data.*, Geophysical Research Letters, **18**, no. 6, 1015 - 1018.

Sze N.D., M.K.W. Ko, D.K. Weisenstein, J.M. Rodriguez, R.S. Stolarski and M.R. Schoeberl (1989), *Antarctic ozone hole: Possible implications for ozone trends in the Southern hemisphere.*, Journal of Geophysical Research, **94**, no. D9, 11521 - 11528.

Thompson A.M., R.D. Diab, G.E. Bodeker, M. Zunckel, G. Coetzee, C.B. Archer, D. McNamara, K. Pickering, E. Mravlag, J. Barsby, F. Sokolic, J. Fishman and D. Nganga (1994), *Total ozone over Southern Africa during SAFARI-92/TRACE-A*, Journal of Geophysical Research, in preparation.

Tolson R.H. (1981), *Spatial and temporal variations of monthly mean total columnar ozone derived from 7 years of BUV data.*, Journal of Geophysical Research, **86**, 7312.

Toon O.B., P. Hamill, R.P. Turco and J. Pinto (1986), *Condensation of $HNO_3$ and HCl in the winter polar stratospheres.*, Geophysical Research Letters, **13**, no. 12, 1284 - 1287.

Toon O.B. and R.P. Turco (1991), *Polar stratospheric clouds and ozone depletion.*, Scientific American, **287**, 40 – 47.

Torres O., Z. Ahmad and J.R. Herman (1992), *Optical effects of polar stratospheric clouds on the retrieval of TOMS total ozone.*, Journal of Geophysical Research, **97**, no. D12, 13015 – 13024.

Toumi R., R.L. Jones and J.A. Pyle (1993), *Stratospheric ozone depletion by $ClONO_2$ photolysis.*, Nature, **365**, 37 – 39.

Tuck A.F. (1989), *Synoptic and chemical evolution of the Antarctic vortex in late winter and early spring, 1987.*, Journal of Geophysical Research, **94**, 11687 – 11737.

Tung K.K., M.W.K. Ko, J.M. Rodriguez and N.D. Sze (1986), *Are Antarctic ozone variations a manifestation of dynamics or chemistry?*, Nature, **322**, 811 – 814.

Tung K.K. and H. Yang (1988), *Dynamic variability of column ozone.*, Journal of Geophysical Research, **93**, 11123 – 11128.

van Loon H. and R.L. Jenne (1972), *The zonal harmonic standing waves in the Southern Hemisphere.*, Journal of Geophysical Research, **77**, 3846 – 3855.

Volz A. and D. Kley (1988), *Evaluation of the Montsouris series of ozone measurements made in the nineteenth century.*, Nature, **332**, 240 – 242.

Wahner A. and C. Schiller (1992), *Twilight variation of vertical column abundances of OClO and BrO in the North Polar region.*, Journal of Geophysical Research, **97**, no. D8, 8047 – 8055.

Waters J.W., L. Froidevaux, W.G. Read, G.L. Manney, L.S. Elson, D.A. Flower, R.F. Jarnot and R.S. Harwood (1993), *Stratospheric ClO and ozone from the Microwave Limb Sounder on the Upper Atmospheric Research Satellite.*, Nature **362**, 597 – 602.

Watson R.T. (1988), *Stratospheric ozone depletion: Antarctic processes. Ozone depletion, greenhouse gases, and climate change.*, Proceedings of a joint symposium by the board on atmospheric sciences and climate and the committee on global change, National Academy Press.

Wellemeyer C.G., S.L. Taylor, C.J. Seftor and R.D. McPeters (1993), *TOMS*

*profile shape error estimates at high latitude.*, Paper presented at 'Atmospheric studies by optical methods' conference, Tromso.

Wirth V. (1993), *Quasi-stationary planetary waves in total ozone and their correlation with lower stratospheric temperature*, Journal of Geophysical Research, **98**, no. D5, 8873 – 8882.

WMO report number 20 (1990), **1**

Wuebbles D.J., F.M. Luther, and J.E. Penner (1983), *Effect of coupled anthropogenic perturbations on stratospheric ozone.*, Journal of Geophysical Research, **88**, no. C2, 1444 – 1456.

Zerefos C.S., A.F. Bais, I.C. Ziomas and R.D. Bojkov (1992), *On the relative importance of quasi-biennial oscillation and El Niño southern oscillation in the revised Dobson total ozone records.*, Journal of Geophysical Research, **97**, no. D9, 10135 – 10144.