

The Role of Gravity Waves in Coupled Middle–Upper Atmosphere Dynamics

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

Scott Leslie England
Department of Physics and Astronomy
University of Leicester

April 2005

BEST COPY

AVAILABLE

PAGE
NUMBERING
AS ORIGINAL

Abstract

The Role of Gravity Waves in Coupled Middle–Upper Atmosphere Dynamics

Scott England

Internal gravity waves which propagate upwards from the lower atmosphere transport energy and momentum into the middle and upper atmosphere. As these waves dissipate, they alter the dynamics and composition of the background atmosphere. Changes in the background also affect the dissipation of these waves. Through this interplay, gravity waves couple the lower and upper atmosphere. Recent improvements in our understanding of gravity waves have allowed detailed physical representations of these waves to be included in global-scale atmospheric models. Three models have been developed for this study which are used to examine the role of gravity waves in the dynamics of the middle and upper atmosphere over a range of time-scales.

The effects of the solar sunspot cycle on gravity waves in the middle atmosphere are explored. Changes in the wave-induced wind variance of up to 10 % are simulated in the mesosphere. It is shown that a strong change in the tropospheric gravity-wave source could explain observed variability in gravity waves in this region, but changes in the mesospheric winds also play a crucial role in the model results.

Varying the amplitude of the lower atmosphere gravity-wave source affects gravity-wave drag and the residual circulation. These are shown to influence conditions in the stratosphere and can alter the strength of sudden warmings. This has important consequences for the whole atmosphere.

The interaction between gravity wave-drag and the diurnal tide is investigated using a coupled middle atmosphere and thermosphere model. This is shown to have a strong influence on dynamics and airglow emissions in the equatorial MLT region. Further constraint of the lower-atmosphere gravity waves and time dependent propagation effects are needed to improve our understanding of the role which gravity waves play in the dynamical coupling of the atmosphere.

ACKNOWLEDGEMENTS

There are a many people without whom this thesis would not exist. My deepest thanks must go to my supervisor, Neil, for giving me a great deal of his time, enthusiasm and assistance over the last three years. My thanks have to go to the whole of the RSPP group, both research and technical, for their support and encouragement. The people from APL, especially Alison, Matt and Alan, have also been instrumental to the work with the CMAT model included in this thesis, and I thank them for all they have done.

Of course, it's been my officemates who have had to put up with me through every day of this work. My thanks must be extended to Ranvir, whose extensive knowledge of all things UNIX has helped me out on many occasions; to Natalie, for importing the many fine wines and been custodian of the very best of my white-board doodles; and Mr Tommi, who has been a great housemate and officemate, and connoisseur of fine pizza and good curry.

Of my other housemates over the years, I'd like to thank Mr Allen for enduring more than his fair share of my ramblings; Jonny for being a great friend, both in Halls and in Ashfield Road; and Evil Fraz for his humour and party spirit. I'd also like to thank Cossie, for persuading me to study Italian, go to the gym and for some fabulous dinner parties.

I'd like to say cheers to Jim and Ather for being great mentors, to Adriana for teaching me Italian and to the whole of Leicester Writers for taking me under their wing.

Hugs have to go to Sarah, for showing me just how good the good times get; to Kate for being outgoing, adventurous and generally lovely; to Maria for her passion for the arts, wonderful conversation, and trying keeping me sane throughout; and to Amie for her endless encouragement and for teaching me to cook some fantastic food.

Last, but most certainly not least, I've got to thanks Flimsy, Father Bear, Banana Woman, Mr Melon, Fraggie, Michael Cat, Alfonse and Flaw Claw for taking good care of me and helping me through it all.

Table of Contents

TABLE OF CONTENTS

List of Common Symbols Used	v
Chapter I Background Theory	1
I.1 Basic Atmospheric Structure	2
I.2 Atmospheric Thermodynamics	4
I.2.1 The Ideal Gas Law	4
I.2.2 Hydrostatic Balance	5
I.2.3 Scale Height	5
I.2.4 Adiabatic Lapse Rate	5
I.2.5 Brunt-Väisälä Frequency	6
I.2.6 Potential Temperature	6
I.2.7 Thermal Conduction	7
I.3 Atmospheric Dynamics	7
I.3.1 The Primitive Equations on a Beta-Plane	8
I.3.2 Geostrophic Balance and Thermal Wind	9
I.3.3 The Transformed Eulerian-Mean Equations	10
I.3.4 Residual Circulation and the Mesopause Anomaly	11
I.4 Atmospheric Radiation	12
I.4.1 The Solar Spectrum and Ultraviolet Absorption	12
I.4.2 Infrared Emission and Local Thermodynamic Equilibrium	13
I.5 Atmospheric Waves	14
I.5.1 Atmospheric Tides	14
I.5.2 Planetary Waves and the Eliassen-Palm Flux	16
I.5.3 Atmospheric Gravity Waves	18
I.5.4 Kelvin Waves	18
I.6 The Ionosphere	19
Chapter II Atmospheric Gravity Waves and Their Representation In Global Scale Models	
II.1 Linear Gravity Wave Theory	30
II.1.1 High-Frequency Waves	34
II.1.2 Medium-Frequency Waves	35
II.1.3 Low-Frequency Waves	36
II.1.4 Conservative Propagation and Dissipation	37
II.2 Gravity Wave Sources	38
II.3 Observations	39
II.3.1 Observational Techniques	40

Table of Contents

II.3.1.1 Mesospheric, Stratospheric and Tropospheric (MST) Radars	40
II.3.1.2 Partial-Reflection Radars	40
II.3.1.3 Ionospheric D1 Technique	41
II.3.1.4 Meteor Radars	41
II.3.1.5 Lidars	41
II.3.1.6 Rockets	42
II.3.1.7 Balloons	42
II.3.1.8 Aircraft	42
II.3.1.9 Optical Measurements	42
II.3.1.10 Spacecraft	43
II.3.2 Observed Gravity-Wave Spectra	43
II.4 Gravity Wave Parameterisations	45
II.4.1 Rayleigh Friction	46
II.4.2 Linear Saturation, Lindzen, 1981	46
II.4.3 Atmospheric Transmission, Matsuno, 1982	48
II.4.4 Semi-Empirical, Fritts and Lu, 1993, Warner and McIntyre, 2001	49
II.4.5 Doppler Spread, Hines, 1997a,b	49
II.4.6 Nonlinear Diffusion, Medvedev and Klaassen, 2000	50
Chapter III The Stratosphere to Thermosphere Energy Variability Experiment (STEVE) Model	
III.1 Global Scale Atmospheric Models	55
III.2 The UK Meteorological Office Stratosphere Mesosphere Model	56
III.3 The STEVE Model	57
III.3.1 The Model Grid	57
III.3.2 Ultra-Violet Absorption	58
III.3.3 Infra-Red Absorption and Emission	60
III.3.4 Gravity-Wave Drag	61
III.3.5 Diffusion	62
III.4 The STEVE Model Climatologies	63
III.4.1 Zonal Mean Temperatures	63
III.4.2 Zonal Mean Wind Structure	64
III.4.3 Zonal Mean Meridional Circulation	66
III.4.4 Diabatic Heating Rates	67
III.4.5 Gravity-Wave Drag, Wind Variances and Eddy Diffusion	69
III.5 The STEVE-3D Model	72
III.6 The STEVE-3D Model Climatologies	73

Table of Contents

Chapter IV Modelling the Effect of the Solar Cycle on Gravity-Wave Activity in the Middle Atmosphere	103
IV.1 Modelling the Effect of the Annual Cycle on Gravity-Wave Activity in the Middle Atmosphere	104
IV.1.1 Observations of the Effect of the Annual Cycle on Gravity-Wave Activity in the Middle Atmosphere	104
IV.1.2 Theories of the Effect of the Annual Cycle on Gravity-Wave Activity in the Middle Atmosphere	105
IV.1.3 STEVE Model Results of the E Effect of the Annual Cycle on Gravity-Wave Activity in the Middle Atmosphere	107
IV.2 Modelling the Effect of Variable Solar UV Flux on Gravity-Wave Activity in the Middle Atmosphere	110
IV.2.1 Observations of the Effect of the Solar Cycle on Gravity-Wave Activity in the Middle Atmosphere	111
IV.2.2 Theories of the Observed Effect of the Solar Cycle on Gravity-Wave Activity in the Middle Atmosphere	113
IV.2.3 STEVE Model Results of the Effect of Variable Solar UV Flux on Gravity-Wave Activity in the Middle Atmosphere	114
IV.3 Exploring a Possible Mechanism for a Solar Cycle in Gravity Wave Activity in the MLT Region	119
IV.3.1 The Impact of Changes in the Middle Atmosphere	120
IV.3.2 The Impact of Changes in the Lower Atmosphere	124
IV.4 Summary	125
Chapter V Sudden Warmings in the Northern Hemisphere Winter Polar Stratosphere: Dependence on the Lower-Atmosphere Gravity-Wave Source Spectral Amplitude	
V.1 Sudden Stratospheric Warmings	146
V.1.1 Observed Properties of SSW	146
V.1.2 The Matsuno Theory of SSW	147
V.2 Previous Studies into the Effect of Gravity-Wave Drag on SSW	149
V.3 Effect of Variable Gravity-Wave Sources on SSW in the STEVE-3D Model	151
V.3.1 A SSW During January in the STEVE-3D Model	152
V.3.2 Theory of the Effect of Gravity-Wave Drag on SSW	152
V.3.3 The Effect of Gravity-Wave Drag on SSW in the STEVE-3D Model	153
V.3.3.1 STEVE-3D Model Configuration	153
V.3.3.2 STEVE-3D Model Results and Discussion	155

Table of Contents

V.4 Summary	158
Chapter VI The Effect of Gravity-Wave Drag on the Equatorial Diurnal Tide and Airglow Emissions in the CMAT Model	
VI.1 Introduction	171
VI.1.1 The Propagating Diurnal Tide in the Equatorial MLT Region	171
VI.1.2 The Atomic-Oxygen Greenline Airglow Emission	172
VI.1.3 The UARS Wind Imaging Interferometer (WINDII)	173
VI.2 Previous Modelling and Observational Studies of the Interaction Between Gravity Waves and Atmospheric Tides in the MLT Region	173
VI.3 The Coupled Middle Atmosphere and Thermosphere General Circulation Model	176
VI.3.1 CMAT Model Description	176
VI.3.2 CMAT Model Configuration	177
VI.3.3 Configuration of the Gravity-Wave Parameterisations Used in the CMAT Model	178
VI.4 CMAT Model Results and Discussion	179
VI.4.1 The Effect of Parameterised Gravity-Wave Drag on the Zonal Mean Zonal Winds	179
VI.4.2 The Effect of Parameterised Gravity-Wave Drag on the Meridional Winds	180
VI.4.3 The Effect of Parameterised Gravity-Wave Drag on Atomic-Oxygen Greenline Emission Rates	183
VI.5 Summary	184
Chapter VII Summary and Future Work	
VII.1 Summary	194
VII.2 Future Work	198
Appendix I Downward Control Theory	205
Appendix II Causality in Middle Atmosphere Dynamics	208
Appendix III Chapman Chemistry	209
References	211

List of Symbols

LIST OF COMMON SYMBOLS USED

c_p	phase speed in the Eulerian frame of reference
\hat{c}_p	phase speed in the Lagrangian frame of reference
c_g	group velocity in the Eulerian frame of reference
f	Coriolis parameter
\mathbf{F}	Eliassen-Palm flux
g	acceleration due to gravity
H	pressure scale height
k	zonal wave number
l	meridional wave number
m	vertical wave number
N	Brunt-Väisälä frequency
p	pressure
Q	diabatic heating rate
T	temperature
u	zonal component of the background wind
v	meridional component of the background wind
$\overline{v^*}$	meridional component of the residual circulation
V	volume
w	vertical component of the background wind
$\overline{w^*}$	vertical component of the residual circulation
X	zonal component of frictional forces
Y	meridional component of frictional forces
z	log pressure height
ϕ	latitude
θ	potential temperature
ρ	density
ω	frequency in Eulerian frame of reference
$\hat{\omega}$	frequency in the Lagrangian frame of reference
Ω	rotation rate

CHAPTER I BACKGROUND THEORY

Over recent years there has been growing public, scientific and political interest in observed changes in the Earth's climate system and the impact which these may have on our lives. The Intergovernmental Panel on Climate Change (IPCC) have highlighted the need to understand the natural climate variability, in order to assess the anthropogenic impacts and to help predict future climate conditions. While some of the processes are considered to be understood by the scientific community, the IPCC 2001 report concluded that the influence of the 11-year solar sunspot cycle on the Earth's climate system was among the least well understood. For this reason the study of how solar variability may affect climate is both important and topical.

Perhaps the best known example of what is believed to be the impact of solar variability on society is the so-called Little Ice Age, which occurred during the in the late 17th and early 18th centuries. Eddy (1976) linked this unusually cold period in Europe to the prolonged absence of observed sunspots reported by Maunder (1894).

While changes in the lower atmosphere may be the most important to society, the most pronounced effects of solar variability are seen in the upper atmosphere. Understanding the coupling between different atmospheric layers and the wider solar-terrestrial system is necessary to further our understanding how the relatively small changes in the total solar energy output can be amplified to produce the observed changes in the lower atmosphere. The middle atmosphere plays a vital role in the coupling and as such will be the main focus of this thesis.

This study will concentrate on the dynamical coupling between different layers of the atmosphere. A key part of this is played by atmospheric gravity waves. Many of these waves are created in the lower atmosphere and propagate up into the middle and upper atmosphere. As these waves dissipate, they affect local wind and temperature conditions, create small-scale turbulence which leads to mixing of the atmosphere and, on a larger scale, drive the global-scale transport of energy and chemical constituents in the mesosphere (Murgatroyd and Singleton, 1961). It is through this circulation and interactions with other atmospheric waves, that gravity waves couple the variability in the middle and upper atmosphere to the lower atmosphere.

Throughout this study, global-scale atmospheric models will be used to quantify the interaction between gravity waves and other components of the atmosphere. Until recently, it was not possible to include a detailed physical representation of these waves in global-scale atmospheric models. Chapter (II) will outline linear gravity wave

theory and introduce the gravity-wave parameterisations which have allowed this representation to be improved. Chapter (III) will introduce two new models which have been developed for this thesis and Chapter (VI) will detail improvements which have been made to a third in order to study gravity-wave interactions with the middle and upper atmosphere in more detail. By comparing the model results to observations and other theoretical work, the improvements in the atmospheric models made by, as well as the limitations of, these parameterisations will be highlighted.

The interplay between gravity waves and the middle atmosphere can be considered over a wide variety of time-scales. Over long time-scales, such as the 11-year solar sunspot cycle, gravity waves interact with the slowly varying background atmosphere. Chapter (IV) of this thesis will examine the effect which these slow changes, induced by changes in solar radiation, have on gravity waves in the middle atmosphere. Conversely, the impact which solar driven changes in the lower atmosphere gravity-wave source have on the middle atmosphere background state will also be investigated.

Planetary-scale Rossby waves, which have periods of days to weeks, are a dominant feature of the winter polar stratosphere. They are also key to both solar effects on climate (e.g. Labitzke and van Loon, 1988) and troposphere-stratosphere coupling (e.g. Thompson *et al.*, 2002). The effect which Rossby-wave – gravity-wave interactions have on this key region will be explored in chapter (V).

The dynamics and composition of the middle and upper atmosphere are strongly affected by atmospheric thermal tides. Both the propagating diurnal tide and atmospheric gravity waves are in part sourced the lower atmosphere and can propagate into the middle and upper atmosphere, where they deposit energy and momentum. Therefore, understanding the interaction between these two is important in understanding how the lower atmosphere can affect the middle and upper atmosphere and, in turn, how these may affect the lower atmosphere. Chapter (VI) will consider the effect of gravity-wave / tidal interactions on the dynamics and composition of the equatorial middle and upper atmosphere.

I.1 BASIC ATMOSPHERIC STRUCTURE

The atmosphere is made up of shells, known as spheres, identified as regions in which the vertical temperature gradient is either positive or negative. The temperature gradient changes sign in transition regions, known as pauses. Figure (I.1) shows a schematic of the typical vertical temperature profile in the neutral atmosphere at mid-

latitudes (Hedin, 1991). The lower atmosphere, referred to as the troposphere, is primarily heated from the ground, so the temperature in this region decreases with altitude. The tropopause, the boundary between the lower and middle atmospheres, has a typical height of 8 km in the polar regions and a height of 18 km near to the equator. The difference in this height is due to differences in the temperature and convective flows in the troposphere between low and high latitudes.

The middle atmosphere is primarily heated by absorption of solar ultraviolet radiation by ozone and the heating from this absorption peaks at ~ 50 km altitude, which creates a local temperature maximum at the stratopause. Within the mesosphere, the global mean temperature falls with altitude to the coldest region of the atmosphere, the mesopause, at a height of around 90 km. The low temperatures in this region are created by a combination of low absorption rates of solar radiation and very efficient cooling by infrared emission from CO_2 . However, the temperature structure of this region is also greatly affected by the global-scale meridional circulation, which causes the summer mesopause to be the coldest part of the atmosphere, shown in figure (I.2).

In the upper atmosphere, above ~ 100 km, the primary heating process is absorption of solar ultraviolet and extreme ultraviolet, which results in an increase in temperature with altitude up to around 200 km. Above 200 km, molecular heat conduction becomes so efficient that the temperature of the neutral atmosphere remains almost constant with altitude. As can be seen in figure (I.1), the temperature of the thermosphere varies greatly with the solar cycle. This is because the ultraviolet and extreme ultraviolet components of the solar spectrum vary greatly throughout this cycle (see figure I.7 and section I.4.1). Both the dynamics and energetics of the neutral thermosphere are also strongly influenced by the charged particles which make up the ionosphere. In turn, the ionosphere is influenced by both the solar cycle and shorter-term variability of the solar wind and interplanetary magnetic field, which provides another source of variability in the thermosphere. Above 400 – 600 km, particles have ballistic orbits and can escape into space. This region is known as the exosphere.

From hydrostatic balance (discussed in section I.2), pressure falls exponentially with height throughout the atmosphere. From this temperature and pressure variation and the ideal gas law, it follows that the density of the atmosphere also falls approximately exponentially with height. As a consequence, around 90 % of the mass of the atmosphere is contained within the lower atmosphere. However, despite their relatively low masses, the middle and upper atmospheres do influence the troposphere

and the surface of the Earth. For example, the ozone layer absorbs ultraviolet solar radiation (e.g. Chapman, 1930). As a result of its importance, the variability of ozone concentrations in the atmosphere has been the subject of extensive study for many years (e.g. Gardiner, 1988).

Below the turbopause, the atmosphere is sufficiently turbulent to ensure its long-lived chemical constituents, such as N_2 , are well mixed. However, many minor species do vary greatly in concentration throughout the atmosphere. The distribution of ozone, figure (I.3) is influenced by transport as well as photochemical production and loss mechanisms (see appendix III). Most water vapour and clouds are confined to the troposphere and lower stratosphere. However, photo-dissociation of methane can create trace amounts of water vapour within the mesosphere. In the polar summer mesopause region, this water can condense to form clouds which can be seen to glow after sunset at ground level, earning them the name noctilucent clouds. Above the turbopause the atmosphere can be considered to be non-turbulent, so the concentration of chemical constituents changes as a function of height, depending on their mass (heavier constituents settle to the bottom), although this picture is complicated by production and loss mechanisms in this region.

I.2 ATMOSPHERIC THERMODYNAMICS

In order to understand the atmospheric processes discussed in this study, a physical description of the atmosphere, including the principal equations of thermodynamics and dynamics must be introduced. This section outlines the equations of thermodynamics and their associated effects. The equations of motion are introduced in section (I.3).

I.2.1 THE IDEAL GAS LAW

The atmosphere is often approximated to an ideal gas, which obeys the law,

$$pV = nRT \quad \text{I.1}$$

where p is the pressure of the gas, V is the volume, n is the number of moles within that volume, R is the universal gas constant and T is the absolute temperature. From this we obtain the relationship between the density, ρ , of the gas and its temperature and pressure,

$$p = \frac{\rho RT}{M} \quad \text{I.2}$$

where M is the mass of one mole of the gas.

I.2.2 HYDROSTATIC BALANCE

In many cases the atmosphere can be said to be in hydrostatic balance. This means that the weight of an air parcel is balanced by a difference in the pressures at the bottom and top of that air parcel, according to,

$$\frac{dp}{dz^*} = -g\rho \quad \text{I.3}$$

where g is the acceleration due to gravity and z^* is the geometric height.

I.2.3 SCALE HEIGHT

Using equations (I.2) and (I.3), it is possible to find the density and pressure of the atmosphere as a function of height, given a known temperature profile. Using the boundary condition that the pressure is equal to a reference pressure p_0 at the ground ($z^* = 0$), the pressure as a function of height for an isothermal atmosphere with temperature T_0 is given by,

$$p = p_0 \exp\left(-\frac{gz^*}{RT_0}\right) = p_0 \exp\left(-\frac{z^*}{H}\right) \quad \text{I.4}$$

where $H = RT_0/g$ is known as the pressure scale height. H is the height over which the pressure decreases by a factor of e . For this isothermal case, a similar relationship applies to the density,

$$\rho = \rho_0 \exp\left(-\frac{gz^*}{RT_0}\right) = \rho_0 \exp\left(-\frac{z^*}{H}\right) \quad \text{I.5}$$

where ρ_0 is the density of the air at the ground. Both the density and the pressure of the atmosphere fall exponentially with height, with a typical H of ~ 7 km in the middle atmosphere (based on an isothermal atmosphere of 260 K). Consequently, $\sim 90\%$ of the mass of the atmosphere is contained within the troposphere.

I.2.4 ADIABATIC LAPSE RATE

The dry adiabatic lapse rate, Γ_a is defined as the reduction in temperature that a dry parcel of air would experience if it was raised adiabatically from its current position. Γ_a is given by,

$$\Gamma_a = -\left(\frac{dT}{dz^*}\right)_{\text{parcel}} = \frac{g}{c_p} \quad \text{I.6}$$

where c_p is the specific heat at constant pressure. It is worth noting that in general this will differ from the atmospheric lapse rate, $\Gamma(z) = -dT/dz^*$ due to non-adiabatic processes in the atmosphere, such as the latent heat released during water condensation.

I.2.5 BRUNT-VÄISÄLÄ FREQUENCY

The Brunt-Väisälä frequency is a useful concept when considering atmospheric waves. If an air parcel, which is initially in equilibrium with its surroundings, is displaced vertically such that it rises adiabatically through a small height δz^* , the parcel will now have a new temperature, due to its adiabatic expansion. If the adiabatic lapse rate is greater than the atmospheric lapse rate, the parcel will have a lower temperature than the surrounding air, and from equation (I.2) it will also have a greater density, so will fall back to its original position (the condition for static stability). However, if the atmospheric lapse rate is less than the adiabatic lapse rate, the parcel will have a lower density than its surroundings and will continue to rise (convectively unstable). In the case of the two lapse rates being equal, there is neutral stability and the parcel will remain at its new height.

From equations (I.6), (I.2) and Newton's Second Law, it can be shown that,

$$\frac{d^2(\delta z)}{dt^2} + N^2(\delta z^*) = 0 \quad \text{I.7}$$

where,

$$N^2 = \frac{g}{T}(\Gamma_a - \Gamma) = \frac{g}{T}\left(\frac{dT}{dz^*} + \frac{g}{c_p}\right). \quad \text{I.8}$$

In a statically stable region, where $N^2 > 0$, equation (I.7) shows that a vertically displaced parcel will undergo simple harmonic motion about its original position with an angular frequency N , known as the Brunt-Väisälä frequency. For the statically unstable case where $N^2 < 0$, N is imaginary which leads to exponential solutions to equation (I.7), corresponding to the displaced parcel continuing to move with an increasing speed.

I.2.6 POTENTIAL TEMPERATURE

Potential temperature, θ is the temperature that a parcel of air would have if it was expanded adiabatically to a reference pressure p_0 (usually taken to be 1013 mb, the mean pressure at sea level). This can be thought of as the parcel being displaced vertically in the atmosphere until it reaches this reference pressure. The potential temperature can be defined as,

$$\theta = T \left(\frac{p_0}{p} \right)^\kappa \quad \text{I.9}$$

where $\kappa = R/c_p$, which is approximately 2/7 for a diatomic gas (such as molecular oxygen or nitrogen).

Equation (I.8) can now be rewritten in terms of potential temperature, such that,

$$N^2 = \frac{g}{\theta} \frac{d\theta}{dz}. \quad \text{I.10}$$

A region which is statically stable is one in which θ increased with height and a region is convectively unstable if θ decreases with height.

I.2.7 THERMAL CONDUCTION

As previously mentioned, thermal conduction in the atmosphere is an important process in maintaining the temperature structure of the thermosphere. The conduction of heat by the different species is given by,

$$\frac{C}{2} nk \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left(\lambda \frac{\partial T}{\partial z} \right) \quad \text{I.11}$$

where the coefficient of thermal conductivity, λ is given by,

$$\lambda = \frac{\sum_j C_j n_j}{\sum_j n_j} T^{0.69} \quad \text{I.12}$$

where n_j is the concentration of the j^{th} species and C_j is given in Table (I.1). From equation (I.12), it can be seen that this process is most efficient when there is a high background temperature and strong vertical temperature gradient, such as in the lower thermosphere. Indeed, above 150 km altitude, where infrared emission by CO_2 is no longer efficient, this becomes the dominant local cooling process in the atmosphere (e.g. Roble, 1995). Above around 200 km, the thermal conduction becomes so efficient that the vertical temperature gradient is reduced to almost zero.

I.3 ATMOSPHERIC DYNAMICS

In this section, the equations of motion that describe the dynamics of the middle atmosphere are introduced, as well as some useful approximations that are needed to introduce this study. By comparing the results from various approximations with observations of the middle atmosphere, key processes will be identified.

I.3.1 THE PRIMITIVE EQUATIONS ON A BETA-PLANE

The equations of motion for a compressible gas in a rotating frame of reference are well known, but in their generalised form are more complicated than necessary for dealing with medium-scale and large-scale atmospheric phenomena. Scale analysis can be used to reveal that it is a reasonable approximation to replace the vertical momentum equation with hydrostatic balance, the Coriolis force associated with the horizontal component of the Earth's rotation can be neglected and that the distance from the Earth's centre to any point in the atmosphere can be approximated to be equal to the Earth's radius. Further, the equations of motion can be written more easily if a new vertical coordinate, z , is defined. This log-pressure coordinate can be found from the inverse of equation (I.4),

$$z \equiv -H \ln(p/p_0). \quad \text{I.13}$$

It can be seen from equations (I.4) and (I.13) that $z = z^*$ if $T = T_0$. Here it is also necessary to define an important operator, known as the material derivative or advective derivative, which is given by,

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \quad \text{I.14}$$

in which the directions u , v and w represent eastward, northward and upwards at the Cartesian points x , y and z . Using these approximations, the equations of motion simplify to a set of equations known as the primitive equations. Rossby (1939) introduced a further approximation to the geometry by considering motion close to a latitude ϕ_0 , such that the Coriolis parameter can be simplified to

$$f = f_0 + \beta y \quad \text{I.15}$$

where $f_0 \equiv 2\Omega \sin \phi_0$, $\beta \equiv 2\Omega \cos \phi_0 / a$, a is the mean radius of the Earth and Ω is the rotation rate of the Earth. Using this geometry, known as the beta-plane, the primitive equations can be written in the form,

$$\frac{Du}{Dt} - fv + \Phi_x = X \quad \text{I.16}$$

$$\frac{Dv}{Dt} + fu + \Phi_y = Y \quad \text{I.17}$$

$$\Phi_z = \frac{RT}{H} = R\theta \exp\left(-\frac{\kappa z}{H}\right) \quad \text{I.18}$$

$$u_x + v_y + \frac{(\rho_0 w)_z}{\rho_0} = 0 \quad \text{I.19}$$

$$\frac{D\theta}{Dt} = Q \quad \text{I.20}$$

(e.g. Holton, 1975), where the suffixes x , y and z denote partial derivatives, $\Phi \equiv gz$ is the geopotential, X and Y are horizontal components of friction or some other non-conservative mechanical forcing, ρ_0 is background atmospheric density, and Q is the local diabatic heating rate. These equations represent momentum balance in zonal and meridional directions, hydrostatic balance, conservation of mass and the relation between diabatic heating and the material rate of change of potential temperature. Further approximations to these equations provide a useful insight into the dynamics of the middle atmosphere.

I.3.2 GEOSTROPHIC BALANCE AND THERMAL WIND

In the absence of friction, large-scale, low-frequency flows in the extra-tropics are often in geostrophic balance. This means that the Coriolis terms in equations (I.16) and (I.17) are balanced by a horizontal gradient in the geopotential. Using this condition and hydrostatic balance, it is possible to derive the thermal wind or thermal windshear equations,

$$f_0 \frac{\partial u_g}{\partial z} = -\frac{R}{H} e^{-\kappa z/H} \frac{\partial \theta}{\partial y} = -\frac{g}{T} \frac{\partial T}{\partial y} \quad \text{I.21}$$

$$f_0 \frac{\partial v_g}{\partial z} = \frac{R}{H} e^{-\kappa z/H} \frac{\partial \theta}{\partial x} = \frac{g}{T} \frac{\partial T}{\partial x} \quad \text{I.22}$$

where u_g and v_g are the horizontal components of the geostrophic wind (in this approximation, the vertical wind is zero). Equations (I.21) and (I.22) show that horizontal temperature gradients give rise to a vertical shear in the horizontal geostrophic wind, and that the direction of this flow depends upon the temperature gradient and the hemisphere in which the gradient occurs. This balance is responsible for creating the cyclonic and anti-cyclonic flows seen around pressure systems in the troposphere.

Under these conditions, the basic wind structure of the middle atmosphere can be described. Figure (I.4) shows the balance of these forces in the mesosphere. The

latitudinal variation in solar heating creates a horizontal pressure gradient which, in purely geostrophic balance, would produce the zonal jets seen in the figure (I.6).

Figure (I.5) shows the zonal mean wind climatology from the UARS Reference Atmosphere Project (URAP; Swinbank and Ortland, 2003) for December. Figure (1.6) shows the zonal mean winds calculated by Geller (1983) for December, assuming pure geostrophic balance. By comparing these two figures, it can be seen that geostrophic balance does not account for the entire wind structure of the middle atmosphere. The URAP climatology shows the mesospheric zonal jets closing around 80 km altitude and their direction reversing around 90 km altitude, whereas the pure radiative equilibrium considered in figure (I.6) indicates that the jets remain open throughout the mesosphere. This phenomenon is part of the mesopause anomaly, which shall be discussed in more detail in section (I.3.4). In order to explain this, the effects of frictional forces upon flows in the middle atmosphere must be considered.

I.3.3 THE TRANSFORMED EULERIAN-MEAN EQUATIONS

In order to describe the effect of frictional forces upon the middle atmosphere, a new set of equations of motion must be introduced, known as the transformed Eulerian mean equations. For the case of a time-averaged, zonally-averaged flow which is quasi-geostrophic (the effects of friction and time dependence have been included), Andrews and McIntyre (1978) defined a residual mean circulation in the meridional plane, (\bar{v}^*, \bar{w}^*) , given by,

$$\bar{v}^* \equiv \bar{v} - \rho_0^{-1} (\rho_0 \bar{v}' \theta' / \bar{\theta}_z)_z \quad \text{I.23}$$

$$\bar{w}^* \equiv \bar{w}_a + (\bar{v}' \theta' / \theta_{0z})_y \quad \text{I.24}$$

where the horizontal bar represents the zonal mean, the primes denote perturbations on this mean and the subscript a denotes the ageostrophic velocity, defined as the departure of the flow from geostrophic balance. Andrews and McIntyre (1978) used this circulation to derive the transformed Eulerian mean equations of motion,

$$\bar{u}_t - f_0 \bar{v}^* = \frac{1}{\rho_0} \nabla \cdot \mathbf{F} + \bar{X} \quad \text{I.25}$$

$$\bar{\theta}_t + \theta_{0z} \bar{w}^* = \bar{Q} \quad \text{I.26}$$

$$\bar{v}_y + \frac{1}{\rho_0} (\rho_0 \bar{w}^*)_z = 0 \quad \text{I.27}$$

$$f_0 \bar{u}_z = -\frac{R}{H} e^{-\kappa z/H} \bar{\theta}_y \quad \text{I.28}$$

where,

$$\nabla \cdot \mathbf{F} \equiv -\left(\rho_0 \overline{u'v'}\right)_y + \left(\rho_0 f_0 \overline{v'\theta'}/\theta_{0z}\right)_z \quad \text{I.29}$$

where \mathbf{F} is known as the Eliassen-Palm flux (EP flux) and is given by,

$$\mathbf{F} \equiv \left(0, -\rho_0 \overline{u'v'}, \rho_0 f_0 \overline{v'\theta'}/\theta_{0z}\right). \quad \text{I.30}$$

In equation (I.29), $\overline{v'\theta'}$ and $\overline{u'v'}$ are the eddy heat and momentum fluxes (see chapter II). The meaning and significance of EP flux is discussed in Section (I.5.2) and Chapter (V).

I.3.4 RESIDUAL CIRCULATION AND THE MESOPAUSE ANOMALY

From equations (I.25) to (I.28) it can be shown that a meridional residual circulation can be created by a frictional force \overline{X} or a divergence of \mathbf{F} . Further, for this circulation to exist, there must be a diabatic heating rate \overline{Q} , which shows that the atmosphere is not in radiative equilibrium (Dickinson, 1969). The origin of these terms will be explored in Sections (I.4) and (I.5), but first their impact upon the middle atmosphere can be set into context. Murgatroyd and Singleton (1961) first calculated the mean diabatic heating rate for the middle atmosphere under solstice conditions. Haurwitz (1961) suggested that this result implied an upwelling of air in the summer hemisphere and a downwelling in the winter, representing a summer to winter meridional circulation. Equation (I.24) shows that such a meridional circulation, \overline{v} , can be created by a zonal frictional drag force \overline{X} . This force would also have the effect of reducing the zonal mean zonal wind, creating the observed jet closure. From the continuity equation (I.27), this meridional circulation would create the upwelling and downwelling suggested by Haurwitz (1961). Such an upwelling in the summer hemisphere would represent an expansion of the air which would result in an adiabatic cooling, which produces the cold summer mesopause see in figure (I.2). Similarly, the downwelling in the winter produces the high temperature observed in the winter mesopause region.

Leovy (1964) introduced a linear friction term (Rayleigh friction) to the quasi-geostrophic equations to reproduce this mechanism of jet closure. However, Rayleigh friction alone cannot account for the zonal jets reducing to 0 ms^{-1} , or the jet reversal seen above 90 km in figure (1.5). Lindzen (1967) and Houghton (1978) proposed that this drag is actually created by dissipating gravity waves. The effects of this dissipation were first parameterised for use in global-scale models using a simplified theory of gravity-wave propagation and dissipation by Lindzen (1981). This parameterisation has been used by many authors to reproduce the observed zonal

wind and temperature structure in the mesopause region. The details of these wave and their representation in global-scale models are discussed in more detail in Section (I.5) and Chapter (II).

I.4 ATMOSPHERIC RADIATION

While the processes described in Section (I.3.4) have a large impact upon the detailed structure of the atmosphere, the overall temperature structure of the atmosphere is controlled by the balance between absorption and emission of radiation. The details of absorption and emission of radiation by the atmosphere are very complex and still a subject of current research, but the fundamental processes are understood. This section will review the solar input and radiative processes which are most relevant to the energy budget of the middle atmosphere.

I.4.1 THE SOLAR SPECTRUM AND ULTRAVIOLET ABSORPTION

Figure (I.7) shows the solar spectrum which is incident on the Earth's atmosphere and its variability with the solar cycle. While the variability in spectral irradiance is of order 0.1 % around the peak in the spectrum, at UV and EUV frequencies the variability can exceed 100 %. To understand the effect this has on the energy input to the atmosphere, the absorption of this spectrum by the atmosphere must be considered.

O₂ and O₃ are the primary absorbers of UV radiation in the middle atmosphere. This absorption can either excite vibrational modes and electron transitions which create discrete absorption bands, or photo-dissociation and photo-ionisation which create absorption continua. Figure (I.8) shows the absorption cross-section of O₃. O₃ absorbs UV in the Hartley and Huggins bands, and absorbs visible and near IR in the Chappuis band. Heating caused by this UV absorption by O₃ peaks around the stratopause. This creates the local temperature maximum seen in figure (I.1). Figure (I.9) shows the absorption cross-section of O₂. O₂ is ionised by wavelengths less than 100 nm and is photo-dissociated in the Schumann-Runge continuum. The Schumann-Runge band is associated with electron transitions and vibrational excitation. The lower-frequency Herzberg continuum is created by electron transitions and ionisation. At 242 nm this ionisation produces ground-state oxygen atoms which are important for the creation of O₃. Above 80 km altitude, radiative absorption by O₂ becomes a significant heat source. Absorption in the Schumann-Runge band provides heating in the upper mesosphere and absorption in the Schumann-Runge continuum heats the lower thermosphere. EUV ionisation of O, O₂ and N₂ becomes a significant heat source above 150 km. As the greatest variability in the solar spectrum due to solar activity

occurs at EUV frequencies, the largest changes in heating occur in the thermosphere. Thermospheric temperatures are therefore highly sensitive to solar activity, shown in figure (I.1). The influence of changes in solar activity on the middle atmosphere is discussed in further detail in chapter (IV).

I.4.2 INFRARED EMISSION AND LOCAL THERMODYNAMIC EQUILIBRIUM

Radiative cooling in the middle atmosphere is dominated by IR emission from CO_2 and O_3 . CO_2 has a broad band emission around $15\text{ }\mu\text{m}$ and a narrow band near $4.3\text{ }\mu\text{m}$, both created by de-excitation of vibrational modes, while O_3 has a strong absorption cross-section around $9.6\text{ }\mu\text{m}$, which contribute to both the heating and cooling of the middle atmosphere. CO_2 emission is an effective source of cooling throughout the middle atmosphere, with peaks around the stratopause and mesopause. The peak in radiative cooling at the stratopause shown in figure (I.10) is caused by the local temperature maximum. The peak near the mesopause is created by an increase in efficiency of radiative cooling in this region of the atmosphere, which is said to be in non-local thermodynamic equilibrium (non-LTE). The atmosphere below 70 km is said to be in local thermodynamic equilibrium (LTE) where collisional excitation and de-excitation dominates over radiative processes such as photon absorption, collisionally activated emission and spontaneous emission. The radiation emitted by such a region will resemble a Planck black body function. In the non-LTE region, above around 80 km , both collisional and radiative excitation and de-excitation are important, which leads to a different kind of emission spectrum. The efficiency of the radiative cooling by CO_2 increases in this non-LTE region due to a decrease in collisional de-excitation of the $\text{CO}_2(\text{O}1^1\text{O})$ state. This, in part, gives rise to the local temperature minimum at the mesopause. In the mid and upper thermosphere, radiative cooling becomes less important and molecular heat conduction becomes the main local cooling mechanism.

Figure (I.10) shows the calculated global mean radiative heating and cooling contributions from the main constituents of the atmosphere and the total heating and cooling rates as a function of height. The difference in the local heating and cooling rates represents a net diabatic heating, discussed in section (I.3.1). This shows that the atmosphere is not in radiative equilibrium. The non-conservative wave processes believed to be responsible for maintaining this state (Dickinson, 1969) are discussed in the next section.

I.5 ATMOSPHERIC WAVES

Sections (I.3) and (I.4) have highlighted the importance of the effects of waves on the middle atmosphere. While there are many classes of atmospheric waves, arguably the most important when considering the large-scale dynamics of the middle atmosphere are atmospheric tides, planetary waves, internal gravity waves and Kelvin waves. All four are introduced below, and gravity waves are discussed in detail in chapter (II).

I.5.1 ATMOSPHERIC TIDES

Atmospheric thermal tides are waves that are driven by solar heating of the atmosphere. They have frequencies that are harmonics of the diurnal cycle. In the middle atmosphere, the diurnal and semi-diurnal tides tend to be the dominant modes. These wave modes are primarily driven by IR heating of H₂O in the troposphere and UV heating of O₃ in the stratosphere. Laplace (1825) showed that these waves can be represented by equations of the form,

$$\tilde{u}_t - f\tilde{v} + (a \cos \phi)^{-1} \tilde{\Phi}_\lambda = 0 \quad \text{I.31}$$

$$\tilde{v}_t + f\tilde{u} + a^{-1} \tilde{\Phi}_\lambda = 0 \quad \text{I.32}$$

$$\frac{\tilde{u}_\lambda}{a \cos \phi} + \frac{(\tilde{v} \cos \phi)_\phi}{a \cos \phi} + \frac{1}{gh} \tilde{\Phi}_t = 0 \quad \text{I.33}$$

where \tilde{u} and \tilde{v} are the horizontal velocity components of the wave, $\tilde{\Phi}/g$ is the depth perturbation of the wave and h is known as the equivalent depth. Chapman and Lindzen (1970) found sinusoidal solutions which are bound at the poles, with zonal wavenumber s and period $2\pi/2\Omega\sigma$ can be found. For such solutions, equations (I.31) to (I.33) simplify to,

$$L\tilde{\Phi} + \gamma\tilde{\Phi} = 0 \quad \text{I.34}$$

which is called Laplace's tidal equation, where L is the operator,

$$L = \frac{d}{d\mu} \left[\frac{(1-\mu^2)}{(\sigma^2 - \mu^2)} \frac{d}{d\mu} \right] - \frac{1}{(\sigma^2 - \mu^2)} \left[\frac{-s(\sigma^2 + \mu^2)}{\sigma(\sigma^2 - \mu^2)} + \frac{s}{1-\mu^2} \right] \quad \text{I.3.5}$$

$\gamma = 4\Omega^2 a^2 / gh$ is known as Lamb's parameter, $\mu = \sin \phi$, s is the zonal wavenumber, and the period is $2\pi/2\Omega\sigma$. The eigenfunctions $\Theta_n^{(\sigma,s)}$ and eigenvalues γ of Laplace's tidal equation satisfy,

$$L\Theta_n^{(\sigma,s)} + \gamma_n^{(\sigma,s)} \Theta_n^{(\sigma,s)} = 0 \text{ for } n=1,2,\dots \quad \text{I.36}$$

The eigenfunctions are called Hough functions, which were tabulated, along with the eigenvalues, by Longuet-Higgins (1968).

The solutions to Laplace's tidal equation show several important properties of atmospheric tides which are relevant to the study in Chapter (VI). Diurnal and semi-diurnal tides can propagate vertically through the middle atmosphere and, in the absence of dissipation, will grow in amplitude as the background density decreases. The propagating diurnal tide is largely confined in the region $\pm 30^\circ$ latitude because the inertial period, f , falls below 24 hours at higher latitudes. The diurnal tide has a vertical wavelength of around 28 km, whereas the semi-diurnal tide has a wavelength of over 100 km. These vertical wavelengths mean that the diurnal tide is most readily excited by IR absorption by H_2O in the troposphere, whereas the semi-diurnal tide is mainly excited by the more vertically extended O_3 layer in the stratosphere and mesosphere, which absorbs mainly in the UV.

Tides can have a large effect on the dynamics and composition of the middle atmosphere. Tidal oscillations in density, temperature and horizontal winds can be observed by ground based instruments and satellites studying the dynamics of the middle atmosphere. Figure (I.11) shows the diurnal tidal oscillations in the meridional winds in the mesosphere, as observed by the UARS satellite. The pattern of cells, anti-symmetric about the equator, is characteristic of the diurnal tide. The meridional wind components associated with this oscillation (over 60 ms^{-1} in some places) can easily dominate over the background meridional winds (typically a few ms^{-1}). In addition to its effects on short time-scales, Akmaev and Shved (1980) showed that tidal oscillations can have a net effect on the composition of the middle atmosphere, as chemical reaction rates can respond non-linearly to changes in the background density and temperature. Lindzen (1968) showed that tidal dissipation in the mesosphere can produce turbulence and net vertical motions, which lead to further changes in composition (e.g. Fauliot, 1997). Tidal dissipation can also lead to acceleration of the background winds, which is believed, in part, to be responsible for driving the westward zonal mean flow in the equatorial mesosphere during equinox, shown in figure (I.5) (e.g. McLandress, 1998). Interactions between atmospheric tides and planetary waves or gravity waves are also believed to play an important role in affecting the structure of the middle atmosphere. Such effects will be discussed in Chapter (VI). In the thermosphere, molecular diffusion damps the shorter wavelength diurnal tide more than the longer wavelength semi-diurnal tide (Hines, 1960), so the semi-diurnal tide becomes the dominant component in the middle and upper thermosphere.

The solutions to Laplace's tidal equation discussed here are idealised, and do not take into account several factors, such as interactions with other atmospheric waves. The

effect of the coupling between atmospheric gravity waves and the diurnal tide on the tidal wavelength and amplitude will be discussed further in Chapter (VI).

I.5.2 PLANETARY WAVES AND THE ELIASSEN-PALM FLUX

Rossby or planetary waves are a common feature of the winter stratosphere and play an important role in the Brewer-Dobson circulation, shown in figure (I.12). Planetary waves can be formed in the troposphere by wind interactions with large-scale topography such as mountains or land-sea heating contrasts or, as proposed by Plumb (1983), by instabilities within the mesospheric jets. Planetary waves can be either stationary or travelling with respect to the ground and forced or unforced (natural modes), but all are large-scale horizontal perturbations which are restored by the latitudinal gradient in the Coriolis force. Such wave motions have the form,

$$\Psi' = \text{Re} \hat{\Psi}(z) e^{ik(x-ct)} \sin ly \quad \text{I.36}$$

where Ψ' is the perturbation of the stream-function $\hat{\Psi}$, k is the zonal wavenumber, l is the meridional wavenumber and c is the phase speed (positive for eastwards motion). Following the method of Charney and Drazin (1961), in which the mean flow is assumed to depend only on z , the wave motion can be written in terms of a departure from the zonal mean flow $\bar{u}(z)$, given by the second order equation in $\hat{\Psi}$,

$$\frac{1}{\rho_0} \left(\frac{\rho_0 f_0 \hat{\Psi}_z}{N^2} \right)_z + \left[\frac{\bar{q}_v(z)}{\bar{u}(z) - c} - (k^2 + l^2) \right] \hat{\Psi} = 0 \quad \text{I.37}$$

where q is Ertel's potential vorticity (Rossby, 1940; Ertel, 1942), defined as,

$$q = \rho^{-1} \omega_a \cdot \nabla \theta \quad \text{I.38}$$

where ω_a is the absolute vorticity. From equation (I.37) it can be seen that if the horizontal phase speed is equal to the mean flow speed at any level, a singularity arises. Planetary waves cannot propagate through this critical level. For vertical propagation to occur, it can also be shown that the phase speed of the waves must obey the condition,

$$0 < \bar{u} - c < \frac{\beta}{k^2 + l^2 + f_0/4H^2 N^2} \equiv \bar{u}_c \quad \text{I.39}$$

where \bar{u}_c is the phase speed relative to the background flow. Equations (I.39), known as the Charney-Drazin criterion (Charney and Drazin, 1961), shows that for a planetary wave to propagate vertically, its phase speed must be westward with respect to the mean wind, and not too fast. For waves which are stationary with respect to the ground, the Charney-Drazin criterion means that vertical propagation can only occur when the mean flow is eastward and not too fast.

The dispersion relation for planetary waves is given by,

$$\omega = k\bar{u} - \frac{\beta k}{k^2 + l^2 + f_0(m^2 + 1/4H^2)/N^2} \quad \text{I.40}$$

where m is the vertical wavenumber. As $k > 0$ and $m > 0$, the phase fronts, $kx - kct + mz = \text{constant}$, tilt westwards with height.

Planetary-wave oscillations of periods close to 2, 5 and 16 days, with zonal wavenumbers 3, 2 and 1 respectively, are a commonly observed feature of the middle atmosphere. In the stratosphere and lower mesosphere, these maximise in the winter hemisphere because of favourable zonal mean wind conditions. Planetary waves have been observed in the upper mesosphere, maximising in the summer hemisphere (e.g. Glass *et al.*, 1975; Mueller and Nelson, 1978). The zonal mean winds in the mesosphere and stratosphere prevent the propagation of these waves from source regions in the troposphere, so the source of these waves has been suggested to be instabilities within the summer mesospheric jet core (Plumb, 1983). Planetary waves can propagate into the lower thermosphere (e.g. Mitchell *et al.*, 1999), but are not believed to be able to reach higher altitudes. However, planetary-wave signatures have been observed in the upper thermosphere (e.g. Chen, 1992; Apostolev *et al.*, 1995, 1998), although these wave signatures are not believed to be the result of *in-situ* planetary wave forcing (e.g. Laštovicka and Sauli (1999).

Planetary waves can drive a residual circulation as discussed in section (I.3.3), through the divergence of Eliassen-Palm flux (EP flux). The EP flux (equations I.29 and I.30) can be considered as a generalised form of group velocity for planetary wave oscillations. For the case of $\nabla \cdot \mathbf{F} = 0$, there is no dissipation and the planetary waves do not drive a residual circulation. In order to drive a circulation, the planetary waves must be dissipative. In the case where the dissipation is approximated by Newtonian cooling and Rayleigh friction (which both have time constants equal to α), the divergence of the EP flux is given by,

$$\nabla \cdot \mathbf{F} = -\frac{\rho_0 \alpha \beta}{\bar{u}_z} |\hat{\Psi}_0|^2 \exp\left(-\frac{\alpha \beta z N^2}{f_0^2 k \bar{u}^2 m}\right) \sin^2 ly \leq 0. \quad \text{I.41}$$

The dissipation of planetary waves affects the atmosphere on both local and global scales. In the polar regions planetary-wave breaking and divergence of EP flux produces meridional mixing of the air in the polar vortex with warmer air from lower latitudes, preventing the winter polar stratosphere from becoming as cold as for the radiative equilibrium case (discussed in more detail in Chapter V). On the global scale, EP flux divergences drives the stratospheric Brewer-Dobson circulation, shown in figure

(I.12), which was proposed by Brewer (1949) and Dobson (1956). Brewer (1949) inferred the upward motion of air at low latitudes shown in figure (I.12) from the lack of moisture in the stratosphere. The poleward motion in the lower stratosphere was then inferred by Dobson (1956) from observations of polar ozone number densities.

I.5.3 ATMOSPHERIC GRAVITY WAVES

Atmospheric gravity waves are smaller-scale than the tides and planetary waves discussed above, but they have a great impact on the structure of the middle atmosphere, especially in the mesopause region (discussed in section I.3.4). Their effects can also be seen on smaller scales, such as in airglow images. Figure (I.13) shows the airglow emission from OH at 87 km (from Nakamura *et al.*, 1998). The density variations due to a passing gravity wave are clearly visible as changes in the emission rates.

The present study will focus on the impact of gravity waves on the global-scale structure of the middle atmosphere. Chapter (II) will discuss in detail gravity-wave sources, their propagation and dissipation, with particular emphasis on their representation within global-scale atmospheric models. Chapter (IV) will focus on the effect of variable solar flux on gravity waves in the middle atmosphere, while Chapters (V) and (VI) will discuss the interactions of gravity waves with planetary waves and tides within global-scale atmospheric models.

I.5.4 KELVIN WAVES

Equatorial Kelvin waves are believed to be important in the driving of the quasi-biennial oscillation (QBO) in the stratosphere (Lindzen and Holton, 1968; Holton and Lindzen, 1972). Kelvin waves are equatorially trapped waves, which are forced by variations in convection in the troposphere. Figure (I.14) shows a schematic of the oscillations in the geopotential and horizontal wind produced by these waves. As can be seen, the horizontal winds induced by Kelvin waves are purely zonal. Viewed in the equatorial plane, their phase fronts tilt eastwards with height (with respect to the background flow) and propagate downwards with time. As their phase propagation direction is perpendicular to their group velocity, their group velocity is upwards and eastwards with respect to the background flow and so they carry eastward momentum from the troposphere into the equatorial middle atmosphere.

Kelvin waves can be divided into two categories. Meso-scale Kelvin waves have zonal wavelengths of order 100 – 1000 km and vertical wavelengths of order 10 km (e.g.

Riggin *et al.*, 1997). Planetary-scale ultra-fast Kelvin waves are globally coherent modes with vertical wavelengths of order 40 – 80 km and periods of several days.

I.6 THE IONOSPHERE

The neutral atmosphere coexists and interacts with a region of charged particles known as the ionosphere. These ions are produced either by collisions with precipitating energetic particles or from photoionisation. The production rate per unit volume of ions by a single wavelength of solar radiation is given by the Chapman production formula,

$$q(z, \theta) = \mu \sigma N_0 I_\infty \exp\left(-\sigma H \sec \theta N_0 e^{-\frac{z}{H}} - z/H\right) \quad \text{I.42}$$

where q is the production rate per unit volume per unit time, θ is the angle between the zenith and the incoming radiation, μ and σ are constants, N_0 is number density of the atmosphere at some reference height and I_∞ is intensity of radiation incident upon atmosphere. This production peaks at a certain altitude for each frequency, forming a Chapman layer. The production is, on average, balanced by ion recombination, with the equilibrium electron density profile shown in figure (I.15).

Collisions with charged particles in the ionosphere play a major role in the momentum and energy budgets of the neutral thermosphere. The effect of the neutrals on the charged particles, and *vice versa*, depends upon the ratio of the collision frequency to the gyro frequency. In the presence of a magnetic field, \mathbf{B} , charged particles undergo gyroscopic motion about the field lines, with rotational 'gyro' frequency, $\omega_i = \mathbf{B}q_i/m_i$, where q_i is the charge and m_i is the mass of the particle. The collision frequency is defined in terms of an ion-neutral or neutral-ion collisional rate coefficient (K_{in} or K_{ni}), such that $\nu_{in} = nK_{in}$ for ion-neutral collisions where n is the neutral concentration. In the thermosphere, the relative concentrations of neutrals and charged particles means that the ion-neutral collision frequency is typically much larger than the neutral-ion collision frequency. Collisions between ions and neutrals can accelerate the neutrals through ion drag, given by

$$\frac{\partial v_n}{\partial t} = \frac{\rho_i}{\rho_n} \nu_{in} (v_i - v_n) \quad \text{I.43}$$

where v_i and v_n are the velocities of the ions and the neutrals, or cause Joule heating, given by

$$\frac{d\varepsilon}{dt} = \frac{\mathbf{J} \cdot \mathbf{E}}{\rho_n}$$

I.44

where ε is the kinetic energy plus enthalpy per unit mass of the neutrals, \mathbf{J} is the current density and \mathbf{E} is the electric field in the ionosphere. There are two major source of electric fields within the ionosphere: interactions with neutrals and solar wind interactions with the magnetosphere. In regions where $\omega_i \ll \nu_{in}$, the charged particle motion is dominated by collisions with the neutrals, but the electrons and ions are often subject to different forcing as their collision and gyro frequencies differ. In the E-region ionosphere, where the conductivity is high, this leads to polarisation and an associated dynamo electric field. Another major electric field exists within the polar cap, which is related to the rate of anti-sunward flow of plasma in the ionosphere and magnetosphere caused by reconnection processes. Joule heating arising from collisions between the accelerated charges and the neutrals in the polar cap region. This, along with precipitating particles in the auroral region, accounts for a large portion of the thermospheric energy budget (e.g. Roble, 1995).

The ionosphere is affected by energetic, dynamic and chemical interactions with the thermosphere and middle atmosphere. A number of global-scale numerical models have been developed to help investigate this interaction, such as the TIME-GCM (Roble and Ridley, 1994) and the CMAT model (Harris, 2001). In addition, the possible effect of the ionosphere magnetosphere system upon the middle atmosphere, through a variety of different coupling mechanisms, is a topic of current research.

J	C_j
N_2	56
O_2	56
O	75.9
He	299
H	379

Table I.1, Numerical coefficients for equation (I.11), after Rees (1989).

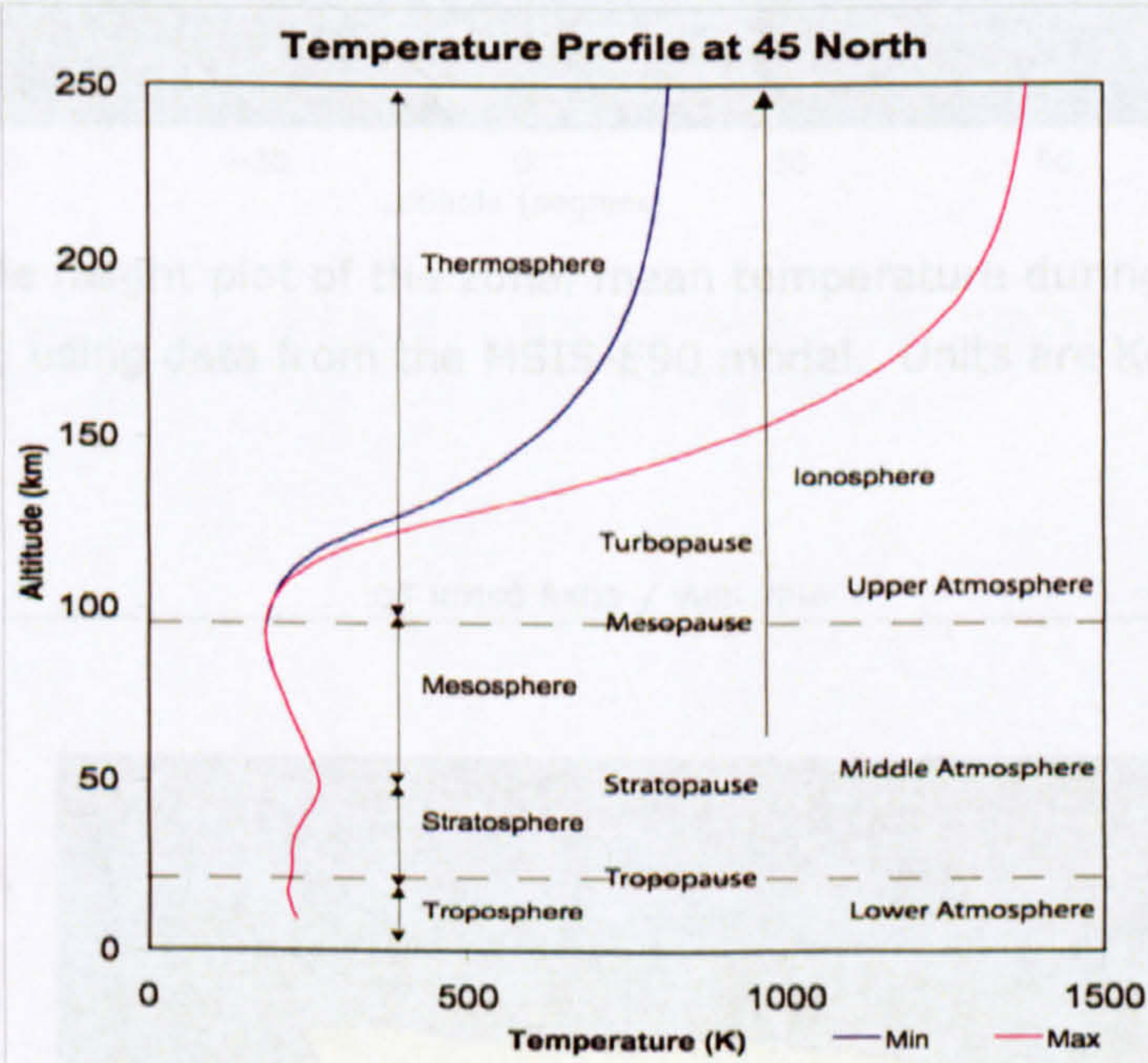


Figure I.1, Zonal mean temperature profile of the atmosphere at 45° North during equinox at solar minimum and maximum, using data from MSIS-E90 empirical model (Hedin, 1991).

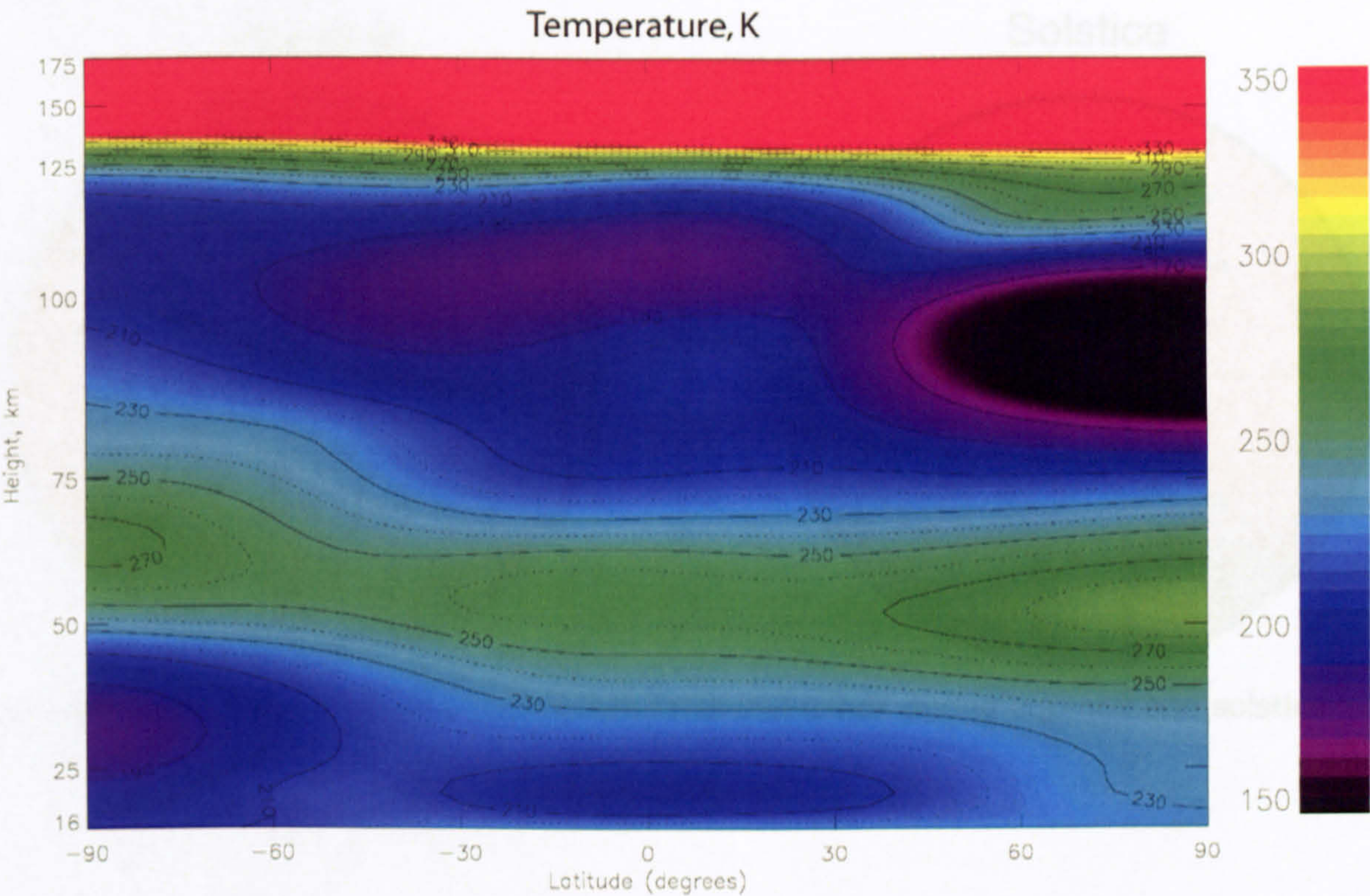


Figure I.2, Latitude height plot of the zonal mean temperature during the June solstice at solar minimum, using data from the MSIS-E90 model. Units are Kelvin.

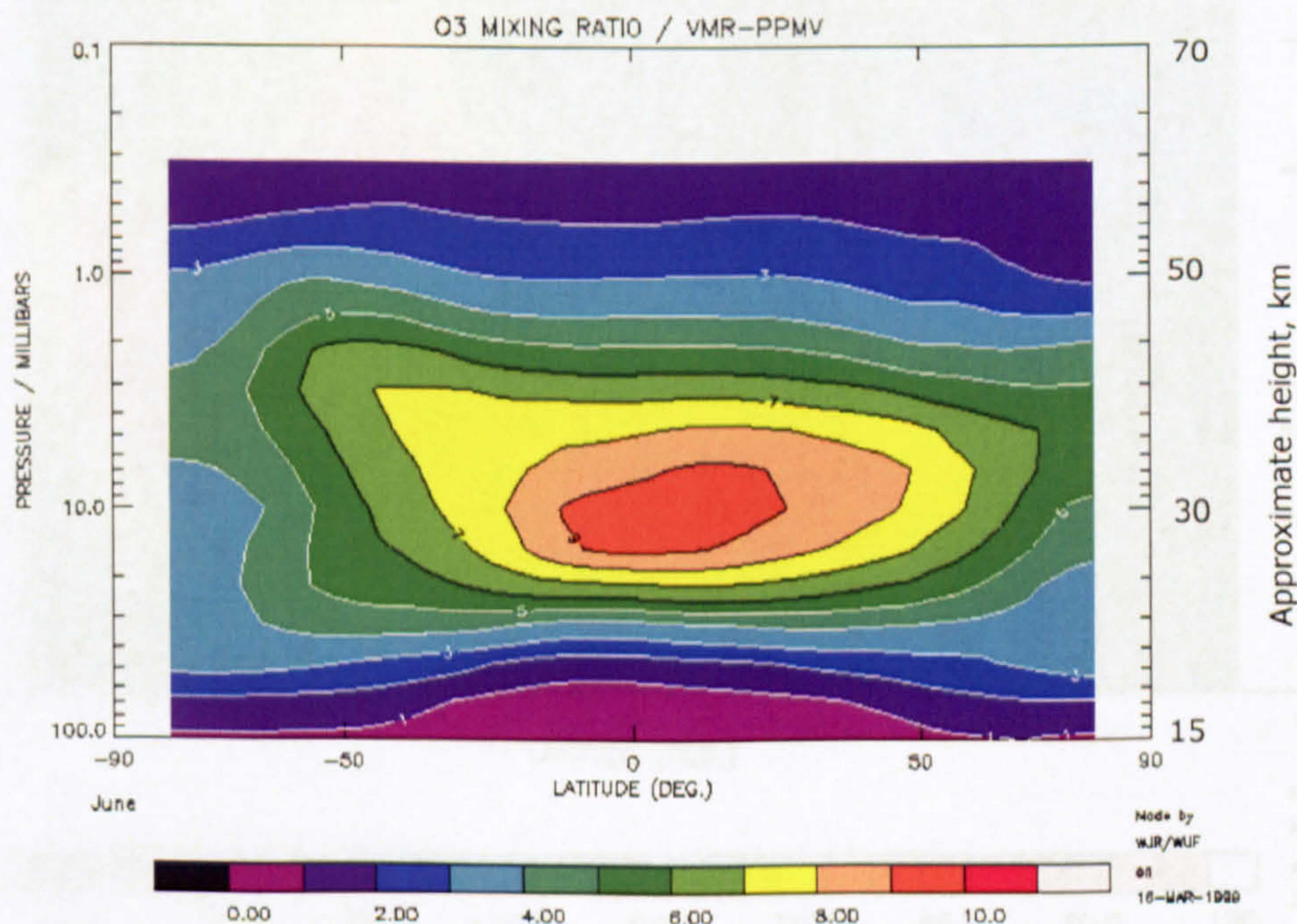


Figure I.3, Ozone volume mixing ratio climatology for June from the UARS Reference Atmosphere Project, extended baseline (Swinbank and Ortland, 2003). Units are parts per million by volume.

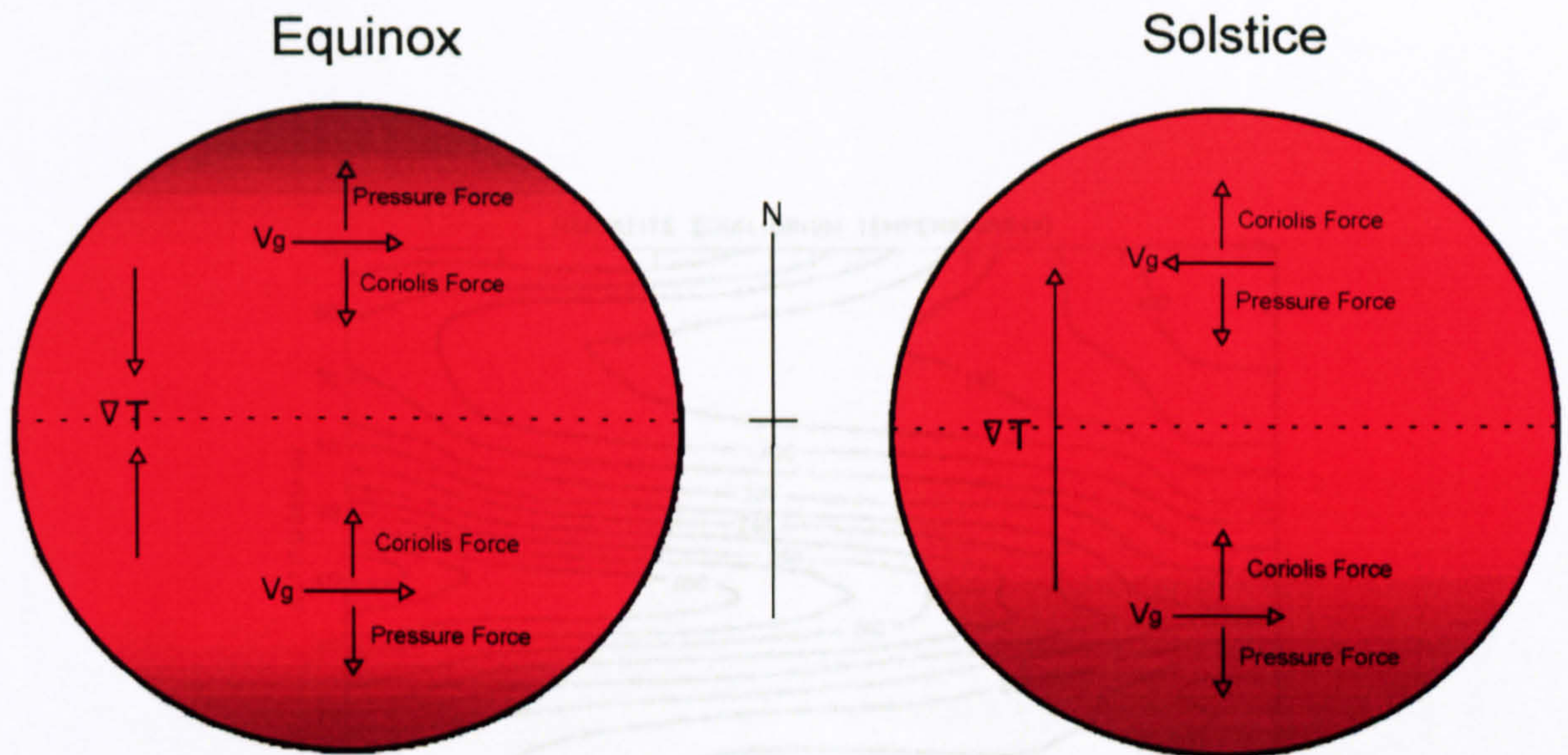
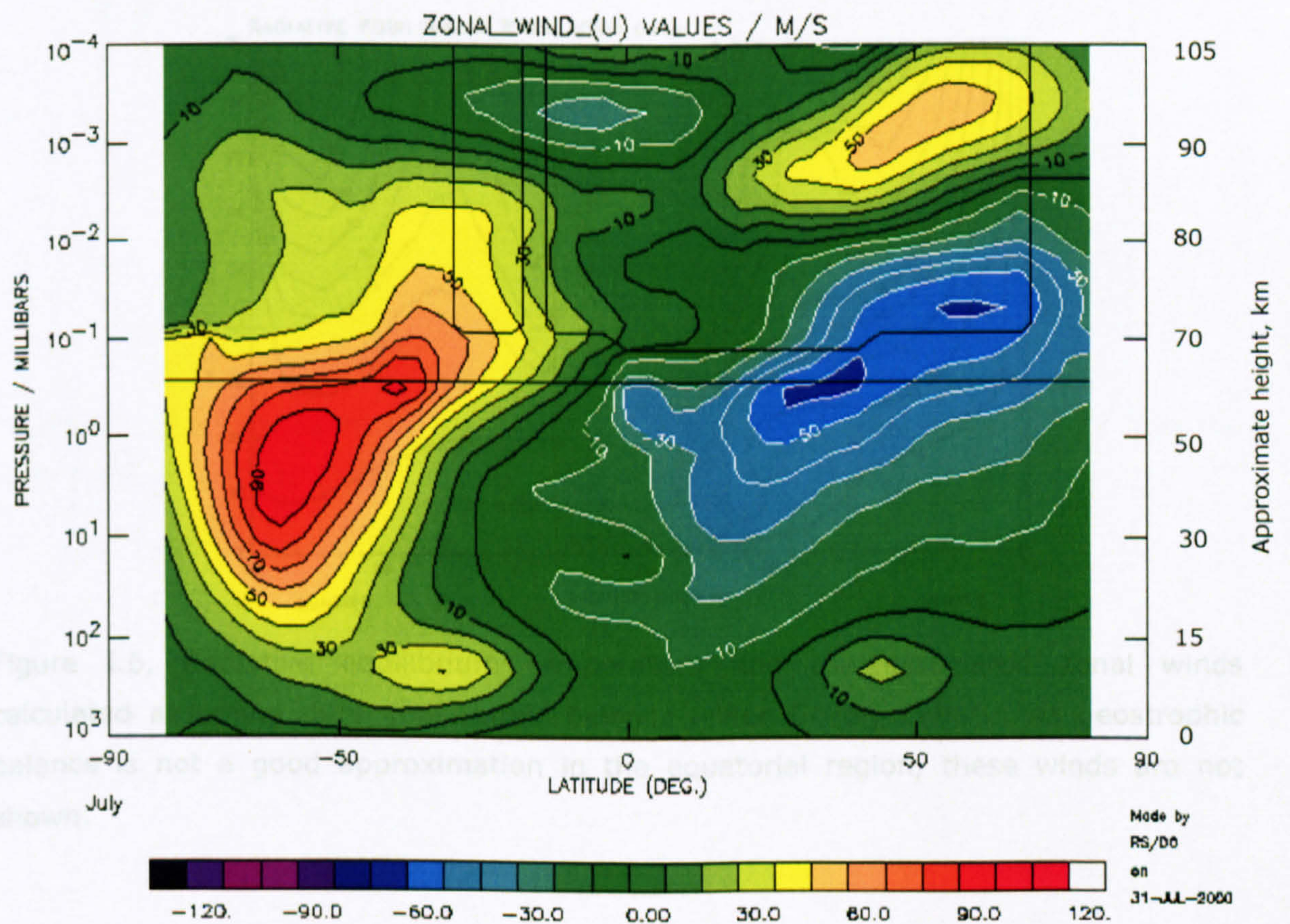


Figure I.4, Illustration of the mesospheric thermal winds during equinox and solstice.



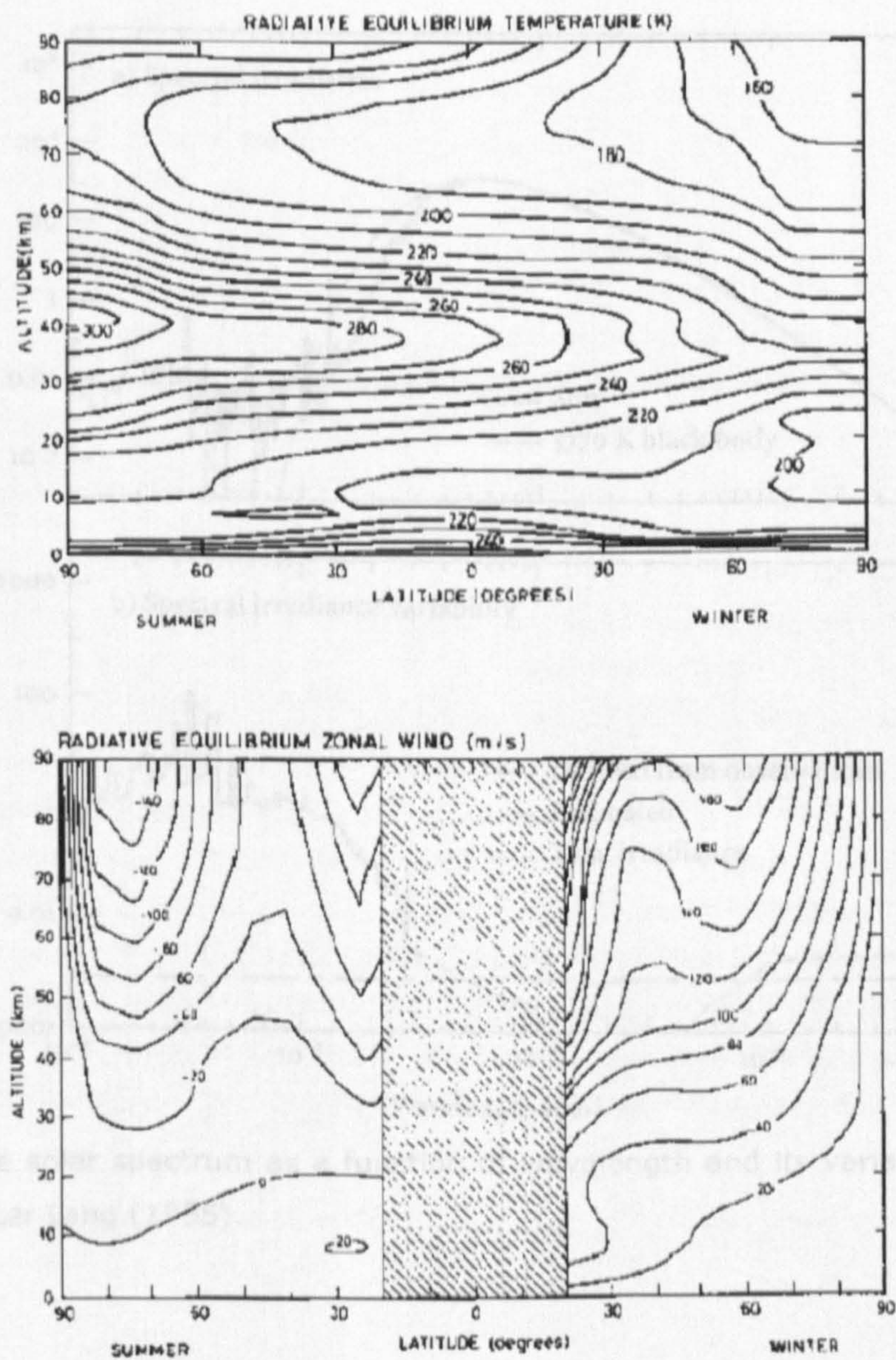


Figure I.6, Radiative equilibrium temperature and the associated zonal winds calculated assuming pure geostrophic balance (after Geller, 1983). As geostrophic balance is not a good approximation in the equatorial region, these winds are not shown.

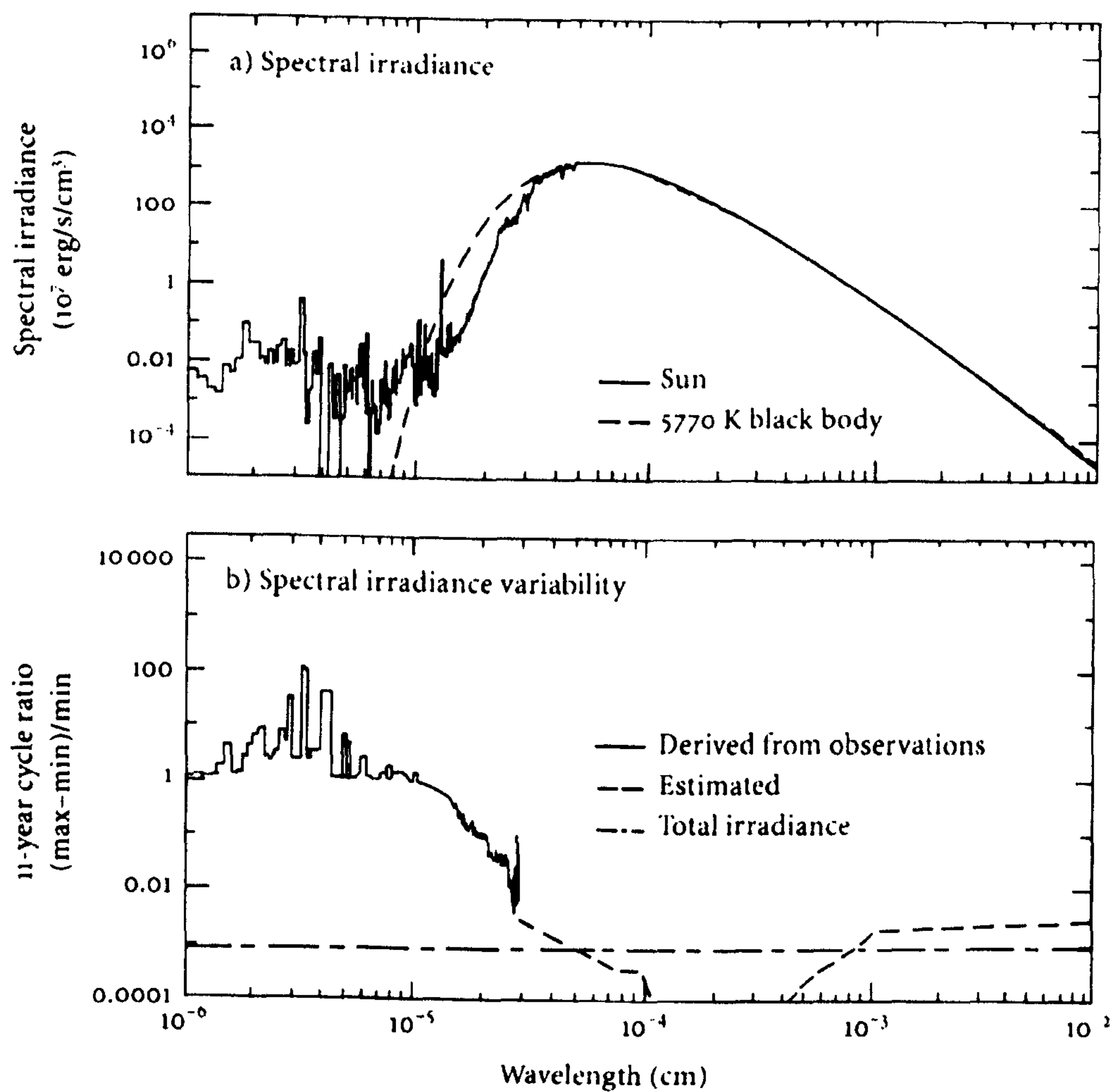


Figure I.7, The solar spectrum as a function of wavelength and its variability over the solar cycle, after Lang (1995).

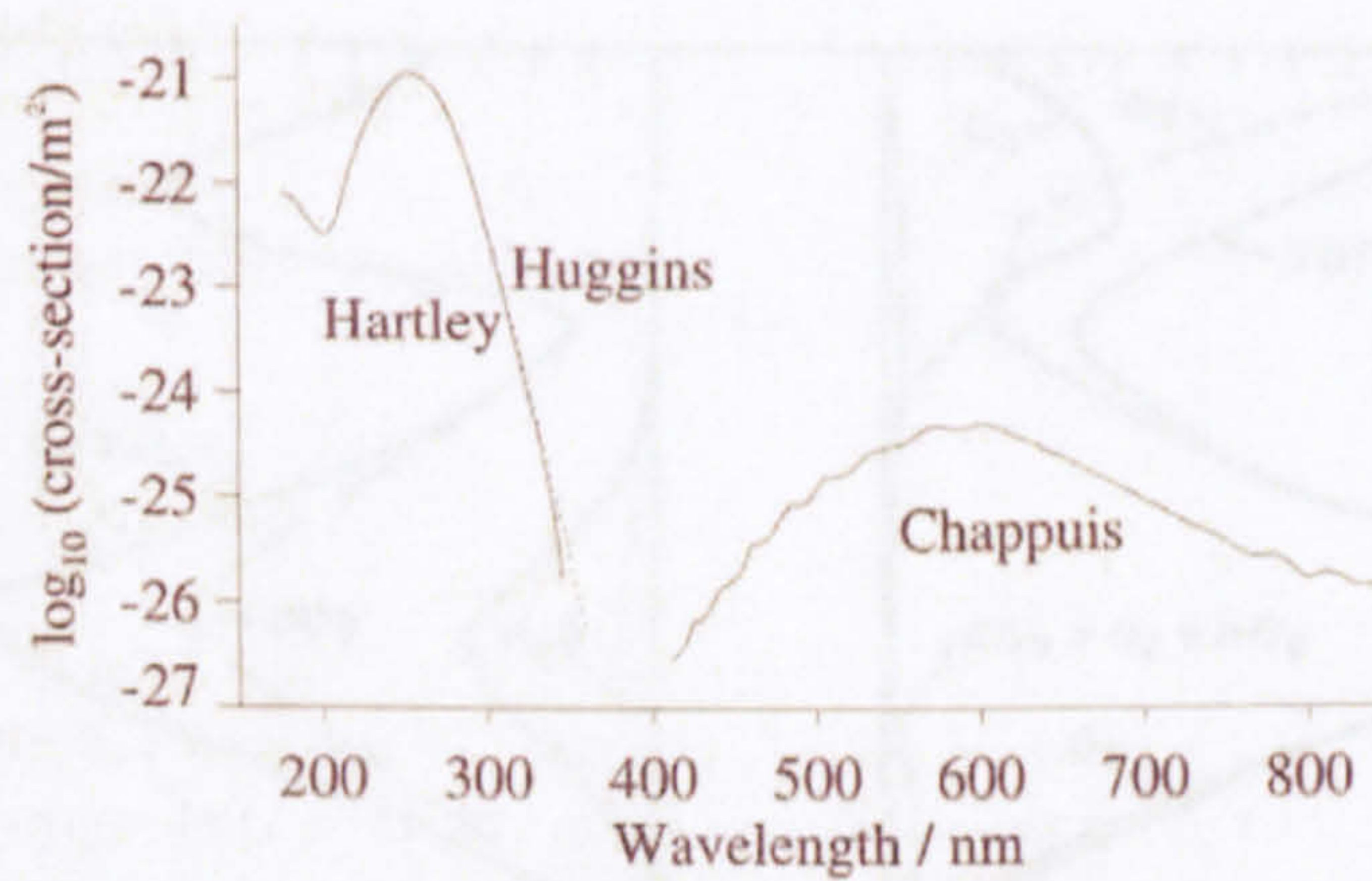


Figure I.8, Absorption cross section of O_3 as a function of wavelength, after Andrews (2000).

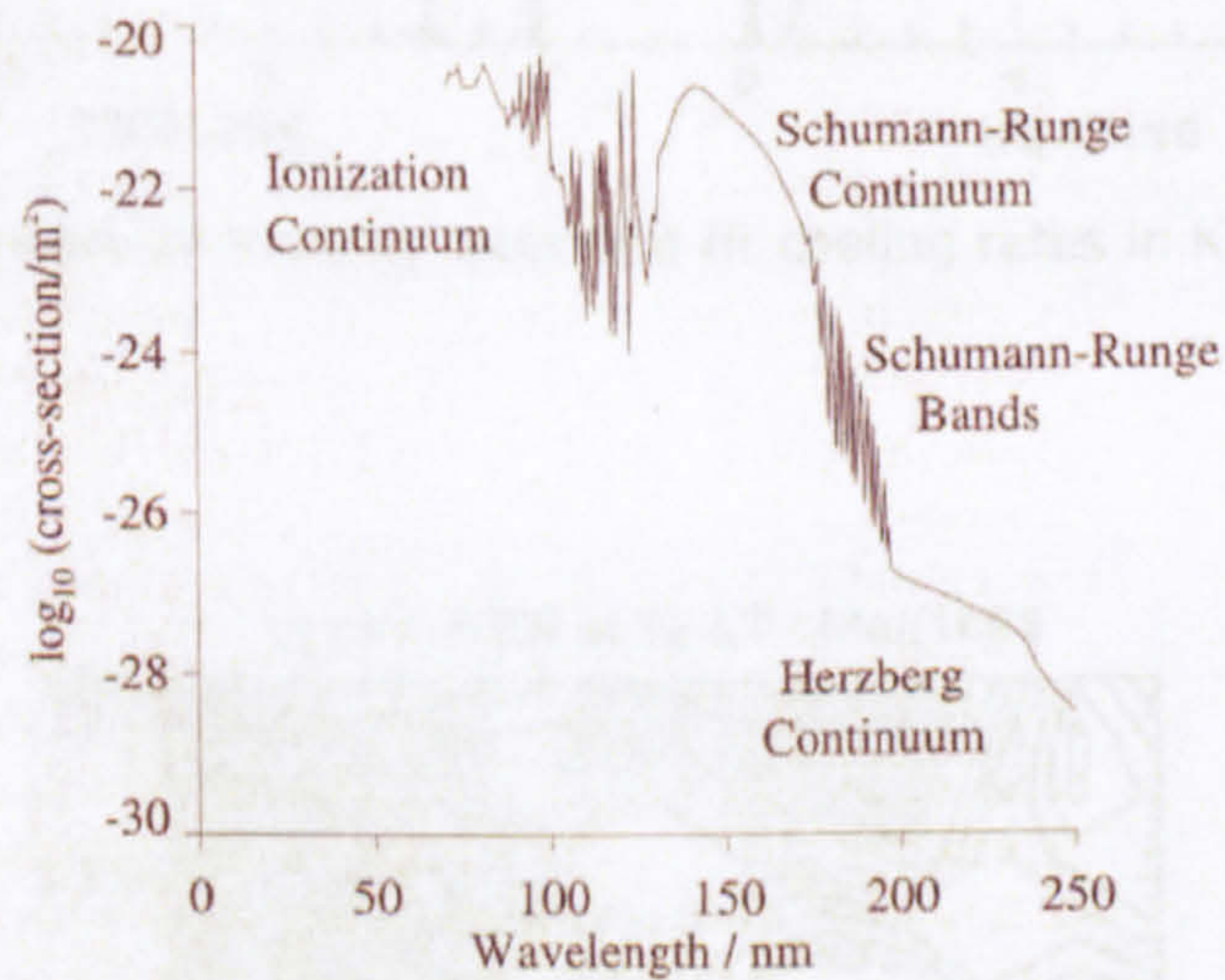


Figure I.9, Absorption cross-section of O_2 as a function of wavelength, after Andrews (2000).

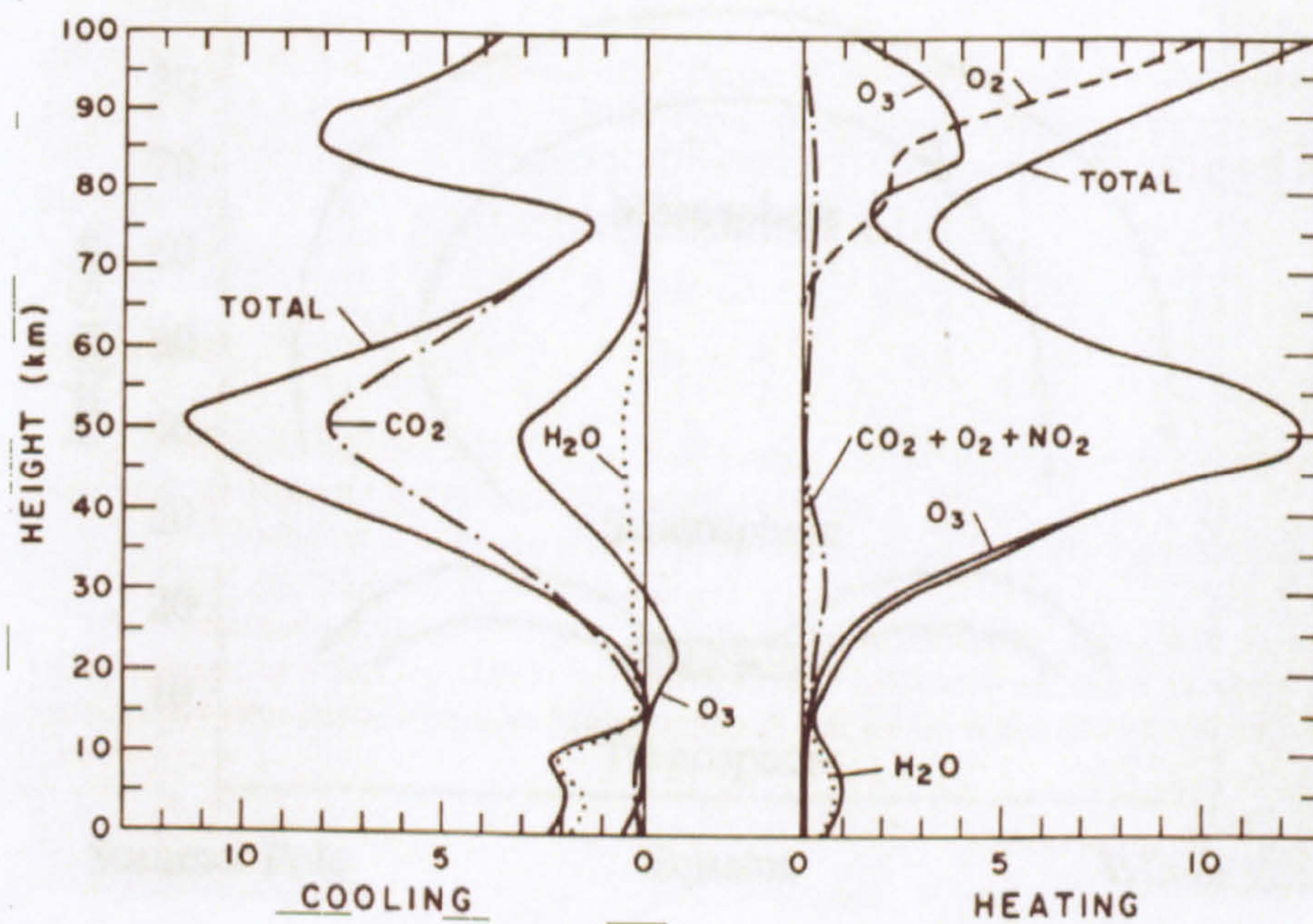


Figure I.10, Global mean UV heating rates and IR cooling rates in K day^{-1} , after London (1980).

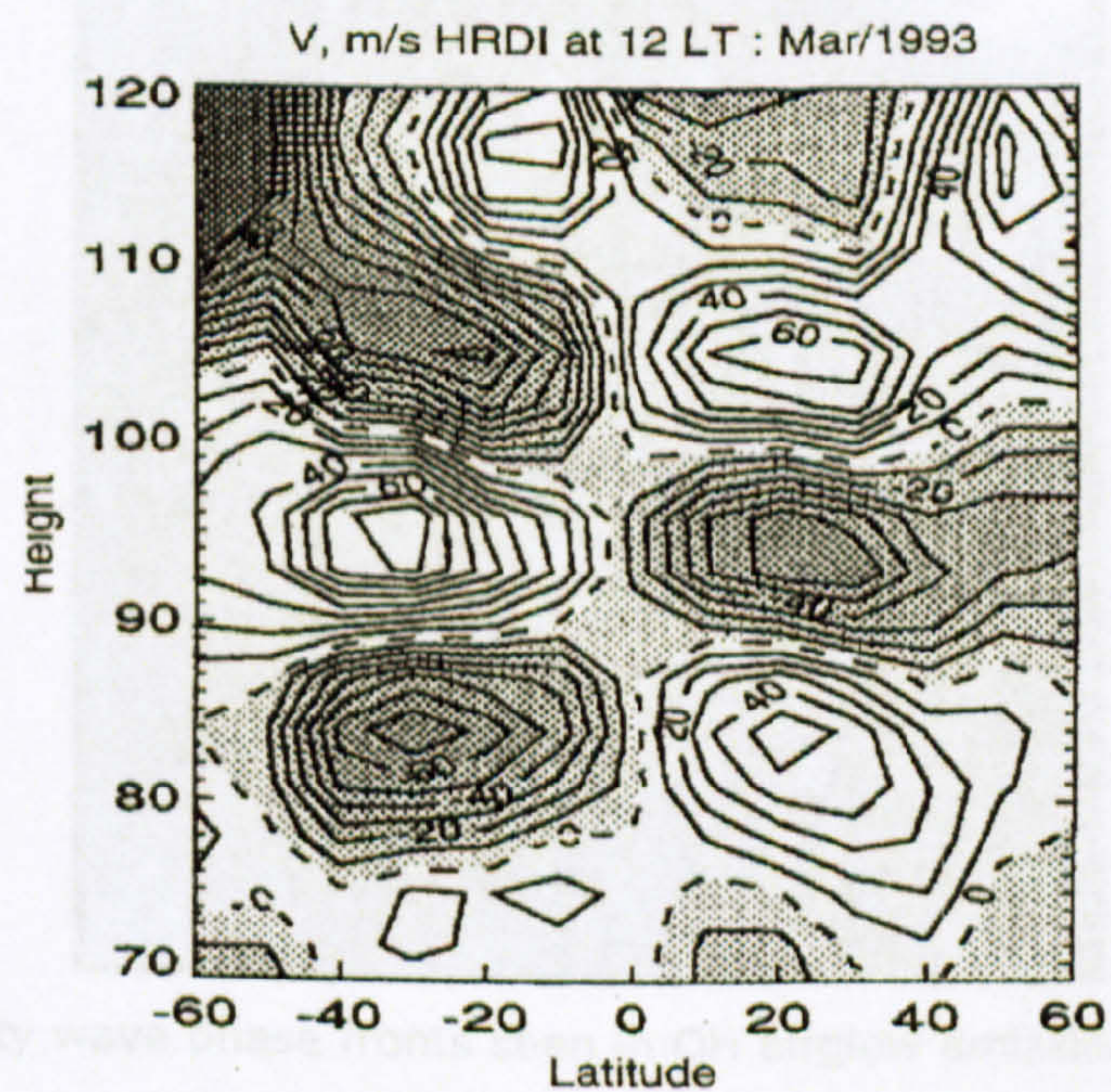


Figure I.11, Meridional winds observed by the UARS High Resolution Doppler Imager (HRDI) for March 1993 at 12 LT, after Yudin *et al.* (1997). Southward winds are shaded.

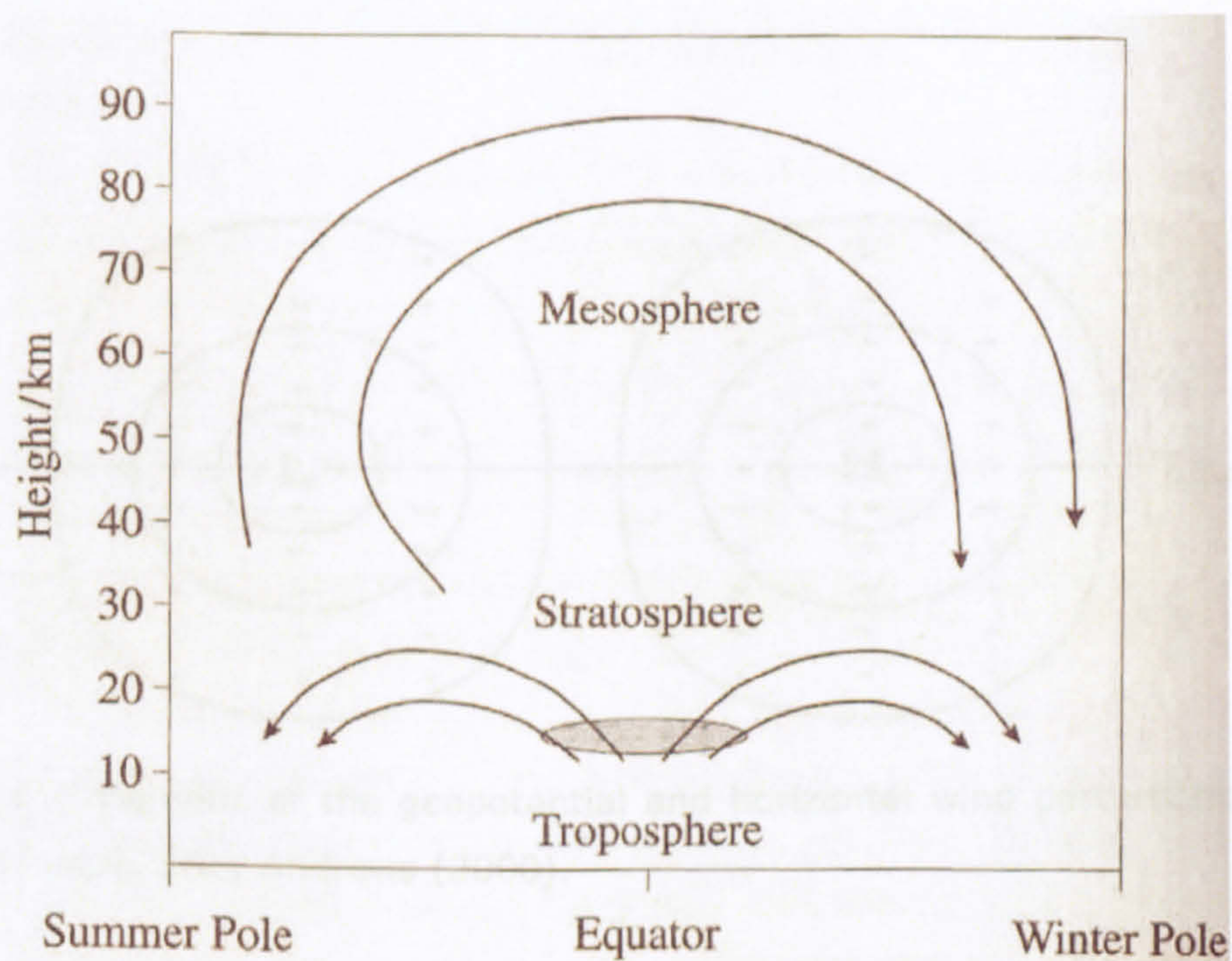


Figure I.12, Schematic of Brewer-Dobson stratospheric and Murgatroyd and Singleton mesospheric circulation, after Andrews (2000). The shaded region indicates the 'cold trap' identified by Brewer (1949).

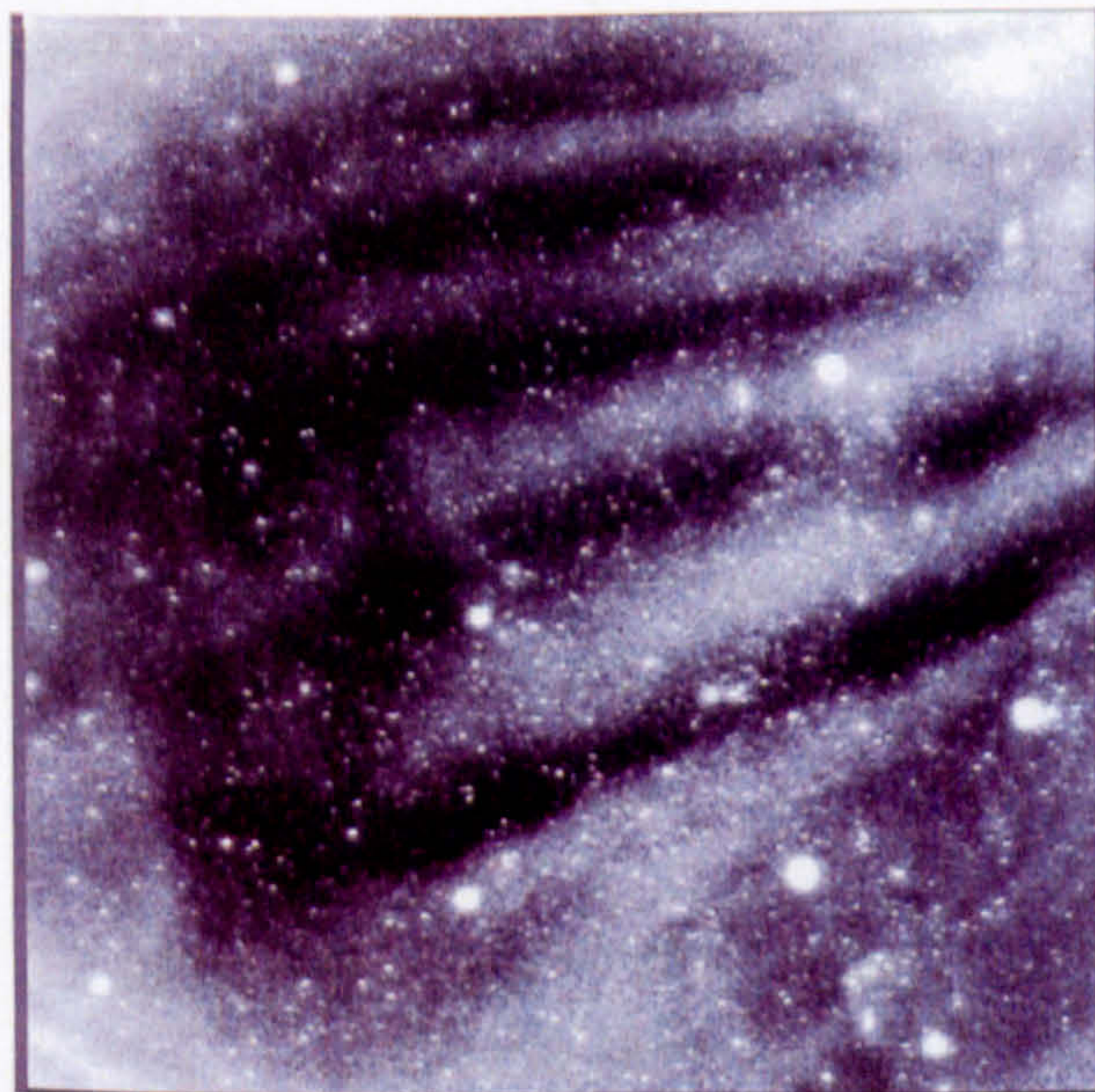


Figure I.13, Gravity wave phase fronts seen in OH airglow emissions at 87 km altitude, after Nakamura *et al.* (1998).

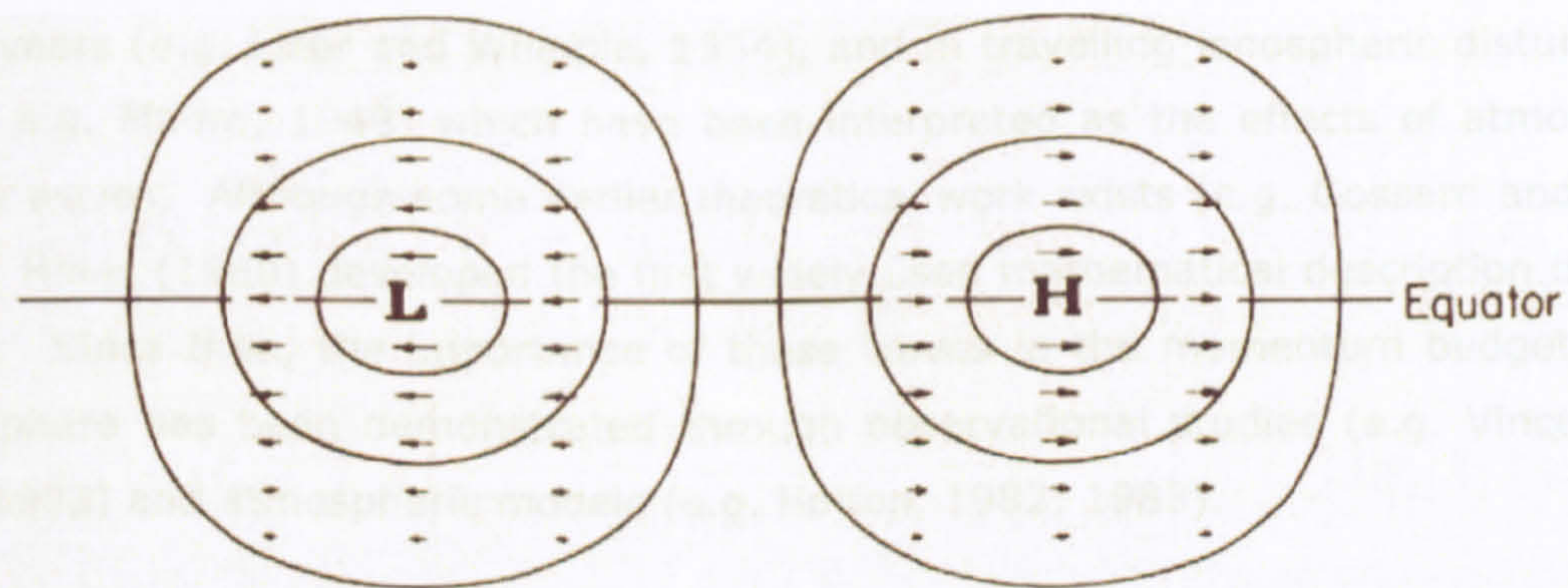


Figure I.14, Schematic of the geopotential and horizontal wind perturbations induced by a Kelvin wave, after Andrews (2000).

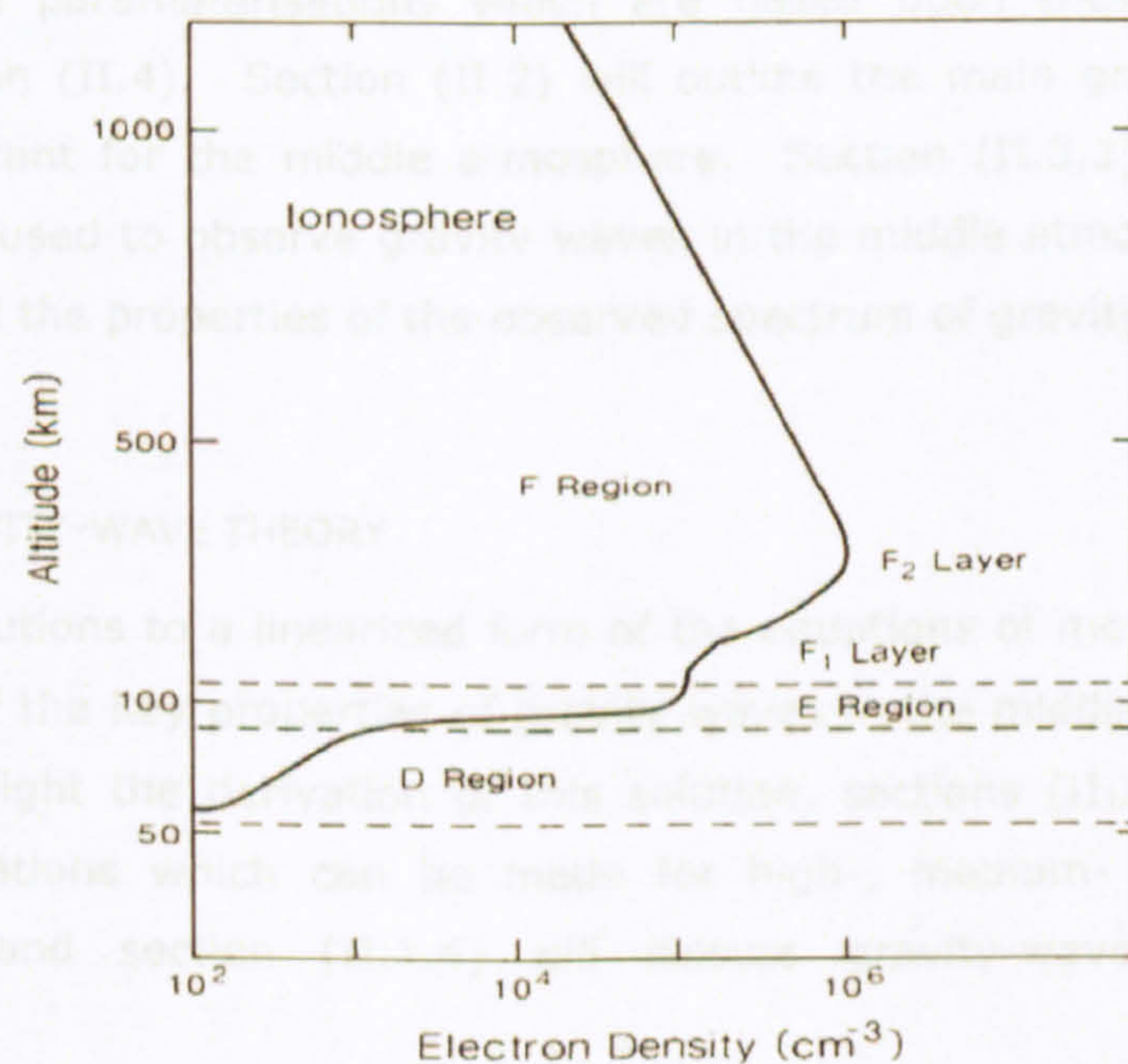


Figure I.15, Electron density profile in the ionosphere, after Kivelson and Russell (1995).

CHAPTER II, ATMOSPHERIC GRAVITY WAVES AND THEIR REPRESENTATION IN GLOBAL-SCALE MODELS

Wave-like features have been observed in meteor trails in the middle atmosphere for many years (e.g. Liller and Whipple, 1954), and in travelling ionospheric disturbances (TIDs, e.g. Munro, 1948) which have been interpreted as the effects of atmospheric gravity waves. Although some earlier theoretical work exists (e.g. Gossard and Munk, 1954), Hines (1960) developed the first widely used mathematical description of these waves. Since then, the importance of these waves in the momentum budget of the mesosphere has been demonstrated through observational studies (e.g. Vincent and Reid, 1983) and atmospheric models (e.g. Holton, 1982; 1983).

This chapter will discuss several theories of gravity-wave activity in the middle atmosphere. These can be broadly categorised into the linear theory of gravity waves, and a number of competing nonlinear theories. Section (II.1) will discuss the linear theory in detail and highlight some key properties of gravity waves. As there is no consensus on the merits of the numerous nonlinear theories (e.g. Medvedev and Klaassen, 1995; Hines, 2002), these will not be discussed at length here, but some of the gravity-wave parameterisations which are based upon these theories will be detailed in section (II.4). Section (II.2) will outline the main gravity-wave sources which are important for the middle atmosphere. Section (II.3.1) will introduce the main techniques used to observe gravity waves in the middle atmosphere and section (II.3.2) will detail the properties of the observed spectrum of gravity waves.

II.1 LINEAR GRAVITY-WAVE THEORY

Gravity-wave solutions to a linearized form of the equations of motion can be used to describe many of the key properties of gravity waves in the middle atmosphere. This section will highlight the derivation of this solution, sections (II.1.1) to (II.1.3) will discuss simplifications which can be made for high-, medium- and low-frequency gravity waves and section (II.1.4) will discuss gravity-wave propagation and dissipation.

The linear theory treats atmospheric gravity waves as small perturbations on a stably stratified background atmosphere which only varies vertically and the restoring force for these waves is the buoyancy that results from the displacement of air parcels from their stable position (hence the term gravity wave). To describe these perturbations, the fluid equations in Cartesian coordinates are used (e.g. Holton 1992),

$$\frac{du}{dt} - fv + \frac{1}{\rho} \frac{\partial p}{\partial x} = X \quad \text{II.1}$$

$$\frac{dv}{dt} + fu + \frac{1}{\rho} \frac{\partial p}{\partial y} = Y \quad \text{II.2}$$

$$\frac{dw}{dt} + \frac{1}{\rho} \frac{\partial p}{\partial z} + g = 0 \quad \text{II.3}$$

$$\frac{1}{\rho} \frac{d\rho}{dt} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad \text{II.4}$$

$$\frac{d\theta}{dt} = Q \quad \text{II.5}$$

along with the equation for potential temperature, in the form,

$$\theta = \frac{p}{\rho R} \left(\frac{p_0}{p} \right)^{\kappa}. \quad \text{II.6}$$

Following Fritts and Alexander (2003), these equations are simplified by assuming the forcing terms X , Y , which may be created by non-conservative wave processes, are equal to zero and the equations are then linearized such that perturbations are taken from a horizontally uniform, hydrostatic background state with wind $(\bar{u}, \bar{v}, 0)$, potential temperature $\bar{\theta}$, pressure \bar{p} and density $\bar{\rho}$ which vary only in z . These equations now have the form,

$$\frac{Du'}{Dt} + w' \frac{\partial \bar{u}}{\partial z} - fv' + \frac{\partial}{\partial x} \left(\frac{p'}{\bar{\rho}} \right) = 0 \quad \text{II.7}$$

$$\frac{Dv'}{Dt} + w' \frac{\partial \bar{v}}{\partial z} + fu' + \frac{\partial}{\partial y} \left(\frac{p'}{\bar{\rho}} \right) = 0 \quad \text{II.8}$$

$$\frac{Dw'}{Dt} + \frac{\partial}{\partial z} \left(\frac{p'}{\bar{\rho}} \right) - \frac{1}{H} \left(\frac{p'}{\bar{\rho}} \right) + g \frac{\rho'}{\bar{\rho}} = 0 \quad \text{II.9}$$

$$\frac{D}{Dt} \left(\frac{\theta'}{\bar{\theta}} \right) + w' \frac{N^2}{g} = 0 \quad \text{II.10}$$

$$\frac{D}{Dt} \left(\frac{\rho'}{\bar{\rho}} \right) + \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} - \frac{w'}{H} = 0 \quad \text{II.11}$$

$$\frac{\theta'}{\bar{\theta}} = \frac{1}{c_s^2} \left(\frac{p'}{\bar{\rho}} \right) - \frac{\rho'}{\bar{\rho}} \quad \text{II.12}$$

where the primed quantities show the perturbations to the background state, c_s is the speed of sound and D/Dt is the linearized form of the time derivative, given by

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x} + \bar{v} \frac{\partial}{\partial y}. \quad \text{II.13}$$

It is worth noting that only the perturbation terms have horizontal derivatives under these assumptions. Following Gill (1982), a WKB approximation is taken, assuming that (\bar{u}, \bar{v}) and N vary slowly over a wave cycle in the vertical such that the background shear terms in equations (II.7) and (II.8) can be neglected. Solutions of the form,

$$\left(u', v', w', \frac{\theta'}{\bar{\theta}}, \frac{p'}{\bar{p}}, \frac{\rho'}{\bar{\rho}} \right) = (\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\theta}, \tilde{p}, \tilde{\rho}) \exp \left[i(kx + ly + mz - \omega t) + \frac{z}{2H} \right] \quad \text{II.14}$$

where $(\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\theta}, \tilde{p}, \tilde{\rho})$ are the absolute and relative perturbations of the wind components, potential temperature, pressure and density. From equation (II.14), the monochromatic wave perturbations have wave-number components (k, l, m) and an Eulerian frequency ω (relative to the ground), can now be sought by substituting equation (II.14) into equations (II.7) to (II.12). This gives a set of equations for $(\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\theta}, \tilde{p}, \tilde{\rho})$,

$$-i\hat{\omega}\tilde{u} - f\tilde{v} + ik\tilde{p} = 0 \quad \text{II.15}$$

$$-i\hat{\omega}\tilde{v} + f\tilde{u} + il\tilde{p} = 0 \quad \text{II.16}$$

$$-i\hat{\omega}\tilde{w} + \left(im - \frac{1}{2H} \right) \tilde{p} = -g\tilde{\rho} \quad \text{II.17}$$

$$-i\hat{\omega}\tilde{\theta} + (N^2/g)\tilde{w} = 0 \quad \text{II.18}$$

$$-i\hat{\omega}\tilde{\rho} + ik\tilde{u} + il\tilde{v} + \left(im - \frac{1}{2H} \right) \tilde{w} = 0 \quad \text{II.19}$$

$$\tilde{\theta} = \tilde{p}/c_s^2 - \tilde{\rho}, \quad \text{II.20}$$

where $\hat{\omega} = \omega - k\bar{u} - l\bar{v}$ is the intrinsic frequency (the frequency observed by a reference frame which moves with the background wind). It is worth noting that the

$\exp\left(\frac{z}{2H}\right)$ term in equation (II.14) means that the gravity-wave perturbations increase

exponentially as the background density falls with height. Combining the equations (II.15) to (II.20) gives

$$N^2 \left(\frac{f^2 - \hat{\omega}^2}{2gH} + k^2 + l^2 + \frac{im(\hat{\omega}^2 - f^2)}{g} \right) = f^2 \left(\frac{1}{4H} - m^2 - \frac{g}{2Hc_s^2} \right) + \dots$$

$$\hat{\omega}^2 \left(\frac{f^2}{c_s^2} - \frac{1}{4H^2} + \frac{g}{2Hc_s^2} + k^2 + l^2 + m^2 - \frac{\hat{\omega}^2}{c_s^2} \right) + im(f^2 + \hat{\omega}^2) \left(\frac{g}{c_s^2} - \frac{1}{H} \right). \quad \text{II.21}$$

Given that $\frac{g}{c_s^2} = \frac{1}{H} - \frac{N^2}{g}$, equation (II.21) reduces to the form,

$$\hat{\omega}^2 \left(k^2 + l^2 + m^2 + \frac{1}{4H^2} - \frac{(\hat{\omega}^2 - f^2)}{c_s^2} \right) = N^2(k^2 + l^2) + f^2 \left(m^2 + \frac{1}{4H^2} \right). \quad \text{II.22}$$

Equation (II.21) is a solution for both acoustic waves ($\hat{\omega} > c_s/2H$) and internal gravity waves ($\hat{\omega} < N$). In the limit $c_s \rightarrow \infty$, equation (II.21) simplifies to the gravity-wave dispersion relation,

$$\hat{\omega}^2 = \frac{N^2(k^2 + l^2) + f^2 \left(m^2 + \frac{1}{4H^2} \right)}{k^2 + l^2 + m^2 + \frac{1}{4H^2}} \quad \text{II.23}$$

which can also be rewritten in terms of the vertical wavenumber in the Eulerian frame,

$$m^2 = \frac{(k^2 + l^2)(N^2 - \hat{\omega}^2)}{(\hat{\omega}^2 - f^2)} - \frac{1}{4H^2}. \quad \text{II.24}$$

The dispersion relation shows that the gravity-wave frequency and wavenumbers are related to the background atmosphere through both the buoyancy frequency N and horizontal wind (\bar{u}, \bar{v}) (through the intrinsic frequency $\hat{\omega}$). From equation (II.24), it can also be seen that for a wave to propagate vertically, for example from the lower atmosphere to the middle atmosphere, k , l and m must be real and $N > \hat{\omega} > |f|$.

For a wave packet, the energy propagation in the Eulerian frame is given by the velocity of the centre of the wave packet (wave group velocity, c_g), given by

$$\begin{aligned} (c_{gx}) &= \frac{\partial \omega}{\partial k} = \bar{u} + \frac{k(N^2 - \hat{\omega}^2)}{\hat{\omega} \left(k^2 + l^2 + m^2 + \frac{1}{4H^2} \right)} \\ c_{gy} &= \frac{\partial \omega}{\partial l} = \bar{v} + \frac{l(N^2 - \hat{\omega}^2)}{\hat{\omega} \left(k^2 + l^2 + m^2 + \frac{1}{4H^2} \right)} \\ c_{gz} &= \frac{\partial \omega}{\partial m} = \frac{-m(\hat{\omega}^2 - f^2)}{\hat{\omega} \left(k^2 + l^2 + m^2 + \frac{1}{4H^2} \right)} \end{aligned} \quad \text{II.25}$$

where $\hat{\omega}$ is taken to be positive by convention. Horizontal propagation of the gravity wave is in the direction (k, l) and m is negative for upward group velocity.

The phase propagation direction is given by $\mathbf{K} = (k, l, m)$, and the intrinsic phase trace speed

$$\hat{c}_{ph} = \frac{\hat{\omega}}{\sqrt{k^2 + l^2}}, \hat{c}_{pz} = \frac{\hat{\omega}}{m} \quad \text{II.26}$$

where the subscripts z and h indicate the horizontal and vertical intrinsic phase trace speeds (the apparent speed of phase fronts observed in the horizontal and vertical planes moving with the background flow). Comparing equations (II.25) and (II.26) it can be seen that the phase propagation direction is at right angles to the group velocity. For example, a gravity wave which has eastward ($k > 0$) and upward ($m < 0$) group velocity has intrinsic phase trace speeds which are eastward ($\hat{c}_{px} = \hat{\omega}/k$) and downward ($\hat{c}_{pz} = \hat{\omega}/m$) relative to the background wind.

The above equations can be applied to gravity waves with a large range of frequencies. By simplifying these equations so that they apply only to gravity waves within a range of frequencies, several further results can be easily obtained. The following sections will show approximations which can be made for high-, medium- and low-frequency waves and highlight some key results.

II.1.1 HIGH-FREQUENCY WAVES

For gravity waves with small vertical wavenumber m , the change in background density is significant over the vertical wavelength of the wave. In the middle atmosphere, where $H \sim 7$ km, this relates to gravity waves with vertical wavelengths ≥ 30 km. In the limit where the vertical wavelength grows and $m \rightarrow 0$, the gravity wave undergoes total internal reflection. This point is referred to as a turning level. In the special case where two turning levels are present in the atmosphere, a gravity wave may be ducted. Equation (II.23) can be simplified in the case of $m = 0$ and ignoring the effects of the Earth's rotation. This gives a maximum horizontal intrinsic phase trace speed, $\hat{c}_{ph\max}$, which is possible prior to the wave undergoing total internal reflection, given by

$$|\hat{c}_{ph\max}| = N \left(k^2 + l^2 + \frac{1}{4H^2} \right)^{-1/2}. \quad \text{II.27}$$

Using this relationship, it is possible to show that waves with horizontal wavelengths < 10 km are easily reflected and remain trapped in the lower atmosphere, so generally only waves with horizontal wavelengths > 10 km are considered to be important for the middle atmosphere (e.g. Marks and Eckermann 1995).

For high-frequency gravity waves where $\hat{\omega} \gg f$, where the Coriolis force can be neglected and for which $m \gg 1/2H$, the dispersion relation can be simplified to,

$$\hat{\omega}^2 = \frac{N^2(k^2 + l^2)}{k^2 + l^2 + m^2} = N^2 \cos^2 \alpha \quad \text{II.28}$$

where α is the angle between the group velocity and the vertical. It can be seen that the plane wave gravity waves have phase propagation directions and group velocities which are dependent on the intrinsic frequency of the wave. In this high-frequency case, the group velocity simplifies to,

$$c_{gh} = \bar{u}_h + \frac{Nm}{k_h^2 + m^2} \frac{m}{(k_h^2 + m^2)^{1/2}}$$

$$c_{gz} = \frac{Nm}{k_h^2 + m^2} \frac{-k_h}{(k_h^2 + m^2)^{1/2}} \quad \text{II.29}$$

where the subscript h denotes the horizontal components of the wave and group velocity.

II.1.2 MEDIUM-FREQUENCY WAVES

For medium-frequency gravity waves with $N \gg \hat{\omega} \gg f$, the linear equations can be greatly simplified and several key results obtained. In the middle atmosphere, where $m \gg k_h$, the dispersion relation can be simplified to,

$$\hat{\omega} = N \left| \frac{k_h}{m} \right| \quad \text{II.30}$$

and equation (II.24) can be simplified to,

$$|m| = \frac{N}{|\hat{c}_{ph}|} \quad \text{II.31}$$

Equations (II.30) and (II.31) show that both the vertical wavenumber m and the intrinsic frequency $\hat{\omega}$ are proportional to the intrinsic phase trace speed $\hat{c}_{ph} = c_{ph} - \bar{u}_h$.

However, as we consider the horizontal wind to vary only as a function of height, k_h and therefore c_{ph} (the Eulerian horizontal phase trace speed = $|\omega/k_h|$), remains constant for a vertically propagating gravity wave. One further important result is that at a critical level, where $c_{ph} = \bar{u}_h(z)$, the vertical wavenumber tends to infinity and the vertical wavelength tends to zero.

For medium-frequency gravity waves, the group velocity simplifies to,

$$(c_{gh}, c_{gz}) = (\bar{u}_h, 0) + \left(\frac{\hat{\omega}}{k_h}, \frac{-\hat{\omega}}{m} \right), \quad \text{II.32}$$

and the intrinsic phase trace speed is given by,

$$\hat{c}_{ph} = \frac{\hat{\omega}}{k_h}, \hat{c}_{pz} = \frac{\hat{\omega}}{m}. \quad \text{II.33}$$

Lighthill (1978) showed that as a gravity wave approached a critical level, its vertical group velocity tends to zero (as $m \rightarrow \infty$), so this theoretical limit is never reached. It is worth noting that as a wave approaches this level, nonlinear dissipation and instability mechanisms, which shall be discussed in section (II.4) become increasingly important.

For medium-frequency gravity waves, simple polarisation relations, which relate the amplitudes of the wave perturbations in different quantities to one another, can be written. One important polarisation relation is given by ,

$$\tilde{w} = -\frac{k_h}{m} \tilde{u}_h = \frac{k_h \hat{c}_{ph}}{N} \tilde{u}_h = \frac{\hat{\omega}}{N} \tilde{u}_h. \quad \text{II.34}$$

From this relation and equation (II.32), it can be seen that as a wave propagates through a region in which $|\hat{c}_{ph}|$ increases with altitude, the vertical wavelength increases (m decreases) and $\hat{\omega}$ increases and therefore c_{gz} increases, whereas the opposite is true in a region in which $|\hat{c}_{ph}|$ decreases with altitude, such as for a wave nearing a critical level. The importance of this change in $\hat{\omega}$, referred to as Doppler shifting, will be discussed in more detail in section (II.4).

II.1.3 LOW-FREQUENCY WAVES

For low-frequency gravity waves with $\hat{\omega} \sim f$, the effects of the Earth's rotation become important. Such waves are called inertia-gravity waves. These waves have scales which can be resolved by some global-scale models so their effects are not included in the gravity-wave parameterizations discussed in section (II.4), but important results can still be yielded by considering linear gravity-wave theory in this frequency range.

The dispersion relation for low-frequency gravity waves can be written as

$$\hat{\omega}^2 = N^2 \frac{k_h^2}{m^2} + f^2. \quad \text{II.35}$$

As discussed in section (II.1.2), a wave which nears a critical level will move towards lower $\hat{\omega}$ and the effects of rotation will become important. From equation (II.35), it can be seen that in the limit of $m \rightarrow \infty$, $\hat{\omega} \rightarrow f$, rather than zero as implied by the medium frequency gravity-wave dispersion relation, so critical coupling does not occur.

II.1.4 CONSERVATIVE PROPAGATION AND DISSIPATION

Gravity waves which propagate without dissipation through a background atmosphere which varies only in the vertical conserve the flux of horizontal momentum which they transport (Andrews and McIntyre, 1978). The vertical flux of this horizontal momentum is given by,

$$(F_{Px}, F_{Py}) = \bar{\rho} (1 - f^2 / \hat{\omega}^2) (\overline{u'w'}, \overline{v'w'}). \quad \text{II.37}$$

A dissipating gravity wave does not conserve these quantities. In the middle atmosphere, this dissipation can occur through IR radiative damping (e.g. Spiegel, 1957; Fels, 1982), wave-wave and wave-mean flow interactions (e.g. Dunkerton, 1982) and wave breaking instability processes which will be discussed in section (II.4.2). In the thermosphere, viscosity, molecular diffusion and ion damping also become important wave damping mechanisms (e.g. Hines, 1960).

The negative of the zonal mean flux of horizontal momentum carried by gravity waves is equivalent to their contribution to the EP flux (see section I.3.3). The divergence of the EP flux provides the wave driving (X and Y in equations II.1 and II.2) on the background flow, given by,

$$(\bar{X}, \bar{Y}) = -\frac{\varepsilon}{\bar{\rho}} \frac{\partial}{\partial z} (F_{Px}, F_{Py}) \quad \text{II.38}$$

where ε is an efficiency or intermittency factor (Holton, 1982) associated with the gravity-wave dissipation averaged over a volume or time interval (this factor must be estimated in many gravity-wave parameterisations is a large contribution to the uncertainty of their estimates and will be discussed in section II.4). It is worth noting that the sign of the wave driving is equal to the sign of \hat{c}_{ph} .

Gravity waves also have an associated energy flux and wave dissipation leads to local heating of the background atmosphere. This effect is often neglected in global scale parameterisations, but has been included in some model simulations and its effects will be discussed in section (II.4).

Gravity-wave dissipation often results in turbulent eddy mixing of energy, momentum and chemical constituents, which is estimated by many gravity-wave parameterisations. Lindzen (1981) showed that for a monochromatic gravity wave, the momentum eddy diffusion coefficient associated with wave dissipation is given by,

$$D_M = \frac{\hat{c}(\bar{X}, \bar{Y})}{N^2} \quad \text{II.39}$$

which is related to the thermal eddy diffusion coefficient via the Prandtl number (the kinematic viscosity divided by the thermal diffusivity) associated with gravity-wave dissipation. The value of this Prandtl number has not been well constrained, but numerical and theoretical models of gravity wave convective instabilities have estimated it to be large and highly variable (Chao and Schoeberl, 1984; Fritts and Dunkerton, 1985; Coy and Fritts, 1999; McIntyre, 1989). Global-scale model studies have shown that to reproduce the observed atmospheric structure, the momentum diffusion coefficient predicted by equation (II.39) must be decreased by some arbitrary scaling factor (Strobel *et al.*, 1985). This again leads to an uncertainty in model estimates of gravity-wave effects on the middle atmosphere. The uncertainties, as well as some limitations of the linear theory of gravity waves will be addressed in section (II.4) and several nonlinear theories will be briefly discussed.

II.2 GRAVITY-WAVE SOURCES

On a global scale, upward propagating gravity waves generated in the troposphere and lower stratosphere are generally considered to be the most important source of gravity waves for the middle atmosphere (e.g. McIntyre, 1987). This is because these waves grow exponentially with altitude and dissipate in the middle atmosphere, while waves generated at higher altitudes, such as those created by auroral processes, or ones which propagate downwards tend not to have a significant momentum input to the middle atmosphere. As this study does not focus on the physics of the troposphere, a detailed discussion of tropospheric gravity-wave sources will not be included here, but the main gravity-wave sources will be noted.

Lower-atmosphere gravity-wave sources can be broken down into five main categories:

1. Orographic gravity waves are generated by wind flowing over mountain ranges or sea-land boundaries. As these waves are generated by stationary sources, they have $c_{ph} = 0$ in the Eulerian frame of reference. Typical horizontal wavelengths for orographic gravity waves are in the range of tens to hundreds of km.

2. Gravity waves can be generated by convective regions, especially deep convection cells in the troposphere and thunderstorms. These waves can have a wide range of horizontal scales.
3. Shear motions can generate gravity waves through nonlinear interactions between the Kelvin-Helmholtz instability and propagating modes (Fritts, 1984; Chimonas and Grant, 1984).
4. As an unbalanced flow (such as an instability in the jet stream) returns to a balanced state through geostrophic adjustment, inertio-gravity waves can be generated. These waves radiate energy away from the instability region.
5. Nonlinear wave-wave interactions may be important in generating and altering the evolution of gravity waves, although their role and importance is still a matter of debate. A number of laboratory and numerical models have shown that they might be important and Klostermeyer (1984, 1990) has argued that there are radar observations which show this wave generation mechanism.

In global-scale atmospheric models, gravity-wave sources can either be prescribed (as is the case with the middle and upper atmosphere models presented in this study), or can be self-consistently calculated (as is the case with some high-resolution models which include tropospheric processes and topography, e.g. Scaife *et al.*, 2002). These differences in gravity-wave source definition can make inter-model comparisons problematic. Further, as the global and temporal picture of gravity-wave sources has been poorly constrained by observations (see section II.3.2), this remains a major source of uncertainty when including gravity-wave effects in global-scale models (see chapter VI for further discussion). However, it is worth noting that global-scale models can also be used to further constrain our global picture of gravity-wave sources, which shall be discussed in chapters (IV), (V) and (VI).

II.3 OBSERVATIONS

The study of atmospheric gravity waves is driven by both theoretical work, such as studies using gravity-waves parameterisations in global-scale models, and through observations of gravity waves and their effects in the atmosphere. The earliest observations of gravity waves were reported by Munro (1948) who observed gravity-wave effects on charged particles in the ionosphere (TIDs) and these were first explained as atmospheric waves by Martin (1950). This section will discuss the observation of tropospherically generated gravity waves, which are observed throughout the middle atmosphere using a variety of different techniques. Section (II.3.1) will outline the main observational techniques which have been used to study

middle atmosphere gravity waves and section (II.3.2) will summarise some key observed properties of gravity-wave spectra.

II.3.1 OBSERVATIONAL TECHNIQUES

Methods of observing atmospheric gravity waves in the middle atmosphere can be split into three main categories: ground based, satellite based and *in situ* measurements. This section will discuss all three categories and highlight the uses of various techniques, although a comprehensive review of every observational technique used to observe gravity waves will not be presented (see the Middle Atmosphere Program handbooks volumes 12-15 for a more detailed review).

II.3.1.1 MESOSPHERIC, STRATOSPHERIC AND TROPOSPHERIC (MST) RADARS

MST radars which operate at UHF and VHF measure winds, waves and turbulent parameters by observing the intensities, Doppler shifts and spectral widths of echoes from refractive index irregularities in the atmosphere. They have typical spatial resolutions of order 30 m and temporal resolutions of order 1 minute along the beam. Using three beams, the three wind components can be resolved. In the troposphere and lower stratosphere MST radars can observe continuously. In the mesosphere MST radars can measure between approximately 75 – 105 km using reflections from ionised meteor trails during the day and between 80 – 105 km altitude during day and night when there is a sufficient electron density gradient. MST radars can be used to measure wind variances with gravity wave periods, momentum fluxes and the kinematic energy of gravity waves. As observations in the upper stratosphere are difficult to make with MST radars (e.g. Riggins *et al.*, 1997), other observational techniques are required to observe gravity waves in this region.

II.3.1.2 PARTIAL REFLECTION RADARS

Partial reflection MF radars using the spaced antenna technique have been used to measure horizontal wind fields in the 60 – 100 km altitude range during the day and 80 – 100 km at night. These radars have relatively good vertical resolution and have been used to measure gravity waves with periods over ~5 minutes (e.g. Vincent and Fritts, 1987). MF radars can operate continuously but the reflections are often weak.

II.3.1.3 IONOSPHERIC D1 TECHNIQUE

LF radio wave absorption in the lower ionosphere can be used to measure drift velocities in the lower thermosphere, around 95 km altitude. This technique uses oblique incidence on the ionosphere, with the receiver located around 100 km away from the transmitter. In the case of the Collm D1 system, discussed in chapter (IV), commercial radio stations are used as the transmitters. This technique has been used to observe gravity waves with periods around 10 – 100 minutes (e.g. Bošková and Laštovička, 2001).

II.3.1.4 METEOR RADARS

Meteor radars measure line-of-sight velocities of meteor trails which are influenced by the background wind and wave fields. The distribution of meteor trails is random in both space and time so the radar measurements must be averaged over space and time to get measurements of the horizontal wind, which means that meteor radars are most suited to measuring longer period oscillations such as inertio-gravity waves and atmospheric tides. However, as meteor radars have good vertical resolution it is possible to resolve gravity wave amplitudes and vertical wavelengths and they can be used to study small-scale wind structures (~10 km) created by gravity waves as well as measuring the power in horizontal wind perturbations created by gravity waves. Meteor radar measurements are typically in the height range 80 – 100 km with vertical resolution around 2 km (e.g. Manson *et al.*, 2004).

II.3.1.5 LIDARS

Two types of lidar are used to observe gravity waves. Rayleigh lidars use Rayleigh scattering from atmospheric molecules to measure neutral densities and temperatures. They have a height range of around 30 – 90 km, allowing them to measure gravity wave induced temperature perturbations in the upper stratosphere. Lidars have excellent height and temporal resolution (~1 km and a few minutes), but are subject to meteorological conditions (cloud cover) and some lidars can only operate during the night. They have been used to measure gravity waves with vertical wavelengths above ~1 km (e.g. Wilson *et al.*, 1991). Resonant scatter from ionised iron and sodium is also used to measure gravity waves where meteor deposition creates layers of metal ions between ~80 – 100 km (e.g. Collins and Smith, 2004).

II.3.1.6 ROCKETS

Payloads onboard sounding rockets and falling spheres deployed by them, can be used to measure temperatures and horizontal winds up to ~ 70 km altitude with vertical resolutions ~ 1 km (e.g. Philbrick *et al.*, 1983). By subtracting a background state from these profiles, and assuming all small-scale perturbations are due to gravity waves, their amplitudes can be measured. Their temporal resolution is poor, with repeat soundings typically every few days during campaigns, so they can be used to infer seasonal and long-term variability. Using neutral density observations, sounding rockets can also be used to measure turbulence created by dissipating gravity waves (e.g. Rapp *et al.*, 2004).

II.3.1.7 BALLOONS

Radiosondes can be used to measure temperature and horizontal winds in the troposphere and lower stratosphere. By subtracting a low-order polynomial fit to the temperature and wind profiles and assuming all perturbations with height are due to gravity waves it is possible to measure short vertical wavelength gravity waves in the lower stratosphere over periods of up to 3 days, allowing them to observe inertio-gravity waves (e.g. Sato *et al.*, 1994). By equipping them with two or more anemometers they can be used to measure temperatures and winds on scales ~ 1 m to obtain information on small scale turbulence created by dissipating gravity waves.

II.3.1.8 AIRCRAFT

Aircraft which fly in the lower stratosphere can carry inertial platforms, air motion sensors, thermometers and trace chemical sensors which allow them to measure horizontal wavelengths of gravity waves and turbulence associated with wave dissipation over small horizontal scales. Aircraft can be used to study localised gravity-wave sources such as mountains (e.g. Gary, 1989).

II.3.1.9 OPTICAL MEASUREMENTS

Using photographic techniques, it is possible to study atmospheric gravity waves through their effects on airglow and noctilucent clouds around the mesopause region. Changes in the background through changes in density and temperature alter the airglow emission rates and opacity of noctilucent clouds (shown in figure II.1). These observations provide detailed information on the horizontal structure of gravity waves and horizontal phase velocities.

II.3.1.10 Spacecraft

Space-based platforms can, in theory, be used to measure the global climatology of gravity waves. However, as gravity waves sources are not continuous and the waves often have short periods and small scales, observations from spacecraft have so-far only sampled a small fraction of the gravity-wave spectrum in the middle atmosphere. Infrared (Gille and Russell, 1984) and microwave limb sounders (Wu and Waters, 1996a, b) and infrared telescopes (Preusse *et al.*, 1999) have been used to measure gravity wave vertical temperature perturbations in the stratosphere and mesosphere. These have been used to measured gravity waves with vertical wavelengths $\sim 5\text{-}50$ km between $\sim 15\text{-}80$ km altitude. Using GPS occultation data temperature profiles (Rocken *et al.*, 1997), global maps of gravity wave potential energies have also been found (Tsuda *et al.*, 2000). This technique measures gravity waves with vertical wavelengths $\sim 2\text{-}10$ km in the stratosphere.

II.3.2 OBSERVED GRAVITY-WAVE SPECTRA

Inferring climatologies from gravity wave observations is problematic as gravity waves exist across a wide range of scales and different observational techniques, which sample only a portion of the whole gravity-wave spectrum, can themselves be responsible for many of the patterns seen in observed data (Alexander, 1998). Further, observations taken using different techniques or at different locations often disagree. However, knowledge of some broad characteristics of gravity-wave spectra is required in order to represent atmospheric gravity waves in global-scale models (see section II.4), and these will be introduced here. Some limitations of using these broad characteristics will be discussed in chapter (VI).

Plotting the observed gravity wave energy as a function of vertical wavenumber, m , yields perhaps the best known feature of gravity-wave spectra. The energy of the gravity-wave spectrum is seen to peak at a characteristic wavenumber m_* , while the momentum of the spectrum peaks at high m values. The energy of the gravity-wave spectrum is observed to decrease as m^{-3} for $m > m_*$ (VanZandt, 1982). This slope is observed to be approximately constant with altitude, as shown in figure (II.2), although the characteristic wavenumber m_* is seen to vary with height (see discussion below). This slope has been reported numerous times and appears, regardless of geographical location or time (e.g. Tsuda *et al.*, 1989) and as such is referred to as the universal spectrum of gravity waves.

Figure (II.2) shows that within the m^{-3} portion of the spectrum, wave amplitudes do not increase with altitude. This limiting of the amplitude growth is referred to as saturation. As altitude increases, and m_* decreases, an increasing portion of the wave spectrum becomes saturated. This saturation implies that some form of energy dissipation is taking place which constrains the amplitude of waves within this portion of the spectrum. For large-scale inertia-gravity waves, the Kelvin-Helmholtz instability is widely believed to be responsible for limiting the amplitude growth of gravity waves. As the onset of convective instabilities are more rapid, these are believed to dominate the smaller-scale waves (see section II.4.2) in the mesosphere and lower thermosphere, which are responsible for the bulk of the gravity-wave momentum flux in this region (Fritts and Werne, 2000).

Observations of gravity waves with low m values (large vertical wavelengths) are constrained by the altitude range over which observations can be made. However, it is generally accepted that the wave energy decreases for values below m_* , which also sets a finite limit on the total energy in the gravity-wave spectrum. This is believed to be due to various saturation processes (see Fritts and Alexander, 2003, for further discussion). Gravity wave horizontal velocity and temperature perturbations have been observed to vary as a function of frequency according to ω^{-p} for frequencies $f < \omega < N$. The variable p has been observed to be approximately 5/3 (e.g. VanZandt, 1982; Balsley and Carter, 1982). Combining these observed spectral properties, Fritts and VanZandt (1993) showed that the total observed gravity wave energy can be written as,

$$E(\mu, \omega, \phi) = E_0 A(\mu) B(\omega) \Phi(\phi) \quad \text{II.40}$$

where,

$$A(\mu) = A_0 \frac{\mu^s}{1 + \mu^{s+t}} \quad \text{II.41}$$

and,

$$B(\omega) = B_0 \omega^{-p} \quad \text{II.42}$$

where $\Phi(\phi)$ describes the azimuthal distribution of propagation directions, $\mu = m/m_*$

and the coefficients are such that $\int_0^\infty A(\mu) d\mu = 1$, $\int_f^N B(\omega) d\omega = 1$, and $\int_0^{2\pi} \Phi(\phi) d\phi = 1$, with

values of the spectral slopes $s = 1$, $t = 3$ and $p = 5/3$.

VanZandt and Fritts (1989) showed that the observed total wave energy density varies with height according to,

$$E(m) = \frac{E_0 A(\mu)}{m_*} \sim \left(\frac{N}{N_0} \right)^{1/2} \exp\left(\frac{z}{H_E} \right) \frac{A(\mu)}{m_*} \quad \text{II.43}$$

and,

$$m_* \sim \left(\frac{N}{N_0} \right)^{3/4} \exp\left(\frac{-z}{H_E} \right) \quad \text{II.44}$$

where $H_E \sim 2.3H$ (Warner and McIntyre, 1996). Equation (II.44) shows that the characteristic vertical wavenumber m_* varies as a function of height. Observations have shown m_* to vary from $\sim 3 \text{ rad km}^{-1}$ near the tropopause to $\sim 1/3 \text{ rad km}^{-1}$ near the mesopause (e.g. Vincent, 1984).

Gravity-wave activity has been observed to vary as a function of latitude. Alexander (1997) used observations from the MLS instrument onboard the UARS satellite to calculate gravity wave temperature variances in the stratosphere, shown in figure (II.3). As can be seen, the peak in gravity-wave activity in the stratosphere tends to occur at mid to high latitudes, with minima near the equator and at high latitudes. It is worth noting that this distribution is highly variable with time, but observations such as this can be useful for constraining gravity-wave sources in global-scale models, which shall be discussed in the following section.

II.4 GRAVITY-WAVE PARAMETERISATIONS

Chapter (I) introduced the importance of atmospheric gravity waves to the dynamics of the middle atmosphere. In the discussion in Section (II.2), we have seen that gravity waves with horizontal wavelengths as low as 10 km can propagate up to the middle atmosphere, however, these horizontal scales are too small to be resolved by most global-scale atmospheric models (the STEVE-3D model presented in chapter III has a meridional grid spacing of 500 km). While the effects of larger-scale gravity waves have been successfully incorporated into some high resolution global-scale models (e.g. Hamilton *et al.*, 1999), gravity-wave parameterisations remain the most computationally efficient way of including a physically reasonable description of the gravity-wave spectrum on the middle atmosphere. A number of gravity-wave parameterisations have been developed, each based on different assumptions of gravity-wave dissipation. However, with the exception of Rayleigh friction (Section II.4.1), all gravity-wave parameterisations are one-dimensional ray-tracing schemes which track the vertical evolution of gravity waves. As there are relatively few observed parameters which allow the stringent testing of the theories behind these parameterisations (Gardener, 1996), a large number of gravity-wave parameterisation

have been developed for use in atmospheric models. The most common gravity-wave parameterisations will be outlined in this section, with particular emphasis on the parameterisations used in this study.

II.4.1 RAYLEIGH FRICTION

Rayleigh friction is the simplest description of gravity-wave effects on the middle atmosphere. It has been used by many atmospheric models (e.g. Leovy, 1964; Holton and Wehrbein, 1980) to reduce the strength of the mid-latitude thermal jets (see section I.3.2) by imposing a force on the horizontal winds which is linearly proportional to, and which opposes the horizontal wind. A height-dependant coefficient of Rayleigh friction is used, such as that shown in figure (III.10), so that this drag force increases with altitude to become strong in the upper mesosphere where the largest differences between the radiative equilibrium and observed winds are seen (see Section I.3.2).

There are a number of problems with this simple description of gravity-wave activity which have been noted by numerous authors (e.g. Lindzen, 1981; Garcia and Solomon, 1985). As the drag is proportional to the horizontal wind speed, Rayleigh friction can reduce the thermal winds in atmospheric models, but not cause the jet closure seen in the MLT region. Further, as Rayleigh friction only considers the *in situ* horizontal wind speeds, the effects of critical levels, which filter out portions of the gravity-wave spectrum at lower altitudes, are not included. This means that atmospheric models using Rayleigh drag cannot reproduce features such as the jet reversal seen in the MLT region.

II.4.2 LINEAR SATURATION, LINDZEN, 1981

As an attempt to improve on the Rayleigh friction representation of gravity-wave drag, Lindzen (1981, Lindzen hereafter) developed the linear saturation theory. This was further developed and first implemented as a full gravity-wave parameterisation by Holton (1982). Using the mid-frequency dispersion relation (equation II.30) and WKB approximation, Lindzen derived an analytical solution to a single gravity wave propagating in a slowly-varying background wind. Under this approximation, gravity waves which do not reach a critical level grow exponentially in amplitude with height. The temperature perturbations induced by these waves varies according to,

$$\frac{\partial T'}{\partial z} \approx i \tilde{T} T^{1/2} \bar{T}^{-1/2} \lambda^{3/2} \exp\left\{i \int \lambda dz\right\} \exp\left\{\frac{z}{2H} (ikx - c_{ph} t)\right\} \cos(l y + \phi) \quad \text{II.45}$$

where Γ is the adiabatic lapse rate, ϕ is the horizontal phase of the wave and the vertical wavelength, λ , is given by,

$$\lambda^2 \approx \frac{N^2 \left(1 + \frac{l^2}{k^2} \right)}{(\bar{u} - c_{ph})^2}. \quad \text{II.46}$$

From equations (II.45 and II.46), it can be seen that the temperature perturbation is related to the background temperature and wind. Lindzen argued that, at a certain height, referred to as the breaking level, these temperature perturbations would become convectively unstable (the temperature lapse rate from the wave perturbation and mean state exceeds the adiabatic lapse rate, sometimes referred to as the saturation limit). At the breaking level, convective overturning and small scale turbulence is assumed to constrain the wave amplitude to that at which instability onset occurs. This small scale turbulence creates both mixing and an acceleration of the background flow. The resultant force exerted on the background flow by the i^{th} gravity wave (in the form presented by Holton, 1982) is given by,

$$F_i = \frac{\varepsilon k (c_{phi} - u)^3}{2N} \left[\frac{1}{H} + \frac{3}{c_{phi} - u} \right] \frac{du}{dz} \quad \text{II.47}$$

and the momentum eddy diffusion coefficient is given by,

$$D_{eddy} = Xk \frac{\hat{c}_{phi}^4}{N^3} \left[\frac{1}{2H} - \frac{3}{2} \frac{du/dz}{\hat{c}_{phi}} \right]. \quad \text{II.48}$$

(Please note this is a rearranged form of the equation presented by Lindzen, from Meyer, 1999a.) As the Lindzen parameterisation assumes that the force $F_i = 0$ below the breaking level, the calculated force is often artificially adjusted to include a smooth tapering across a few kilometres below the breaking level in order to avoid numerical instabilities associated with the sudden onset of gravity-wave drag.

Many authors have used this parameterisation in atmospheric models to reproduce the observed jet reversal in the MLT region and associated mesopause temperature anomaly (e.g. Garcia and Solomon, 1985). However, several limitations to the theory can be noted here:

- The Lindzen theory provides no constraints on the smooth tapering below the breaking level or the intermittency factor used in equation (II.47), which may be set differently by different models in order to tune the model results to agree with observations.
- From equation (II.47) it can be seen that for any given breaking gravity wave, the force imposed on the atmosphere will tend to maximise when $c_{phi} - u$ maximises.

This means that, in the absence of critical levels, the gravity-wave drag calculated using this parameterisation will tend to damp resolved wave motions in the atmosphere (chapter VI will discuss the implications of this in further detail).

- The Lindzen parameterisation only considers wave-mean-flow interactions and has no representation of wave-wave interactions. While there is still debate over the importance of wave-wave interactions as a gravity-wave dissipation mechanism, it has been argued by a number of authors (e.g. Medvedev and Klaassen, 2000) that this simplification leads the Lindzen parameterisation to overestimate the height at which wave breaking occurs.

II.4.3 ATMOSPHERIC TRANSMISSION, MATSUNO, 1982

Matsuno (1982, Matsuno hereafter) developed the atmospheric transmission gravity-wave parameterisation, which treats gravity-wave propagation through the atmosphere in terms similar to ray optics. Matsuno assumed that viscosity, both eddy and molecular, is present everywhere in the atmosphere. The Matsuno parameterisation assumes that the eddy viscosity in the atmosphere is produced by gravity-wave dissipation and that this viscosity, along with molecular viscosity, causes further gravity-wave dissipation. The Matsuno parameterisation requires an initial diffusion coefficient to be prescribed at each height level.

Following Plumb and McEwan (1978), Matsuno calculated the momentum flux of the gravity waves in terms of the transmissivity τ of the atmosphere, given by,

$$\tau(\hat{c}_{phi}, \theta, z_1, z_2) = \exp \left[- \int_{z_2}^{z_1} \frac{2\nu_{total}(z')N^3}{k[c_{phi} - u(z')\cos\theta]^4} dz' \right] \quad \text{II.49}$$

where θ is the angle between the wave vector and the background wind, z' is the vertical coordinate between z_1 and z_2 , and ν_{total} is the total viscosity, given by

$\nu_{total} = \nu_{eddy} + \nu_{molecular}$. From equation (II.49) it can be seen that the transmissivity tends to zero as the viscosity of the atmosphere increases or as c_{phi} approaches zero.

Once τ has been calculated, the corresponding acceleration of the background flow can be found from,

$$X = - \frac{F(c_{phi}, \tilde{u})}{p} \frac{\partial \tau}{\partial z} \quad \text{II.50}$$

where $F(c_{phi}, \tilde{u})$ is the momentum flux at the source level and $\tilde{u} = (u'w'2N/kx)^{1/3}$, where $u'w'$ is the momentum flux of the gravity wave at level z' .

Meyer (1999a, b) used a hybrid of the Lindzen and Matsuno gravity-wave parameterisations, along with the Global Scale Wave Model (Hagan *et al.*, 1995) to study gravity wave interactions with planetary waves and the propagating diurnal tide respectively. Figure (II.4) shows how the Lindzen and Matsuno parameterisations are combined in this Meyer (1999a; Meyer hereafter) parameterisation. The implementation of this hybrid parameterisation in the CMAT model will be discussed in further detail in chapter (VI).

II.4.4 SEMI-EMPIRICAL, FRITTS AND LU, 1993, WARNER AND MCINTYRE, 2001

Both Fritts and Lu (1993) and Warner and McIntyre (2001) have developed gravity-wave parameterisations which use continuous spectra, rather than the discrete spectra of the Lindzen and Matsuno parameterisations, and which use observations, rather than theory, to constrain gravity-wave propagation and dissipation. These parameterisations have the advantage that they do not rely on any one gravity-wave dissipation theory. However, limitations in the observations of gravity waves means that these parameterisations do not necessarily represent gravity-wave dissipation in a more realistic manner than any other, and models using these parameterisations may not necessarily reproduce the observed features of the middle atmosphere as well as those which use parameterisations based on theoretical dissipation mechanisms (e.g. McLandress, 1998).

II.4.5 DOPPLER SPREAD, HINES, 1997A,B

Hines (1997a, b; Hines hereafter) argued that when the root mean square of the gravity wave induced winds, σ , becomes comparable to the horizontal phase velocity of the low end of the gravity-wave spectrum, nonlinear interactions resulting from horizontal advection act to Doppler shift the spectrum towards the high m tail. Hines used this process to approximate all nonlinear interactions, which is one limitation of the parameterisation (Hines 1996, 1999). This Doppler spreading is a conservative process, so wave breaking is imposed by removing all waves above a cut-off wavenumber m_M . This cut-off moves to lower wavenumbers with increasing altitude.

The associated momentum deposition is given as function of the changing spectrum $M_j(m_i)$. The vertical flux of horizontal momentum, F_j , is given by,

$$F_j = \rho_i \hat{\sigma}_{ji}^2 k_j \int_{m_m}^{m_j} M_j(m_i) \frac{dm_i}{m_i} \quad \text{II.51}$$

where the subscript j denotes different azimuth directions for each frequency i . From this the acceleration of the background flow can be found by,

$$X_j = -\frac{1}{\rho} \frac{\partial F_j}{\partial z}. \quad \text{II.52}$$

The Hines parameterisation also includes self-consistent calculations for eddy diffusion and local heating produced by gravity-wave dissipation. This parameterisation has been used successfully in a number of models (e.g. Akmaev, 2001a, b), and while it reproduces the observed mesopause zonal mean wind structure, it does not produce a power-law spectrum at high m values (Hines, 1993) and the m_M cut-off is somewhat arbitrary. Further, there are several other parameters used in the Hines parameterisation (referred to as fudge factors by Hines) which must be arbitrarily set. This leads to uncertainty in the calculated gravity-wave drag, heating and eddy diffusion (Hines 1997a, b).

II.4.6 NONLINEAR DIFFUSION, MEDVEDEV AND KLAASSEN 2000

Medvedev and Klaassen (1995, MK95 hereafter) extended the nonlinear diffusive spectral interaction theory of Weinstock (1982, 1990, 1993) to include nonlinear Doppler effects, and Medvedev and Klaassen (2000; MK2000 hereafter. Please note that while most of this parameterisation was first detailed by Medvedev *et al.*, 1997, MK2000 is the more widely recognised publication and will be used here in order to maintain consistency with other literature) used this to form the basis for their gravity-wave parameterisation. This theory assumes that as a wave propagates through the atmosphere, nonlinear (off-resonance) interactions with spectral components of similar or smaller vertical scales (higher m) and lower frequencies (lower ω) force the wave closer to its overturning state. Further, this background wind variance acts to create a nonlinear diffusion which causes gravity-wave dissipation. This can cause wave saturation, similar to the Lindzen theory, or cause wave obliteration. Obliteration occurs when the interactions are especially strong, such as when the wave approaches a critical level or its overturning level predicted by linear theory (this theory predicts that the wave often decays to zero before reaching these levels). Doppler spreading of the spectrum to higher values of m , which is key to the Hines parameterisation, was considered by MK95. However, Eckermann (1997) and Broutman *et al.* (1997) argued that the Hines theory overestimates the effect of Doppler spreading, and MK95 argued that only the portion of the spectrum with higher m and lower ω can Doppler spread any given wave component. Using this assumption, MK2000 showed that the effects of Doppler spreading were negligible and could be excluded from the parameterisation, which is almost entirely at odds with the Hines theory.

MK95 showed that, in the absence of a background wind shear, their theory self-consistently produces saturated spectra with the m^{-3} tail for high values of m and a near-linear dependence on m for lower values, in good agreement with observations.

MK95 showed that the evolution of the gravity-wave spectrum with height can be described by,

$$\frac{dS(m_R)}{dz} = \left(-\frac{\rho_{0z}}{\rho_0} + \frac{m_{Rz}}{m_R} - \beta \right) S(m_R) \quad \text{II.53}$$

where S is the power spectral density (PSD, defined as the squared modulus of the Fourier transform of the time series, or the amount of power per unit frequency as a function of frequency) of the horizontal wind associated with gravity-wave spectrum at a height z , m_R is the real part of the nonlinear Doppler shifted vertical wavenumber, and β is the coefficient of nonlinear damping due to interactions of the spectral component m_R with the rest of the spectrum. m_R is related to the linear vertical wavenumber m by,

$$m_R = m\sqrt{\pi}\alpha \exp(-\alpha^2) \operatorname{erfi}(\alpha) \quad \text{II.54}$$

where,

$$\operatorname{erfi}(\alpha) = \frac{2}{\sqrt{\pi}} \int_0^\alpha \exp(-x^2) dx \quad \text{II.55}$$

is the error function of an imaginary argument. Here α is a dimensionless parameter which MK2000 showed is proportional to the root of the effective Richardson number ($R_i = N^2 / (\partial u / \partial z)^2$) for the component m (and therefore to the stability of the wave), given by,

$$\alpha = \frac{N}{\sqrt{2}m\sigma} = \frac{c_{ph} - \bar{u}}{\sqrt{2}\sigma} \quad \text{II.56}$$

where σ^2 is the horizontal wind variance created by the components of the spectrum with vertical wavenumbers larger than m_R and is equal to,

$$\sigma^2 = \int_{m_R}^{\infty} S(m') dm'. \quad \text{II.57}$$

MK95 showed that the nonlinear damping coefficient is given by,

$$\beta = \sqrt{2\pi} N \sigma^{-1} \exp(-\alpha^2). \quad \text{II.58}$$

From equation (II.58) it can be seen that as α decreases, the damping rate increases dramatically.

From equations (II.52) to (II.58) and the mid-frequency dispersion relation, the evolution of the spectrum can be found. It is worth noting that the only required input to this set of equations, and therefore the only unknown in this parameterisation, is the source spectrum.

Once the evolution of the spectrum is known, the associated momentum deposition can be found from,

$$X(m_R) = -\frac{\beta(m_R)kS(m_R)}{m_R}. \quad \text{II.59}$$

MK2000 simplified the above equations for use in atmospheric models, using a discrete spectrum of gravity waves. They showed that the nonlinear Doppler shifting of the spectrum could be ignored and so the linear vertical wavenumber m can be used in place of m_R throughout (for full details of the implementation of this parameterisation, see MK2000). This theory has been further extended by Medvedev and Klaassen (2003) to include a self-consistent calculation of the thermal eddy diffusion coefficient (see Section III.3.5) and gravity wave heating of the background atmosphere.

MK2000 showed that in the case of only one gravity wave being present, their parameterisation acts in a similar manner to the Lindzen parameterisation, although it produces vertical drag profiles which are more smooth and it predicts that gravity-wave dissipation begins below the breaking level predicted by linear theory. A full spectrum version of the MK2000 parameterisation has been shown to produce realistic middle atmosphere circulation patterns in a number of global-scale models (e.g. Medvedev *et al.*, 1997; Medvedev *et al.*, 1998). The implementation of this parameterisation in the STEVE and STEVE-3D models will be discussed in Chapter (III) and in the CMAT model in Chapter (VI). These models will be used to investigate the interaction between gravity waves and the mean flow (Chapters IV to VI), planetary waves (Chapter V), and the propagating diurnal tide (Chapter VI).



Figure II.1, Gravity wave (larger) and instability (smaller) structures in noctilucent clouds observed over Turku, Finland, on 21-22 July 1989, after Fritts *et al.* (1993).

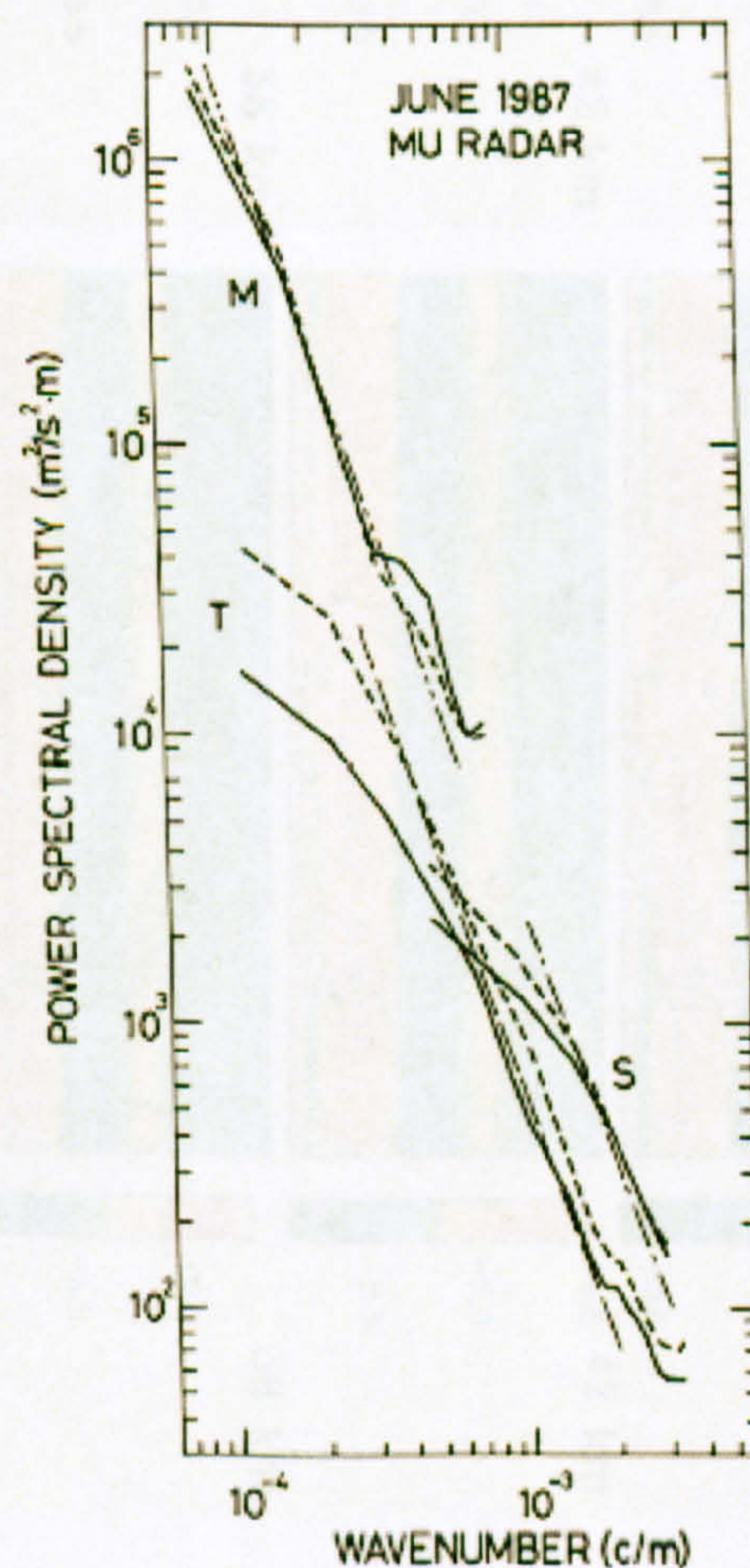


Figure II.2, Velocity spectra assumed to represent the horizontal velocity perturbations created by gravity waves in the troposphere (T), stratosphere (S) and mesosphere (M) and theoretical curves with amplitudes of $N^2/6m^3$, from Tsuda *et al.* (1989). Units are m^3/s^2 against cycles per metre.

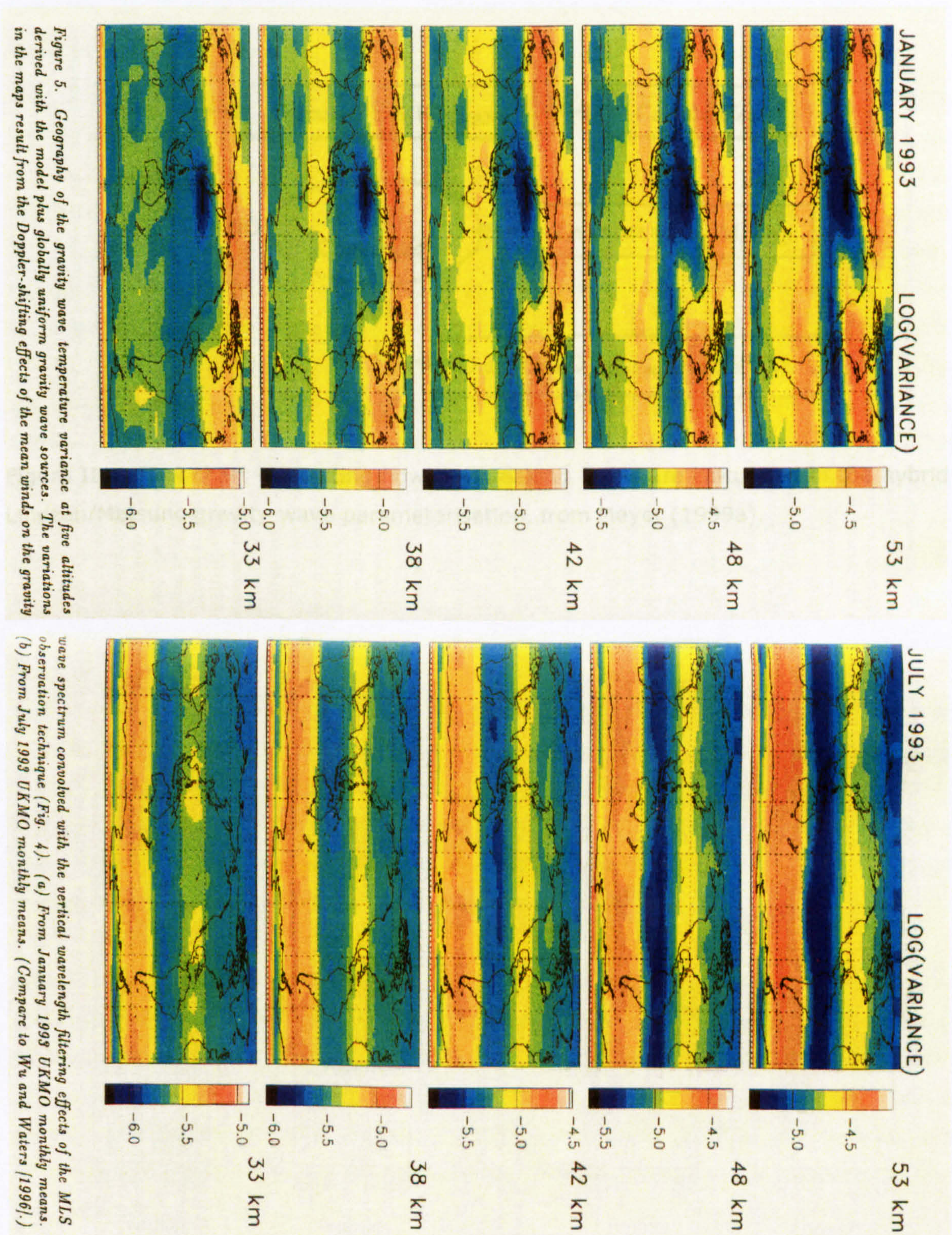


Figure II.3, Latitude / longitude distribution of the gravity wave temperature variances at 33, 38, 42, 48 and 53 km altitude derived by Alexander (1997) using observations from the MLS instrument during January and July 1993 and UKMO monthly mean temperatures.

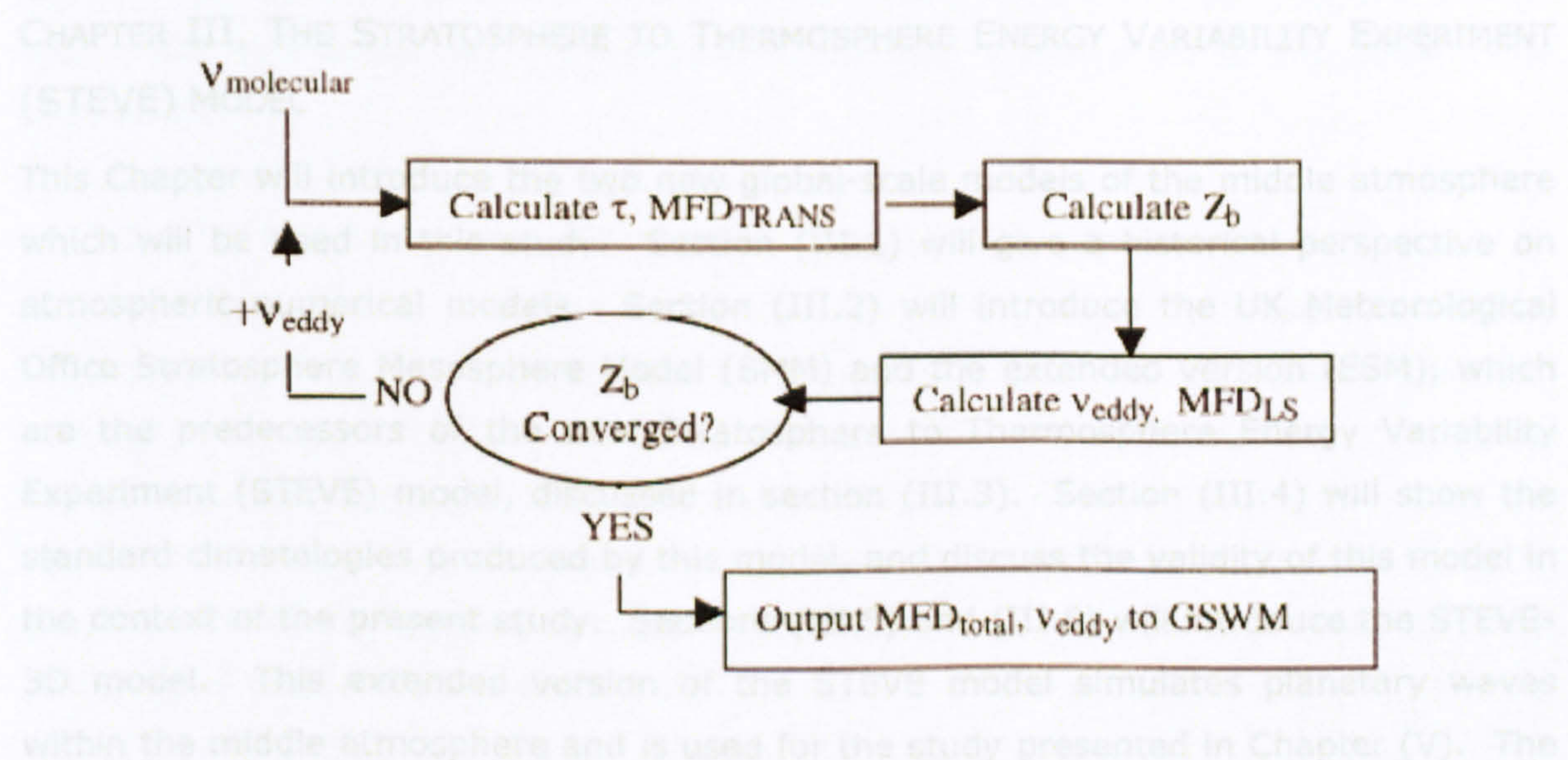


Figure II.4, Flow chart illustrating how the breaking levels are calculated in the hybrid Lindzen/Matsuno gravity-wave parameterisation, from Meyer (1999a).

III.1 GLOBAL-SCALE ATMOSPHERIC MODELS

Global-scale atmospheric models can be divided into two broad categories: weather models (e.g. Ghemey et al., 1950), which are used for operational forecasts; and climate models (e.g. Phillips, 1956), which are used to study fundamental atmospheric processes and to complement observations (e.g. Kalnay et al., 1996). Weather models clearly play an important role in daily life, but climate models have also had a large impact on policy making and public debate (e.g. IPCC, 2001).

This study will focus on the use of climate-type models to simulate the atmosphere. The first of these models was developed by Phillips (1956), who used a 2-layer model of the atmosphere to numerically solve the quasi-geostrophic equations of motion across a portion of one hemisphere. Further development of such models has led to results such as Manabe and Wetherald (1975), who were the first to use a climate model to predict the impacts of rising carbon dioxide levels on the troposphere and stratosphere.

Today, a hierarchy of climate models exist. These range from relatively simple models such as the STEVE model (see Section III.3), which can be used to perform long-term simulations, to ultra-high-resolution models (e.g. Satoh et al., 2004), which use some of the world's most-powerful computers to resolve small-scale features such as clouds and gravity waves. This study will use three different atmospheric models, of varying

CHAPTER III, THE STRATOSPHERE TO THERMOSPHERE ENERGY VARIABILITY EXPERIMENT (STEVE) MODEL

This Chapter will introduce the two new global-scale models of the middle atmosphere which will be used in this study. Section (III.1) will give a historical perspective on atmospheric numerical models. Section (III.2) will introduce the UK Meteorological Office Stratosphere Mesosphere Model (SMM) and the extended version (ESM), which are the predecessors of the new Stratosphere to Thermosphere Energy Variability Experiment (STEVE) model, discussed in section (III.3). Section (III.4) will show the standard climatologies produced by this model, and discuss the validity of this model in the context of the present study. Sections (III.5) and (III.6) will introduce the STEVE-3D model. This extended version of the STEVE model simulates planetary waves within the middle atmosphere and is used for the study presented in Chapter (V). The CMAT model, used in Chapter (VI), has been extensively documented by Harris (2001) and Harris *et al.* (2002), so only a relatively brief introduction to this model will be given in Chapter (VI).

III.1 GLOBAL-SCALE ATMOSPHERIC MODELS

Global-scale atmospheric models can be divided into two broad categories: weather models (e.g. Charney *et al.*, 1950), which are used for operational forecasts; and climate models (e.g. Phillips, 1956), which are used to study fundamental atmospheric processes and to compliment observations (e.g. Kalnay *et al.*, 1996). Weather models clearly play an important role in daily life, but climate models have also had a large impact on policy making and public debate (e.g. IPCC, 2001).

This study will focus on the use of climate-type models to simulate the atmosphere. The first of these models was developed by Phillips (1956), who used a 2-layer model of the atmosphere to numerically solve the quasi-geostrophic equations of motion across a portion of one hemisphere. Further development of such models has lead to results such as Manabe and Wetherald (1975), who were the first to use a climate model to predict the impacts of rising carbon dioxide levels on the troposphere and stratosphere.

Today, a hierarchy of climate models exist. These range from relatively simple models such as the STEVE model (see Section III.3), which can be used to perform long-term simulations, to ultra-high resolution models (e.g. Satoh *et al.*, 2004), which use some of the worlds most power computers to resolve small scale features such as clouds and gravity waves. This study will use three different atmospheric models, of varying

complexity, to investigate the interaction between unresolved atmospheric gravity waves and the dynamics, energetics and chemical composition of the background atmosphere.

III.2 THE UK METEOROLOGICAL OFFICE STRATOSPHERE MESOSPHERE MODEL

The SMM (e.g. Fisher, 1987) is a development on the original numerical model by Murgatroyd (1971). It is a global-scale mechanistic model of the stratosphere and mesosphere, which includes no explicit representation of the troposphere. The model solves the primitive equations on a log-pressure grid of 33 levels from 100 – 0.01 mb (~16 – 70 km at ~2 km spacing) with a horizontal resolution of 5° latitude by 5° longitude (it is worth noting that different variants of this model have used different grids, e.g. Arnold and Robinson, 1998).

The model solves the equations of motion using a fourth-order finite difference scheme in the horizontal plane, second-order in the vertical and a leap-frog time integration with a time step of 4 minutes. An eighth-order Shapiro (1970) filter is used to suppress nonlinear instabilities on scales of two grid lengths or less. The model uses a radiation scheme developed by Shine (1987), which calculates heating due to UV absorption by ozone and atomic oxygen (modified from Strobel, 1978, using ozone data from Klenk *et al.*, 1983, and Keating and Young, 1985) and infra-red radiative transfer matrix inversion calculation to account for the absorption and emission from carbon dioxide, ozone and water vapour. A Rayleigh friction approximation to gravity-wave drag, with a friction coefficient based on Holton (1976) is used in the model. Later versions, such as that used by Austin and Butchart (1992), also include interactive ozone chemistry and minor species transport.

Arnold and Robinson (1998) adapted the SSM into the Extended UK Meteorological Office Stratosphere Mesosphere Model (ESM), by extending its altitude range up to ~140 km. This version of the model used a shorter time-step of 90 s, to provide stability in the thermosphere region where wind speeds and diffusion rates can become very large. In the extended region of the model, the effects of molecular diffusion were included, and the radiation scheme used a combination of Strobel (1978) and Newtonian cooling to a background temperature set by MSIS-E90. For simplicity, the interactive chemistry and tracer transport were removed from the model.

The SSM and ESM models have been well documented and used for a variety of studies. O'Neill and Pope (1988) used the SMM to study dynamical effects around the time of a stratospheric sudden warming (see Chapter V for further discussion), Austin and Butchart (1992) used the SSM to investigate ozone chemistry in the polar region

and Arnold and Robinson (1998, 2000, 2001) have used the ESM to study the effects of solar variability on the middle atmosphere (see Chapter IV for more details). The following section will introduce the new STEVE model, which is a further development of the ESM / SMM family of models.

III.3 THE STEVE MODEL

The STEVE model was developed to study the possible effects of variable solar UV flux with the solar cycle on gravity wave's in the mesosphere and lower thermosphere. In order to simulate a complete 11 year solar cycle, or several successive cycles (such as in Chapter IV), the model needs to be computationally fast. This places some limitations on the complexity of the model. The first restriction is that the model is 2 dimensional (zonally averaged), which means that it is not able to resolve planetary waves or tides. There is no interactive chemistry within the model, which greatly simplifies the computation, but places limitations on the effects of variable UV flux which can be simulated. In addition, despite the model extending to thermospheric altitudes, no coupling between the ionosphere and thermosphere is included.

In order to simulate the effects of gravity wave activity more self-consistently, the Rayleigh friction approximation used in the ESM has been replaced with an implementation of the MK2000 gravity-wave parameterisation. The model extends from the lower stratosphere to the middle thermosphere, which ensures that the numerical effects of the boundaries have as little impact on the mesosphere and lower thermosphere (which is the focus of the study in Chapter IV) as possible. The model also includes a recently revised atmospheric radiation scheme which is necessary for simulating the effects of the solar cycle in UV on the middle atmosphere (Dobbin, 2004; private communication).

Figure (III.1) shows an overview of the components of the STEVE model, which will be described in detail in the following sections.

III.3.1 THE MODEL GRID

The STEVE model uses the numerical scheme from the SMM. The primitive equations are solved on a spatial grid of 65 vertical log pressure levels from 100 mb to 10^{-6} mb ($\sim 16 - 200$ km) at one quarter pressure scale height resolution, 36 latitudinal points at 5 degree spacing (87.5°S to 87.5°N) and is zonally averaged. Figure (III.2) shows the pressure at each model level and figure (III.3) shows the approximate altitude of each model level at 45°N during equinox at solar minimum ($F_{10.7}$, 10.7 cm solar flux = $70 \times 10^{-22} \text{ Js}^{-1}\text{m}^{-2}\text{Hz}^{-1}$; A_p geomagnetic index = 1). The scale height in the middle

atmosphere is often approximated to be constant (~ 7 km), which is represented by the region of approximately constant slope in figure (III.3). In the lower thermosphere, the rapid increase in mean temperature with height leads to a rapid increase in the scale height and decrease in the slope in figure (III.3).

There is zero vertical flux at the upper boundary of the model, which is a reasonable approximation as the vertical temperature gradient tends to zero in the middle and upper thermosphere. For calculations involving thermal gradients, the temperatures above the upper boundary are given by the MSIS-E90 empirical model (Hedin, 1991). The temperature and wind fields are linearly extrapolated when calculations require values below the bottom boundary of the model. The geopotential height of the bottom boundary is defined by the assimilated Stratospheric Sounding Unit data (Miller *et al.*, 1980), using data from 1993. Figure (III.4) shows the zonal mean geopotential height of the 100 mb pressure level at equinox and solstice. As can be seen, both latitude and season affect the geopotential height of the bottom boundary of the STEVE model. In agreement with equation (I.18), where the temperatures of the lower atmosphere are high, the geopotential height of the 100 mb pressure level is elevated compared to locations where the temperature of the lower atmosphere is low. The effect of the latitudinal gradient of the geopotential height of the bottom level of the model will be discussed more in Section (III.4).

The STEVE model calculates the temperature and wind fields every 90 s, which is suitable for the resolution of the model and the numerical scheme used. The inputs to the equations of motion are updated at different time intervals. The radiative heating and cooling is updated daily, the gravity-wave drag and eddy diffusion is calculated every 900 s and the vertical diffusion updates every 90 s. These intervals provide a balance between the computational efficiency required for the model with numerical stability of the model.

III.3.2 ULTRA VIOLET ABSORPTION

The heating rates due to the absorption of UV by O_3 in the Chappuis, Huggins and Hartley bands, O_2 in the Schumann-Runge band, and both O_3 and O_2 in the Herzberg continuum, are calculated using a parameterisation developed from Strobel (1978). The parameterisation used in the model has been adapted from Strobel (1978) by Apruzese *et al.* (1982) and Harris (2001). This modified parameterisation uses heating efficiencies from Mlynczak & Solomon (1993) to account for airglow and Strobel (1978) to account for photo-dissociation of O_2 by absorption in the Schumann-Runge band. This has been further revised for the STEVE model to include variations in each solar

UV frequency band as a function of solar activity, using data from the UARS SOLSTICE instrument (Rottman, 1985).

Heating produced by absorption in the Chappuis band is given by

$$\frac{Q_c}{[O_3]} = 1.05 \times 10^{-15} \exp(-2.85 \times 10^{-21} C_{O_3}), \quad \text{III.1}$$

where Q_c is the heating rate in 0.1 Wm^{-3} , $[O_3]$ is the local number density of O_3 in cm^{-3} and C_{O_3} is the total slant column density of O_3 . The heating rate caused by absorption in the Huggins band is given by

$$\frac{Q_{HC}}{[O_3]} = \frac{1}{C_{O_3}} [4.66 \times 10^3 - 7.8 \times 10^2 \exp(-1.77 \times 10^{-19} C_{O_3}) - 3.88 \times 10^3 \exp(-4.22 \times 10^{-18} C_{O_3})], \quad \text{III.2}$$

where Q_{HC} is the heating rate in 0.1 Wm^{-3} . For the Hartley band, the heating rate is calculated using the fluxes and absorption cross-sections of Strobel (1978). In the Herzberg continuum, the diabatic heating rate is given by,

$$Q_{HERZ} = 1.5 \times 10^7 (6.6 \times 10^{-24} O_2 + 4.9 \times 10^{-18} O_3) \times \exp(-6.6 \times 10^{-24} C_{O_2} - 4.9 \times 10^{-18} C_{O_3}), \quad \text{III.3}$$

where Q_{HERZ} is the heating rate in Wm^{-3} and O_2 is the local number density of O_2 .

The volume mixing ratio of O_2 used in the STEVE model is obtained from the MSIS-E90 empirical model. Figure (III.5) shows the zonal mean O_2 volume mixing ratio during June solstice at solar minimum. Below 100 km, O_2 forms approximately 20 % of the atmosphere by volume, but above the turbopause this value falls with height, while other species, such as O increase as discussed in Chapter (I). The O_3 climatology used in the STEVE model is based on the climatology from the Solar Backscattered Ultra-Violet instrument on Nimbus 7 (Klenk *et al.*, 1983) which was extended using data from Keating & Young (1985) as detailed in Shine (1987). Figure (III.6) shows the O_3 volume mixing ratio used by the STEVE model during June solstice at solar minimum. In Section (III.4) the climatologies of the STEVE model are compared with, amongst others, those of the UARS Reference Atmosphere Project (Swinbank and Ortland, 2003). It is therefore important to show that the zonal mean O_3 climatology used as an input to the STEVE model is in good agreement with that of the UARS Reference Atmosphere Project, shown in figure (III.7). The peak in O_3 concentration observed in the equatorial stratosphere is linked to the peak in photochemical production in this region. There is also a significant O_3 concentration in the polar night in the stratosphere, which is transported into this region by the Brewer-Dobson circulation. The total heating produced by absorption of UV by O_2 and O_3 in the STEVE model is

shown in Section (III.4). The local minimum in O_3 concentrations, created by a peak in OH density in the upper mesosphere (associated with the photo-dissociation of H_2O , e.g. Garcia and Solomon, 1985) is also seen within the O_3 climatology used in the STEVE model.

III.3.3 INFRA RED ABSORPTION AND EMISSION

Radiative heating and cooling in the LTE and non-LTE regions by $15.6\ \mu m$ absorption and emission by CO_2 is calculated using the radiative transfer Curtis matrix inversion parameterisation of Fomichev *et al.* (1998). The heating and cooling from the $9.6\ \mu m$ band of O_3 is calculated using the parameterisation of Fomichev and Shved (1985). In the LTE region, this parameterisation uses matrices of transmission calculated from the line spectra of CO_2 and O_3 for several volume mixing ratios. The parameterisation interpolates between these matrices to calculate the transmittance between two atmospheric layers and the local radiative cooling this produces. In the non-LTE region, the effects of radiative absorption between energy levels, spontaneous and induced emission and deactivation through particle collisions are taken into account in the calculations.

Figure (III.8) shows the global mean CO_2 concentrations used in the STEVE model (after Fomichev *et al.*, 1998). As can be seen, CO_2 is well-mixed below ~ 75 km. Above this altitude, a combination of photodissociation and its relatively high molecular weight (leading to a relatively short scale-height) result in its concentration falling with altitude.

The infra-red radiative heating and cooling parameterisation of Fomichev *et al.* (1998) is only valid below 150 km. Above this level, the radiative cooling rates in the STEVE model are assumed to decrease exponentially with height, which is consistent with the decrease in efficiency of radiative cooling above 150 km discussed in Chapter (I). Temperatures below 100 mb are required for the calculation of infra-red fluxes between the bottom level of the model and the lower atmosphere. To prevent any sharp gradients between the temperatures at the 100 mb and those in the levels below, the temperatures below 100 mb are linearly extrapolated from the bottom 2 levels of the model. This approximation has implications for the temperatures and cooling rates at the bottom boundary of the model which will be discussed in Section (III.4). The infra-red heating and cooling calculated by the STEVE model, and the net diabatic heating rates are presented in Section (III.4).

III.3.4 GRAVITY-WAVE DRAG

Acceleration due to atmospheric gravity waves is calculated using the parameterisation of MK2000, which has been discussed in detail in Chapter (II). An isotropic spectrum of 60 waves is used, with 15 frequencies in each of the four cardinal directions, which is sufficient to produce a realistic gravity-wave drag profile, while being computationally fast (Medvedev *et al.*, 1997). A modified Desaubies source spectrum (Fritts and VanZandt, 1993) is used, as in Medvedev *et al.* (1997). The power spectral density of each source wave, $S(m)$, is given by,

$$S(m) = S_0 \frac{(m/m_*)^s}{1 + (m/m_*)^{s+t}}, \quad \text{III.4}$$

where S_0 is the amplitude of the spectrum, m is the vertical wavenumber of each wave, m_* is the characteristic wavenumber ('knee' of the spectrum, discussed in Section II.3.2), $s = 1$ for a modified Desaubies spectrum, and $t = 3$, which defines the ' m^{-3} ' power law of the high wavenumber tail of the spectrum (e.g Fritts and VanZandt, 1993). Following Medvedev & Klaassen (2000), $m_* = 0.006 \text{ m}^{-1}$ and m values in the range $2\pi/900 \text{ m}^{-1}$ to $2\pi/19 \text{ km}^{-1}$ are used, with a characteristic horizontal wavenumber $k_h = -2\pi/300 \text{ km}^{-1}$ (Allen and Vincent, 1995; Fritts and Chou, 1987; Tsuda *et al.*, 1991). Following Medvedev *et al.* (1998), a latitudinally varying source amplitude S_0 is used which peaks at mid-latitudes. Medvedev *et al.* (1998) argued that this was necessary to account for the strong gravity wave source associated with the tropospheric jet streams. A mean value of $S_0 = 100 \text{ m}^3 \text{ s}^{-2}$ is used, as in Medvedev *et al.* (1997). No time variation is included in the source spectrum, which is discussed in Chapter (IV). Figure (III.9) shows the power spectral density as a function of horizontal phase speed for the source spectrum in STEVE at the equator, compared with those used in Medvedev and Klaassen (2000). A value of the Brunt-Väisälä frequency $N = 2 \pi/8$ minutes and a background wind speed of 0 ms^{-1} have been assumed. As can be seen, the two spectra have the same overall shape, with the ' m^{-3} ' tail at high wavenumbers clearly visible. The differences in the spectrum arise from the different source amplitudes used by Medvedev and Klaassen (2000) and Medvedev *et al.* (1997).

The MK2000 parameterisation does not take into account the damping of gravity waves by ion-neutral collisions, which becomes an important source of damping in the thermosphere. In order to prevent them from growing to unrealistic amplitudes in the upper levels of the STEVE model, the gravity wave amplitudes are decreased

exponentially with height above 120 km, with a scale height equal to the pressure scale height of the atmosphere.

A small Rayleigh friction term is used to simulate the effects of planetary waves which are not represented in the STEVE model. This is needed to reduce the zonal wind strength in the stratosphere, which is important for the overall dynamics of the middle atmosphere and the filtering of the gravity wave spectrum. This drag also provides damping in the upper levels of the model, where the parameterised waves are heavily damped as described above. The Rayleigh friction coefficient is plotted as a function of height in figure (III.10).

The STEVE model can calculate gravity-wave drag using either the MK2000 scheme and weak Rayleigh friction described above, or with just a stronger Rayleigh friction term (three times the profile in figure III.10). The differences between these two will be discussed in more detail in Chapter (IV).

III.3.5 DIFFUSION

As discussed in Chapter (I), molecular diffusion plays an important role in maintaining the thermospheric temperature structure. The vertical diffusion of O₂, N₂ and O is included within the STEVE model, using the data and calculation of Rees (1989; see Table I.1 and equation I.11). There is an additional vertical diffusion component due to gravity-wave dissipation which is included within the STEVE model. This is either calculated from the drags produced by the MK2000 parameterisation and the eddy diffusion coefficient calculation of Medvedev and Klaassen (2003), or are prescribed when the STEVE model is run using only Rayleigh friction. The calculation presented by Medvedev and Klaassen (2003) has the form,

$$D_{zz} = \sum_j (2m_j N)^{-1} a_{hj} , \quad \text{III.5}$$

where D_{zz} is the vertical eddy diffusion coefficient. Figure (III.11) shows the prescribed globally averaged eddy diffusion coefficient used in the model, which is based on the climatology calculated by Garcia and Solomon (1985). The vertical eddy diffusion coefficients calculated by the STEVE model are presented in Section (III.4).

Horizontal diffusion caused by gravity-wave dissipation and molecular diffusion is not calculated as the large horizontal grid spacing (~500 km) makes this insignificant compared with the vertical diffusion (grid spacing ~2 km). However, as there are no planetary waves in the STEVE model, a horizontal diffusion term is included to simulate the horizontal mixing caused by breaking or diverging planetary waves. This mixing is

required to maintain the temperature structure in the winter polar stratosphere and the zonal jets in the middle atmosphere.

III.4 THE STEVE MODEL CLIMATOLOGIES

The STEVE model calculates a number of zonally averaged global fields, some of which can be directly compared to observations, such as temperatures and winds, and some of which cannot be directly measured, such as the heating rate due to UV absorption by O_3 . This section will highlight the key model outputs for both solstice and equinox conditions and compare them with observations and other model results.

III.4.1 ZONAL MEAN TEMPERATURES

Figure (III.12) shows the zonal mean temperatures for the March equinox at solar minimum calculated by the STEVE model and figure (III.13) shows the temperatures from the MSIS-E90 empirical model for the same conditions. The temperatures calculated by the STEVE model are in overall agreement with those of the MSIS-E90 model during equinox, although there are several key differences. Below 25 km altitude the STEVE model overestimates the temperatures, especially in the equatorial region, where a local minimum is seen in the MSIS plot. This local minimum, around the equatorial tropopause is a well-known phenomena, which is created by the adiabatic expansion of air in the tropospheric Hadley circulation. As the effects of the troposphere are not included in the STEVE model, it is not surprising that this feature is not replicated. In addition, as shall be shown in section (III.4.4), the extrapolation of temperatures below the bottom boundary of the STEVE model leads to an incorrect estimate of the IR cooling in the lower stratosphere, which is another cause of the relatively high temperatures seen in this region. The other main area of significant differences in the estimates of the temperatures is in the thermosphere, above 130 km, where the STEVE model overestimates the temperatures by tens of K. This is likely to be due to the lack of IR cooling due to NO and $O(^3P)$ in the model, which has been shown to be a significant source of cooling in this region (e.g. Roble, 1995) and the inadequacy of the UV heating scheme used in the STEVE model at this altitude (see Section III.4.4).

Figure (III.14) shows the zonal mean temperature structure during the June solstice at solar minimum calculated by the STEVE model and figure (III.15) shows the temperatures from the MSIS-E90 model for the same conditions. The agreement between the two is not as good as during equinox, although many of the key features of the MSIS-E90 plot are seen in the STEVE model. Importantly, the STEVE model is able to reproduce local temperature minimum around the summer mesopause, which

indicates that the gravity-wave parameterisation in the model is driving a realistic meridional circulation in this region. However, the summer mesopause in the STEVE model is too warm and the winter mesopause too cold. In addition, the winter polar stratosphere and lower thermosphere temperatures are too high and low respectively. These are both due to an unrealistically high horizontal diffusion which exists in the lower half of the model (discussed in Section III.3.5), which diffuses heat into and out of the polar regions. This horizontal diffusion is required to maintain a realistic wind structure in the stratosphere and lower mesosphere in the absence of planetary waves, which create an amount of horizontal mixing when they break.

Figure (III.16) shows the annual cycle in temperatures calculated by the STEVE model during solar minimum conditions at 67.5 ° North and figure (III.17) shows the annual cycle in MSIS-E90 for the same conditions. The transition between equinox and solstice conditions can be seen in both models. The STEVE model is able produce most of the main features of the seasonal cycle in temperatures seen in the MSIS-E90 model, but significantly less seasonal variability is seen in the STEVE model between approximately 80 and 100 km. This is an effect of the artificially high horizontal diffusion used in the model, which is used to compensate for the lack of planetary wave activity.

III.4.2 ZONAL MEAN WIND STRUCTURE

Figures (III.18 to III.20) show the zonal mean zonal wind structure calculated by the STEVE model, from the HWM empirical model and from the extended URAP during the March equinox at solar minimum. During equinox the zonal mean zonal winds produced by the STEVE model show some of the features of the observed winds, but there are several key discrepancies. At mid and high latitudes in the middle atmosphere, the eastward jets seen in the HWM and URAP climatologies, which are driven by the horizontal temperature and pressure gradients are present in the STEVE model, although they are generally stronger than those in the HWM and URAP data, which is a feature of many atmospheric models. In the case of the STEVE model, the lack of planetary wave dissipation in the model is a likely cause for this discrepancy (planetary wave dissipation acts to slow the zonal mean winds and reduces the temperature gradient between the pole and lower latitudes). At lower latitudes in the middle atmosphere, weaker eastward winds are seen in the STEVE model, similar to those in the HWM empirical model and URAP climatology, although the URAP climatology shows stronger winds at low than mid and high latitudes. In the stratosphere, a weak westward flow is seen in the HWM and URAP climatology, which is largely driven by momentum deposition from the equatorial diurnal tide, which will

be discussed in more detail in Chapter (VI). As the STEVE model does not contain any representation of the diurnal tide, it is perhaps not surprising that this feature is absent in the model results. In the mesopause and lower thermosphere, a reversal of the eastward jets is seen in the HWM and URAP data, although this extends beyond the limit of the URAP climatology. This reversal is driven by both gravity wave and tidal dissipation. This reversal is seen in the STEVE model, which agrees well with the HWM data, but is lower in magnitude than that seen in the equatorial region in the URAP climatology. Above around 120 km, the zonal mean wind structure in the STEVE model is dominated by the strong Rayleigh friction present in the model at these altitudes and the model does not reproduce the HWM data. At these altitudes, ion-neutral collisions play a dominant role in the momentum budget of the thermosphere (e.g. Roble, 1977) and these effects are not included in the STEVE model. In addition, there is known to be a persistent, non-growing numerical instability in the STEVE model associated with the upper polar boundary of the model which produces winds that are unrealistic (see figure III.18). The studies in Chapters (IV) and (V) will only deal with model results from below 110 km to avoid using the data from this region.

Figures (III.21 to III.23) show the zonal mean zonal wind structure from the STEVE model, the HWM empirical model and from the extended URAP during the June solstice at solar minimum. The STEVE model is in better overall agreement with the HWM and URAP zonal winds at solstice than equinox, but some differences remain. The middle atmosphere jets are in good overall agreement with the HWM and URAP winds, with reversals seen in the lower thermosphere in both the HWM and STEVE models, although the URAP climatology shows a stronger jet reversal, especially in the summer hemisphere, than either of the STEVE or HWM. Above around 120 km the winds in the STEVE model are heavily damped as discussed above. At the equator a numerical instability is seen in the STEVE model. This is related to the advection scheme used within the model, which becomes unstable when strong horizontal gradients are present, although this instability does not grow any stronger than seen in figure (III.21).

It is worth noting that while the HWM empirical model is based on observed horizontal winds, recent climatologies (e.g. Portnyagin *et al.*, 2004) have shown that this model does not work well in the high latitude MLT region. Figure (III.53) shows monthly mean annual cycle in the zonal winds measured by a number of MF and meteor radars at Arctic latitudes (after Portnyagin *et al.*, 2004). In all five of the data sets show that the jet reversal in the summer time is at least 10 km lower than that seen during winter. This feature is not reproduced by the STEVE model, or the HWM empirical model. As the CMAM model has been shown to produce a zonal wind structure which

agrees with the HWM and STEVE models, and does not show the seasonal asymmetry in the jet reversal heights when the MK2000 scheme is used (Medvedev and Klaassen, 2003), a possible short-falling in the gravity-wave parameterisation seems to be the likely cause. Despite the apparent inadequacy of the HWM in the high latitude MLT region, this model does provide horizontal winds at all latitudes and heights simulated by the STEVE model and as such will still be used as a reference in later chapters of this thesis.

Figure (III.24) shows the zonal mean meridional wind structure during the March equinox at solar minimum calculated by the STEVE model and figure (III.25) shows the meridional winds from the HWM. As meridional thermal winds are driven by longitudinal gradients, created by diurnal heating and large scale waves, both of which are absent in the STEVE model, it is not surprising that there is little agreement between the meridional winds in the STEVE model and the HWM and therefore no further comparison will be made here. In the zonally averaged case, all meridional winds are driven by frictional forces, so perhaps a better comparison can be made between the meridional circulation driven by those forces (see Section I.3.3) and the Murgatroyd and Singleton circulation which is inferred from observations. The meridional circulation within the STEVE model will be shown in the following section. The zonal mean meridional and vertical winds, which are needed to calculate this circulation, are shown in figures (III.24, III.26, III.27 and III.28). The meridional winds have typical values of $10 - 20 \text{ ms}^{-1}$ around the mesopause and the vertical winds have typical values of $10 - 20 \text{ mms}^{-1}$ in the polar mesosphere, both of which are consistent with the residual circulation calculated in other atmospheric models (e.g. Rosenfield *et al.*, 1987; Shine, 1989; Fetzer and Gille, 1996) and observations (e.g. Solomon *et al.*, 1986).

III.4.3 ZONAL MEAN MERIDIONAL CIRCULATION

As discussed in sections (I.3.4 and III.4.3), gravity-wave dissipation drives a meridional circulation in the mesosphere (Murgatroyd and Singleton circulation, shown in figure I.12), which is responsible for the mesopause temperature anomaly. This circulation can be seen within the STEVE model by using the meridional and vertical winds described above to plot the zonal mean, time-averaged meridional circulation. Figure (III.29) shows the meridional circulation for the March equinox averaged over one day and figure (III.30) shows the meridional circulation for the June solstice. At equinox, the circulation pattern is seen to rise in the equatorial middle atmosphere, move poleward and descend at the pole, while at solstice, the circulation pattern rises in the summer polar region, travels to the winter polar region and descends, in

agreement with the Murgatroyd and Singleton circulation. The ability of the STEVE model to reproduce the shape, sense and approximate strength of this meridional circulation is important as perturbations on this circulation form a major part of the results in Chapters (IV and V). Further, this shows that the zonal mean, time averaged gravity wave forcing calculated by the STEVE model (Section III.4.5) is of the correct form.

III.4.4 DIABATIC HEATING RATES

As discussed in section (I.3.4), the residual circulation drives the middle atmosphere away from radiative equilibrium. This means that there is a net diabatic heating rate in localised regions, where the radiative and chemical heating and cooling processes do not balance. This section will present the local heating and cooling rates due to UV and IR absorption and emission in the STEVE model and the net diabatic heating rate. As these parameters are not currently observable, the STEVE model results will be compared to other global-scale model results.

Figure (III.31) shows the zonal mean atmospheric heating rate due to UV absorption calculated by the STEVE model for the March equinox at solar minimum. The greatest heating due to absorption of solar UV occurs in the thermosphere and is responsible for the high temperatures in this region. The peak in heating rate around the stratopause is associated with the high ozone concentration in this region and is responsible for creating the local temperature maximum at the stratopause seen in figure (III.12). The peak heating rate occurs at the equator for all altitudes, which corresponds to the solar zenith angle of 0 ° during equinox. However, while the solar radiation incident on the top level of the model is symmetrical about the equator during equinox, the heating rates due to UV absorption are not perfectly symmetrical because the chemical concentrations used by the STEVE model are based on the observations which are not perfectly symmetrical about equator (e.g. figure III.6).

Figure (III.32) shows the zonal mean atmospheric heating rate due to UV absorption calculated by the STEVE model for the June solstice at solar minimum. Here the peak in heating due to UV absorption occurs in the summer hemisphere at mid to high latitudes, with maxima around the stratopause and in the thermosphere. Figure (III.33) shows the zonal mean heating rate due to absorption of solar UV calculated by the Extended Canadian Middle Atmosphere Model (CMAM) for the June solstice at solar minimum (Fomichev *et al.*, 2002), which extends from the ground to 2×10^{-7} mbar (~ 210 km) and has been shown to reproduce many of the observed features of atmosphere. The two models show a good overall agreement in their estimates of the

heating rate due to UV absorption, with peaks of around 15 Kday^{-1} at the high latitude stratopause and several hundred in the thermosphere. However, the CMAM model predicts the peak in heating at the stratopause to occur in the summer polar region, whereas the STEVE model predicts this peak to be located around 50°N . As this peak is centred close to the peak in the ozone layer it seems likely that this discrepancy is due to differences in the O_3 profiles used by the two models. In the thermosphere, the efficiency of heating by absorption of EUV is highly dependent on the chemistry of the thermosphere, and as neither model contains interactive chemistry in this region, it is expected that neither estimate will be accurate in this region.

Figure (III.34) shows the zonal mean net atmospheric heating rate due to IR absorption and emission calculated by the STEVE model for the March equinox at solar minimum. Three maxima are seen, in the equatorial stratopause region, in the polar mesosphere and in the lower thermosphere, each corresponding to a high temperature region (figure III.12). The exponential decay in the cooling rates above 150 km, discussed in section (III.3.3), is clearly visible.

Figure (III.35) shows the zonal mean atmospheric heating rate due to IR absorption and emission calculated by the STEVE model for the June solstice at solar minimum. Again the regions of high IR cooling correspond to the regions of high temperature (figure II.14) and again the exponential decay above 150 km is clearly visible. Unlike solar UV absorption, IR emission is dependant on the temperature of the background atmosphere, so comparing this result with the equivalent CMAM result (figure III.36) is a more stringent test of the STEVE model. There is good overall agreement between the two model results in the middle atmosphere, with maxima in the IR cooling in the summer stratopause (around 10 Kday^{-1}), in the winter mesosphere (around 15 Kday^{-1}) and a net IR absorption in the summer mesopause region. However, the IR cooling in the thermosphere, above 120 km, is much larger in the STEVE model than the CMAM model. The temperatures calculated by both models (figure III.14 and Fomichev *et al.*, 2002) are comparable in this region, so these differences are likely caused by the differences between the IR cooling schemes used by the two models.

Comparing figure (III.31) with figure (III.34) and figure (III.32) with figure (III.35), it can be seen that the regions of high IR emission tend to coincide with the regions of high UV absorption, although there are several differences which lead to the net diabatic heating rate, shown in figures (III.37) and (III.38). Above 150 km, a strong net heating is present at all latitudes which are in sunlight. This net radiative heating is balanced by the efficient molecular diffusion in this region (see Section I.2.7). Below this level, the net diabatic heating is maintained by the meridional circulation shown in

the previous section. In regions of rising air, such as the summer mesosphere during solstice or the equatorial mesosphere during equinox, there is a net adiabatic cooling through expansion of the air. This reduces the background temperature to below its radiative equilibrium value, which reduces the local IR cooling rate and produces a net diabatic heating in these regions. In regions of descending air, a net diabatic cooling is seen because the background temperature is raised above the radiative equilibrium temperature through adiabatic heating (via compression), which in turn increases the local IR cooling rate.

The agreement between the circulation patterns (Section III.4.3) and the diabatic heating rates shows that the STEVE model is responding correctly to the gravity wave driven circulation in the middle atmosphere. The following section will show the gravity-wave drags calculated by the STEVE model and the associated horizontal wind variances and vertical eddy diffusion coefficients which these waves produce.

III.4.5 GRAVITY-WAVE DRAG, WIND VARIANCES AND EDDY DIFFUSION

The MK2000 gravity-wave scheme, described in detail in the previous Chapter, produces estimates of the gravity-wave drag, the horizontal wind variance produced by gravity waves oscillating the background atmosphere and the vertical eddy diffusion coefficient associated with the mixing produced by dissipating gravity waves. All three of these quantities can be measured, or inferred from ground based instruments, but only for localised regions. The outputs produced by the STEVE model presented here will be compared with other global-scale model results and, where possible, with observed values. Only the zonal wave drag and wind variance results will be presented because the meridional winds in the STEVE model do not compare well with observations of the wind at any one location (see Section III.4.2) and therefore there is expected to be poor agreement between the meridional wave drag and wind variances and observations.

Figure (III.39) shows the zonal mean zonal gravity-wave drag calculated by the STEVE model for the March equinox during solar minimum. In the stratosphere and lower mesosphere the acceleration due to dissipating gravity waves is small because of the relatively high atmospheric density in this region. The peak in gravity-wave drag from the parameterised waves occurs in the upper mesosphere and lower thermosphere, where their effect on the background winds and meridional transport is greatest. The gravity-wave drag at mid to high latitudes is predominantly westward, opposing the eastward wind in this region (figure III.17). In the equatorial region a weak westward

drag is seen, which is responsible for creating the westward flow seen in this region. Above 150 km the strong Rayleigh friction is seen.

Figure (III.40) shows the zonal mean zonal gravity-wave drag calculated by the STEVE model for the June solstice during solar minimum. The parameterised gravity-wave drag in the winter hemisphere is eastwards and stronger than during equinox, with two maxima, one at 75 km and one at 100 km altitude (the seasonal cycle in gravity wave activity will be discussed in Chapter IV). This feature is a common feature of other models which implement this parameterisation (e.g. Medvedev and Klaassen, 2003), although it is not seen in observations. The MK2000 scheme uses only 15 frequencies of gravity waves in its source spectrum, so it is likely that only a few discrete waves reach this altitude region. The absorption or reflection of just one or two waves may be responsible producing the local minimum around 85 km altitude. The effects of this feature are also seen in other quantities calculated by the gravity wave scheme, which are discussed below.

In the summer hemisphere, the westward zonal wind (figure III.21) filtering of the gravity-wave spectrum gives rise to an eastward drag, maximising around 100 km altitude. Figure (III.41) shows the gravity-wave drag calculated by Fomichev *et al.* (2002) using the extended CMAM model and the Hines (1997a, b) parameterization and figure (III.42) shows the gravity-wave drag calculated by Medvedev and Klaassen (2003) using the CMAM model and the MK2000 parameterization. While all three models produce zonal wind and temperature climatologies which are in reasonable agreement with observations (excluding the effect discussed above), each produces a different estimate of the zonal mean gravity-wave drag. These differences can be attributed to differences in the dissipation mechanisms used in the different parameterisations, different source spectra and different background wind fields. Despite these difference, all three model results share several common features. The peak in gravity-wave drag in each model is around the mid-latitude mesopause region, with stronger eastward drag in the summer hemisphere than the westward drag in the winter. The two models using the MK2000 parameterisation show a double peak in the gravity-wave drag in the winter hemisphere, with the higher altitude peak being smaller than the lower altitude peak.

Gravity wave momentum deposition can also be estimated from ground based radar observations of gravity wave momentum fluxes. In the lower stratosphere the Middle and Upper Atmosphere (MU) radar at 35 °N used by Fritts *et al.* (1990) estimated the mean flow forcing to be around $1 \text{ ms}^{-1}\text{day}^{-1}$, while measurements in the mesosphere (71 – 80 km) by Nakamura *et al.* (1993) at the same latitude using an MU radar

estimated the zonal gravity-wave drag to be up to $61 \text{ ms}^{-1}\text{day}^{-1}$. This increase in gravity-wave drag with height and the approximate values are broadly consistent with all three model results described above.

Figure (III.43) shows the zonal mean zonal wind variance calculated by the STEVE model for the March equinox during solar minimum and figure (III.44) shows the zonal mean zonal wind variance for the June solstice. The regions of high zonal wind variance tend to coincide with the regions of high gravity-wave drag shown in figures (III.39) and (III.40) because the MK2000 scheme uses the horizontal wind variance as a major parameter in the calculation of the gravity-wave dissipation, as discussed in Chapter (II). The general increase in wind variance with altitude is related to the increase in gravity wave amplitude as the density increases with altitude. The seasonal variability in wind variance due to gravity wave activity will be discussed in Chapter (IV), but it is worth noting here that the typical values calculated by the STEVE model are consistent with those observed.

Figure (III.45) shows the zonal mean eddy diffusion coefficient calculated by the STEVE model for the March equinox during solar minimum and figure (III.46) shows the eddy diffusion coefficient for the June solstice, which can be compared with the globally averaged profile shown in figure (III.10). The zonal mean profiles for both equinox and solstice calculated by the STEVE model show that the eddy diffusion coefficient peaks at mid to high latitudes in the mesopause region, coinciding with the peak in gravity-wave dissipation (figures III.39 and III.40). The globally averaged profile based on Garcia and Solomon (1985) shows the same vertical structure and typical values as that calculated by the STEVE model, although it does not reflect the latitudinal variation in the eddy diffusion coefficients. Figure (III.47) shows the zonal mean eddy diffusion calculated by Meyer (1999b) using the GSWM and the Meyer (1999a) gravity-wave scheme for June. Despite using very different gravity-wave parameterisations, both models predict peaks in the eddy diffusion coefficient of similar values (within a factor of 2, which is well within that constrained by observations and high resolution fluid models) and with peaks in the mid to high latitude mesopause region. The twin maxima in the eddy diffusion coefficients calculated during winter in the STEVE model are not seen in the Meyer (1999b) results, presumably due to differences in the gravity-wave schemes used as the strong smoothing function which is applied to the Meyer (1999b) results (see Section II.4.2).

It seems likely that both models will overestimate the eddy diffusion coefficient as both parameterisations assume that all energy from dissipating gravity waves is converted into eddy-scale turbulence, with no estimate of secondary gravity waves produced by

this dissipation. This is supported by estimates based on MU radar observations (e.g. Fukao *et al.*, 1994) and models of meteor trail diffusivities (e.g. Hall, 2002) which estimate the peak value of the eddy diffusion coefficient to be several tens of m^2s^{-1} . The effect of overestimating the eddy diffusion coefficient on global-scale atmospheric models will be discussed in more detail in Chapter (VI), but it is worth noting here that it effects the chemical composition of the middle atmosphere and the characteristics of short wavelength resolved waves and is therefore of less importance to the STEVE model than the CMAT model discussed in Chapter (VI).

As discussed in Chapter (II), much research is still ongoing into both the observed global structure of gravity wave activity and the assumptions used in the gravity-wave parameterisations. The STEVE model has been shown to calculate gravity-wave drags, horizontal wind variances associated with gravity waves and eddy diffusion coefficients which are consistent (within a factor of 2) with other global-scale models and a meridional circulation in the mesosphere which produces temperatures and zonal which are consistent with observations. The study presented in Chapter (IV) will focus on variations in gravity wave activity in the upper stratosphere and mesosphere over decadal timescales, for which the STEVE model is ideally suited. Only results from the STEVE model in the height range 40 – 110 km will be considered in this study as it has been shown in this Chapter that the model does not represent the lower stratosphere and middle thermosphere well.

III.5 THE STEVE-3D MODEL

An extension of the STEVE model, STEVE-3D, has been developed to study the effect of gravity wave activity in the mesosphere on planetary waves in the stratospheric winter polar region in Chapter (V). As the STEVE model has already been discussed at length, only the key improvements to the model will be detailed here. Climatologies from the STEVE-3D model will be presented in section (III.6).

For the STEVE-3D model, the zonally averaged grid from STEVE has been replaced by a three dimensional grid with the same vertical and latitudinal scale and 5° longitudinal resolution (72 grid points). This significantly increases the computation required to run atmospheric simulations, making it unsuitable for studies such as that presented in Chapter (IV), but enabling the model to resolve planetary waves in the middle atmosphere. These planetary waves are introduced through the longitudinally and temporally varying geopotential height of the lower boundary of the model. For this, a non-zonally averaged form of the SSU data used in the STEVE model is employed, shown in figure (III.48) for the 21st December 1992. While the zonal mean geopotential acts in the same way as discussed in section (III.3.1), longitudinal

planetary wave structures can be clearly seen in the geopotential height data, especially in the winter polar region. It is important to note that at present the STEVE-3D model has only been tested with a single 90 day data set from 1st December 1992 until 28th February 1993. Further, for numerical stability, the model starts from a zonal mean climatology and the lower boundary geopotential height perturbations are introduced gradually over 10 days. This means that only data from January 1993 will be considered to represent the winter period in this study. Future development of the STEVE-3D model will focus on validating the model with other lower boundary data sets such as the National Centre for Atmospheric Research NCEP reanalysis data set (Kalnay *et al.*, 1996) which will enable the model to simulate over 40 years of atmospheric conditions (see Chapter VII for further details).

The other key development required for the STEVE-3D model also relates to planetary wave activity. The zonally averaged STEVE model used an unrealistically high horizontal diffusion coefficient (discussed in Section III.3.5) to compensate for the lack of planetary wave activity in the model. As the STEVE-3D model simulates planetary wave activity, this is no longer required. This means that the representation of the middle atmosphere, especially in the stratospheric polar regions where the STEVE model was shown to be deficient, should be more accurate in the STEVE-3D model than its zonally averaged counterpart. Results from the STEVE-3D model will be presented in the following section and briefly compared with observations and other model results to show these improvements.

All other aspects of the STEVE-3D model are identical to the STEVE model and as such will not be discussed further. It is worth noting that this means that the diurnal cycle in solar heating is not included and therefore atmospheric tides are absent from the STEVE-3D model. This enhances the computational efficiency of the model and also allows planetary wave / gravity wave coupling to be studied (Chapter V) in the absence of planetary wave / tidal coupling or gravity wave / tidal coupling (discussed in Chapter VI). It is hoped that future development of the STEVE-3D model will allow the inclusion of atmospheric tides so that their effect on the results in Chapter (V) can be clearly assessed.

III.6 STEVE-3D MODEL CLIMATOLOGIES

This section will present the results from the STEVE-3D model for the simulation of January 1993 and highlight the advantages of this model over the STEVE model. The results will focus on the stratosphere and mesosphere as the STEVE and STEVE-3D models do not simulate the thermosphere well, as discussed earlier. As the results of

the STEVE model have already been presented in detail, only key differences between the results of the two models will be addressed here.

Figure (III.49) shows the zonal mean temperatures calculated by the STEVE-3D model, averaged over January to remove any transient wave effects. While the absolute values differ slightly, the main features of this plot have been discussed in section (III.4.1). Comparing figures (III.49) and (III.14) and (III.15) it can be seen that the inclusion of planetary waves in the STEVE-3D model and the removal of the high horizontal diffusion coefficient means that the winter polar stratospheric temperatures calculated by the STEVE-3D model are in better agreement with those of the MSIS-E90 results. This is an important result as the representation of the winter polar stratosphere is key to the study presented in Chapter (V). Similarly, the local temperature maximum in the winter mesosphere has increased and is in better agreement with the MSIS-E90 temperatures than the STEVE model result.

Figure (III.50) shows the daily averaged, polar stereographic plot of the temperatures for the 9th January at 10 mb (~30 km altitude) calculated by the STEVE-3D. A strong planetary wave structure can be seen in the temperature field in the winter hemisphere, while the summer hemisphere is more symmetric with longitude as expected (e.g. figure III.48). The planetary wave structure in the winter hemisphere creates a region of warm air centred on Siberia and a cold region centred over Greenland. Comparing this to the NCEP reanalysis database for the same period, shown in figure (III.51), a good overall agreement in the temperatures and the planetary wave structure is seen, showing that the STEVE-3D model provides a good representation of the stratospheric winter polar region.

Figure (III.52) shows the zonal mean zonal winds calculated by the STEVE-3D model, averaged over January. These winds show the same basic features as those of the STEVE model with a strong eastward jet in the winter mesosphere and a westward jet in the summer mesosphere, although the strength of these has been reduced by the planetary wave dissipation. In the lower stratosphere an eastward flow is seen in both hemispheres as in the STEVE model and the HWM and URAP data. It is worth noting that averaging over one month has removed the numerical instability at the equator from figure (III.52), although it is still present in the data.

In summary, the STEVE-3D model reproduces the main features seen in the STEVE model in the mesosphere including the meridional circulation driven by dissipating gravity waves. The inclusion of planetary wave activity in the STEVE-3D model has improved the representation of the stratosphere, allowing the STEVE-3D model to

reproduce the observed structure of the winter polar stratosphere during January. The STEVE-3D model is therefore an ideal tool for the study presented in Chapter (V), which will explore the coupling between planetary waves and gravity waves in the middle atmosphere and the effect of this coupling on the winter polar stratosphere region.

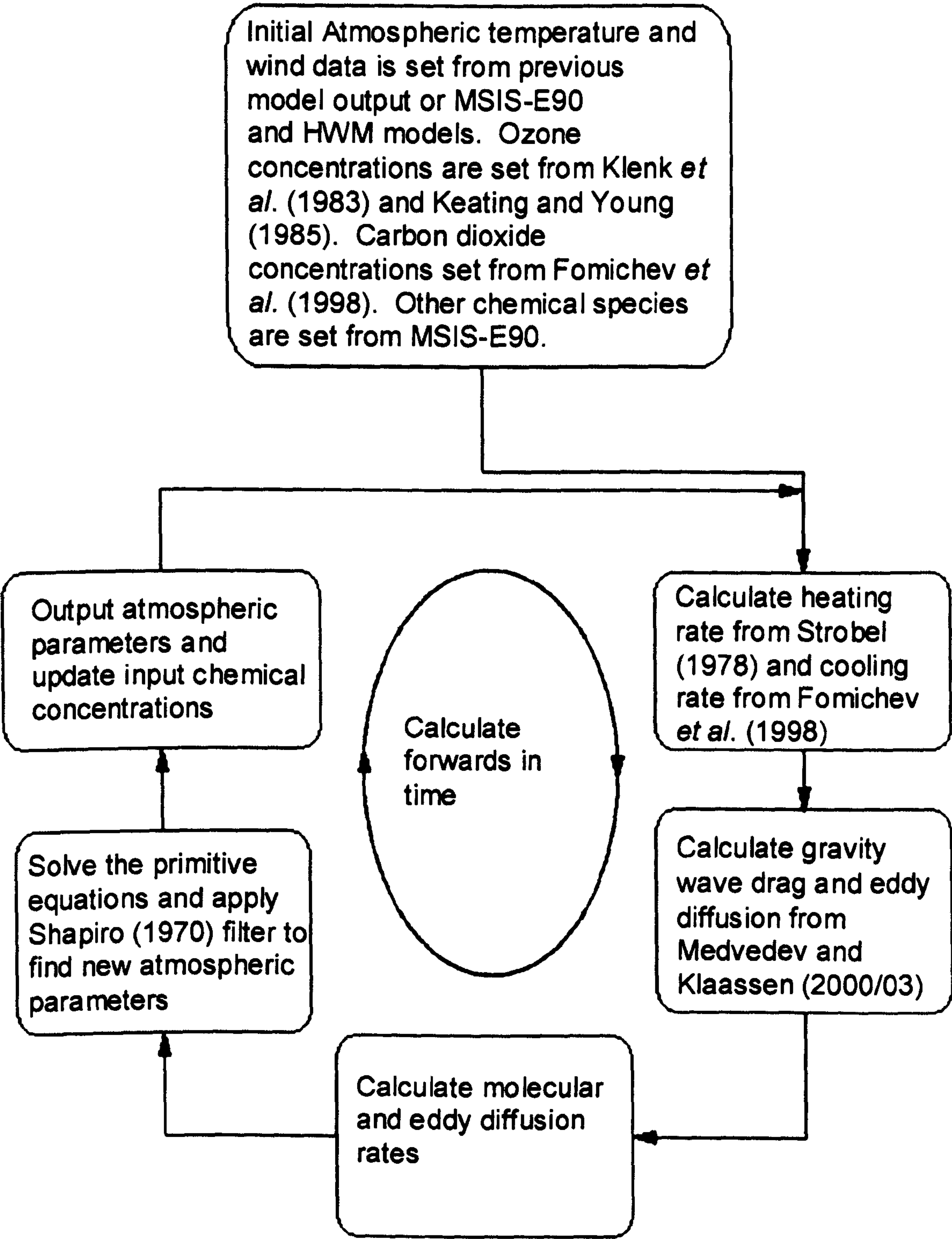


Figure III.1, Schematic of the key components of the STEVE model

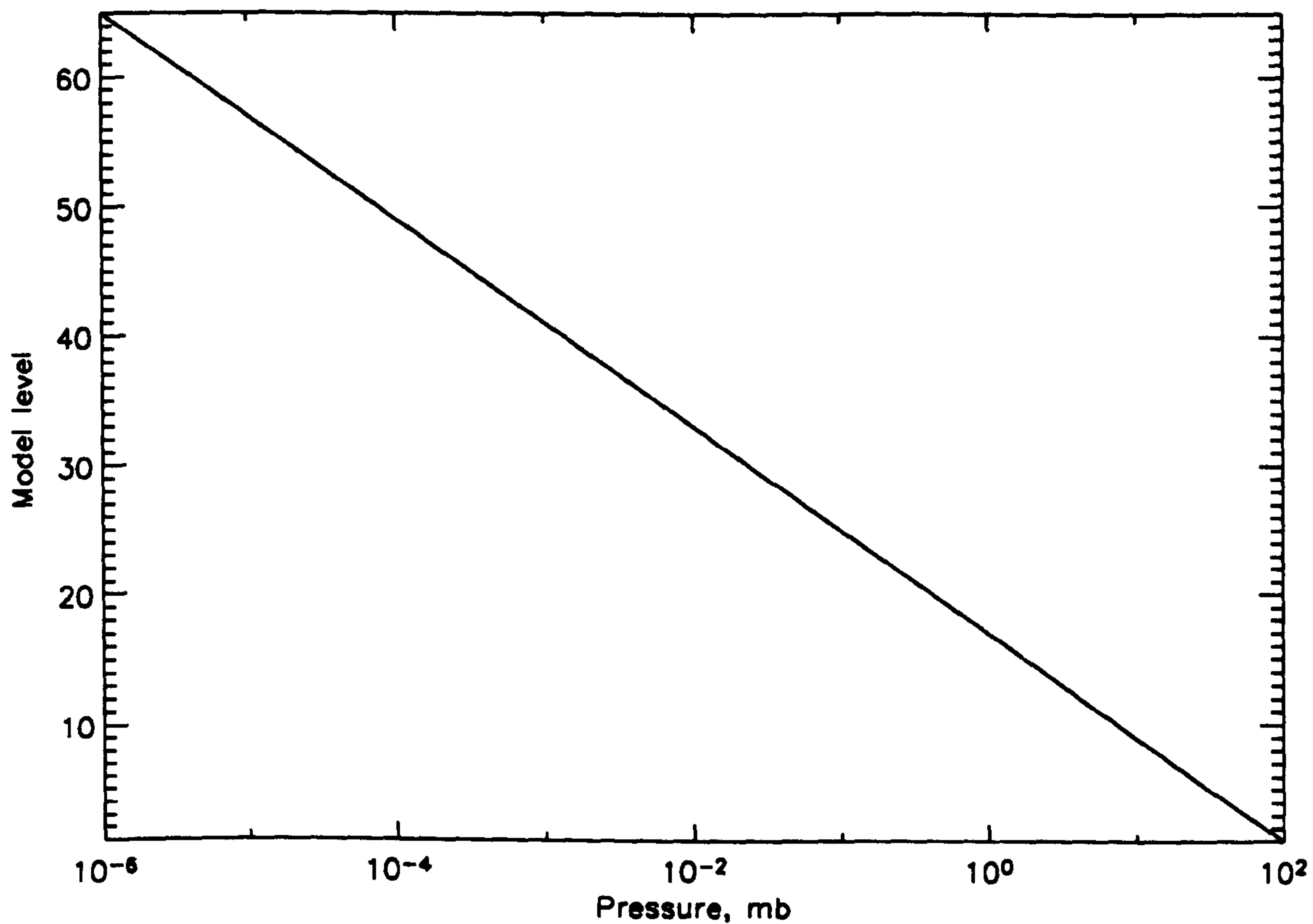


Figure III.2, Constant pressure levels used by the STEVE model.

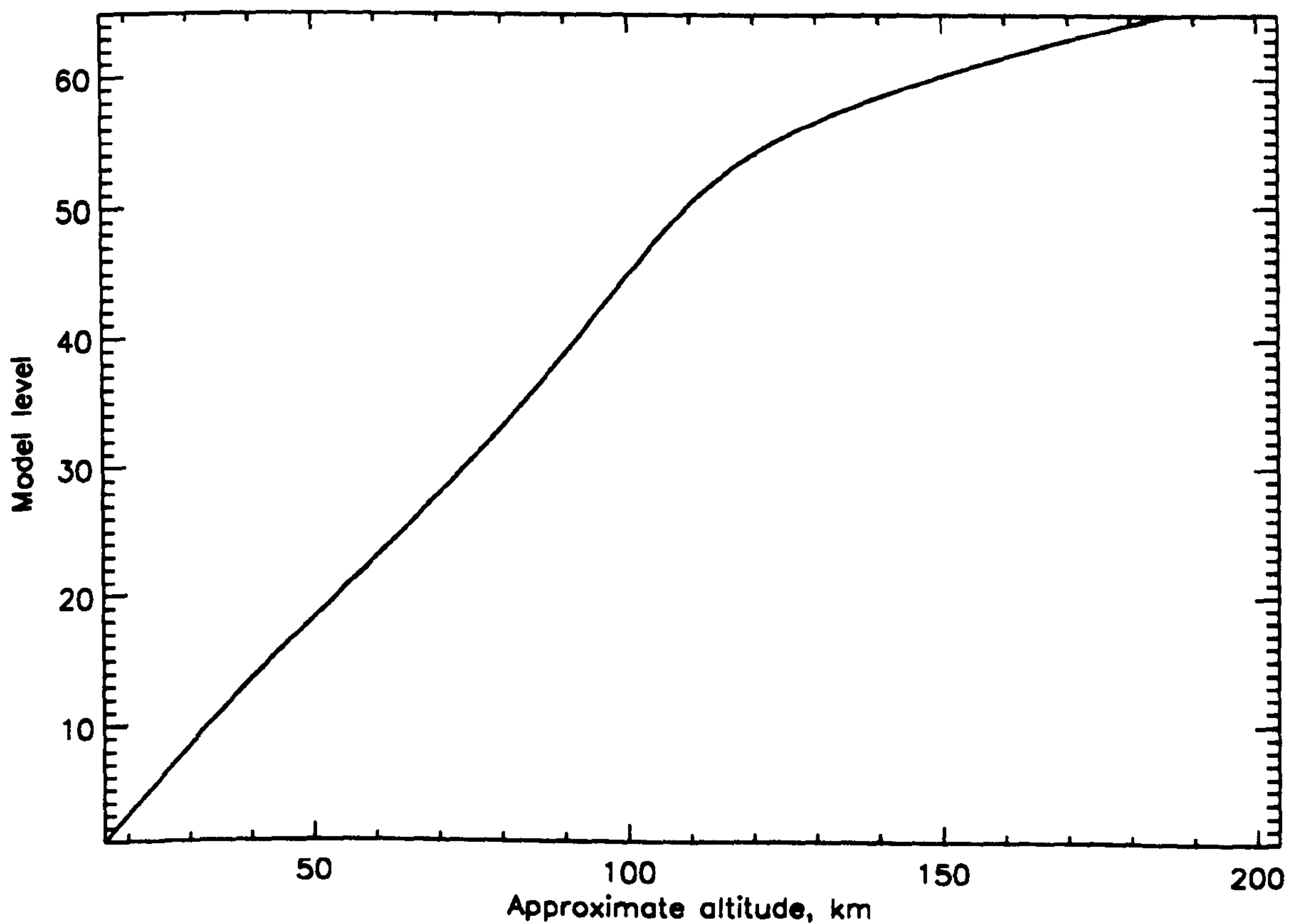


Figure III.3, Approximate altitude of each constant pressure level in the STEVE model at 45 °N during equinox at solar minimum. Data from the MSIS-E90 model, Hedin (1991).

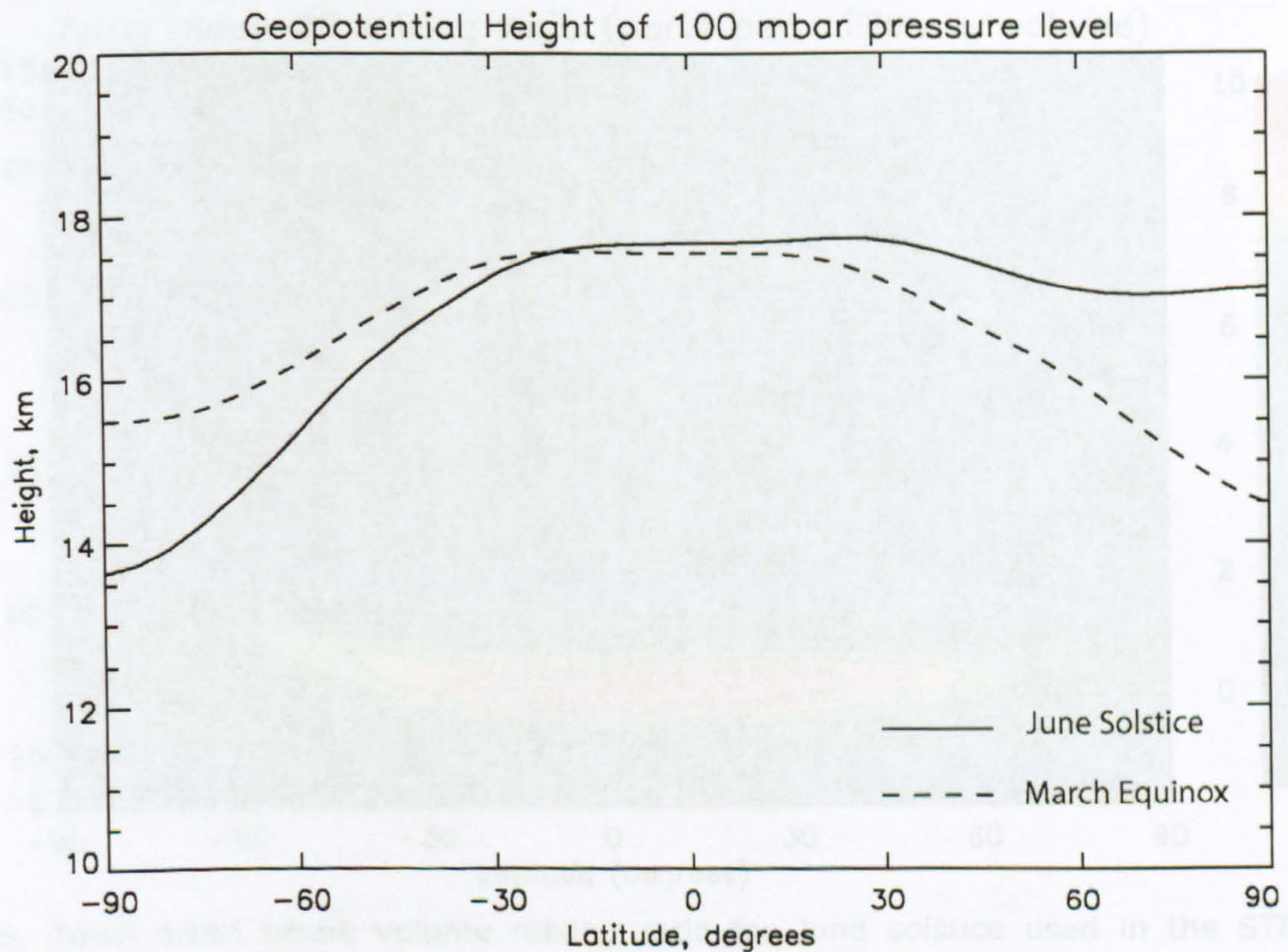


Figure III.4, Zonal mean geopotential height of the 100 mb pressure level as a function of latitude for the March equinox and June solstice during 1993. Data from the assimilated Stratospheric Sounding Unit (Miller *et al.*, 1980).

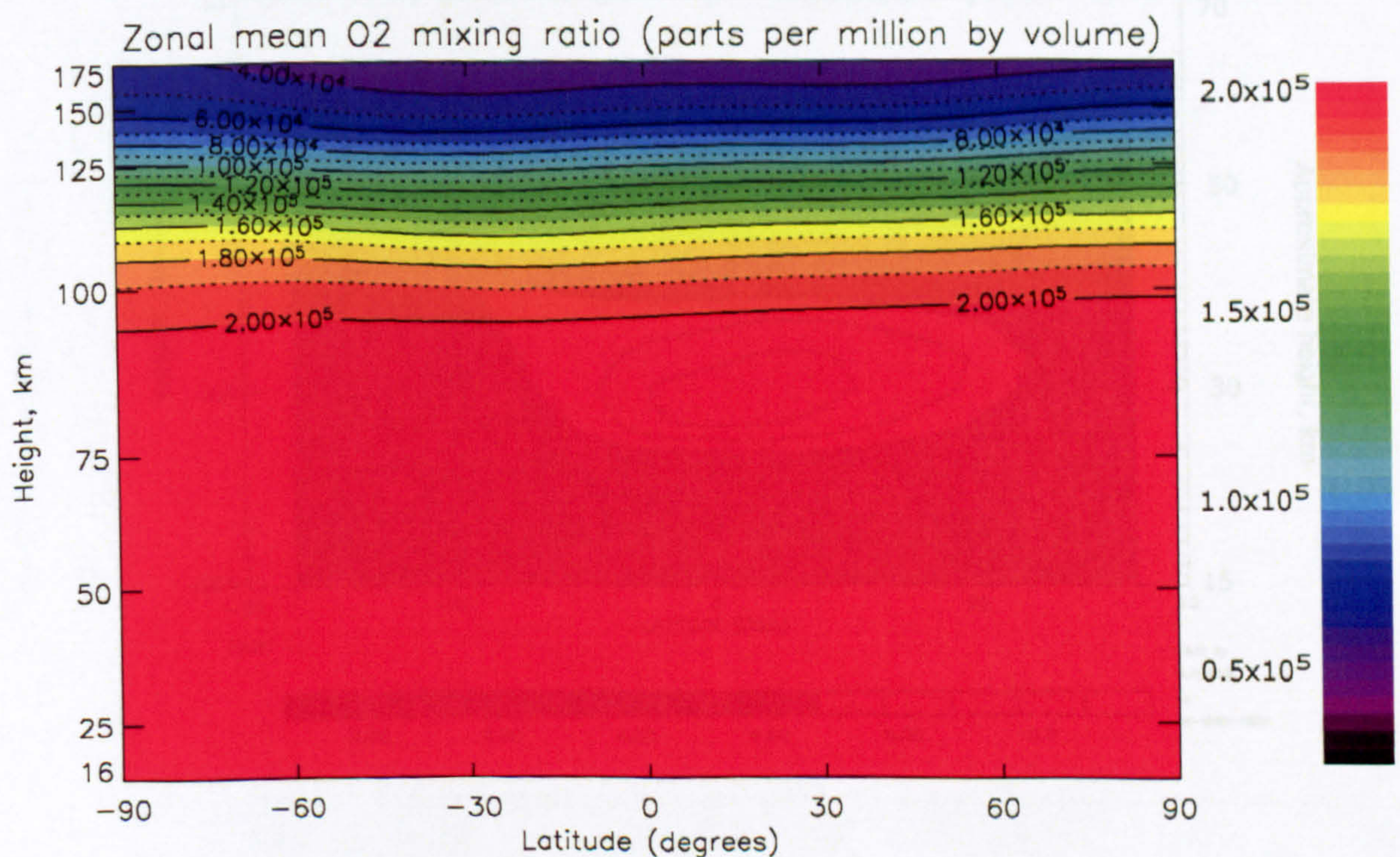
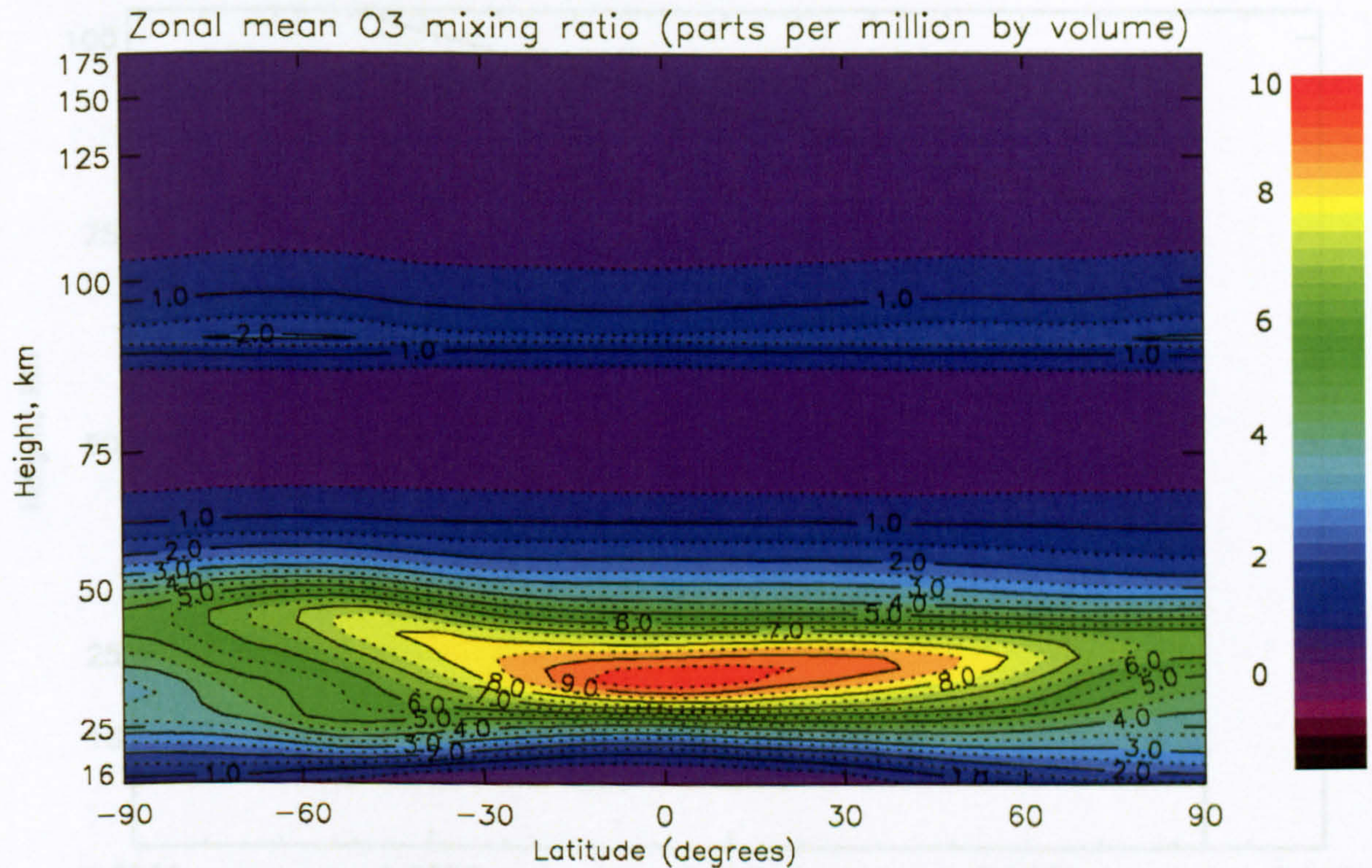


Figure III.5, Zonal mean molecular oxygen volume mixing ratio for June solstice used in the STEVE model. Data from MSIS-E90 empirical model (Hedin, 1991).



III.6, Zonal mean ozone volume mixing ratio for June solstice used in the STEVE model. Data from Klenk *et al.* (1983) and Keating & Young (1985).

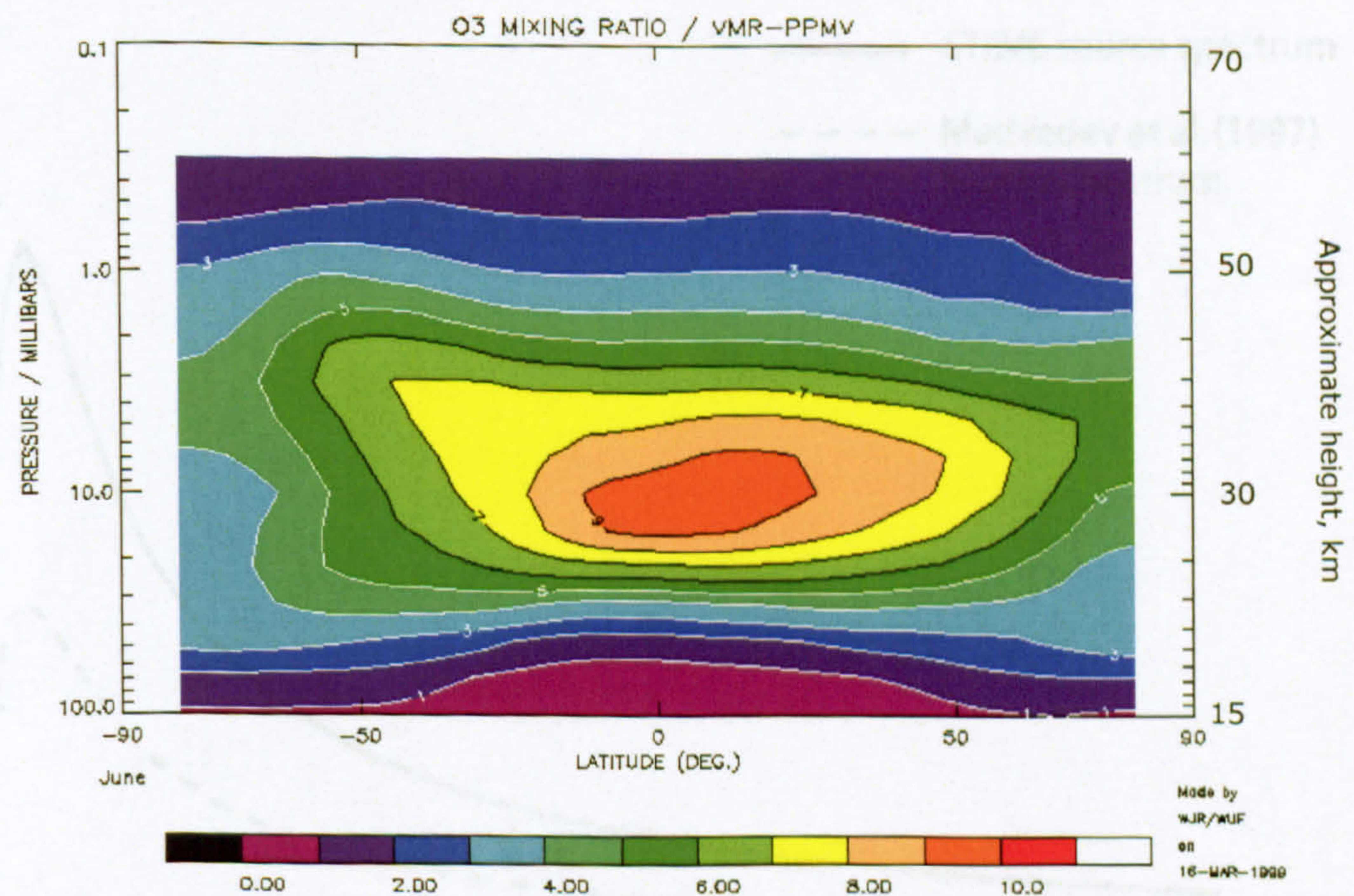


Figure III.7, Ozone volume mixing ratio climatology for June from the UARS Reference Atmosphere Project, extended data, Swinbank and Ortland (2003). Units are parts per million by volume.

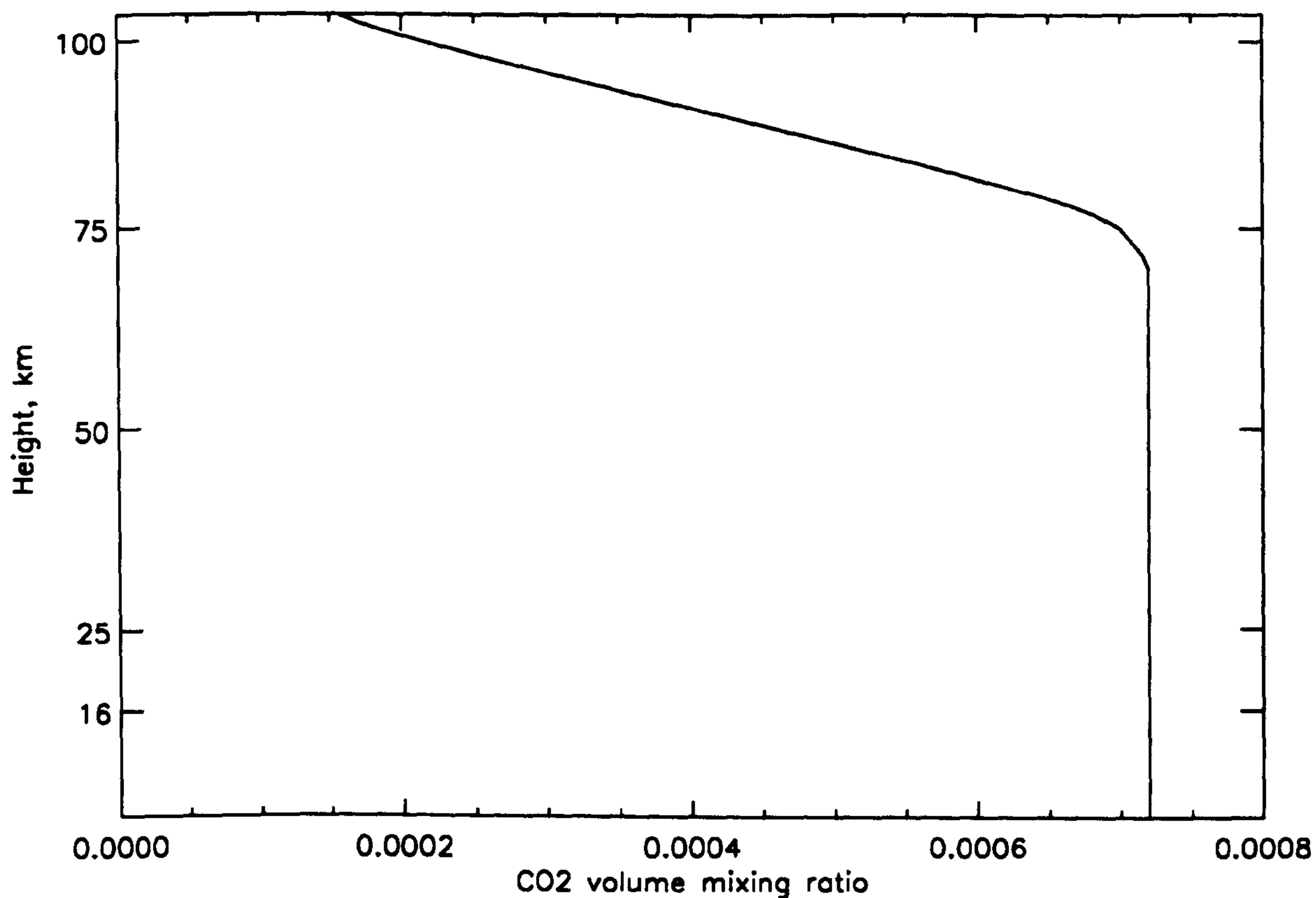


Figure III.8, Global mean carbon dioxide mixing ratio, data from Fomichev *et al.* (1998).

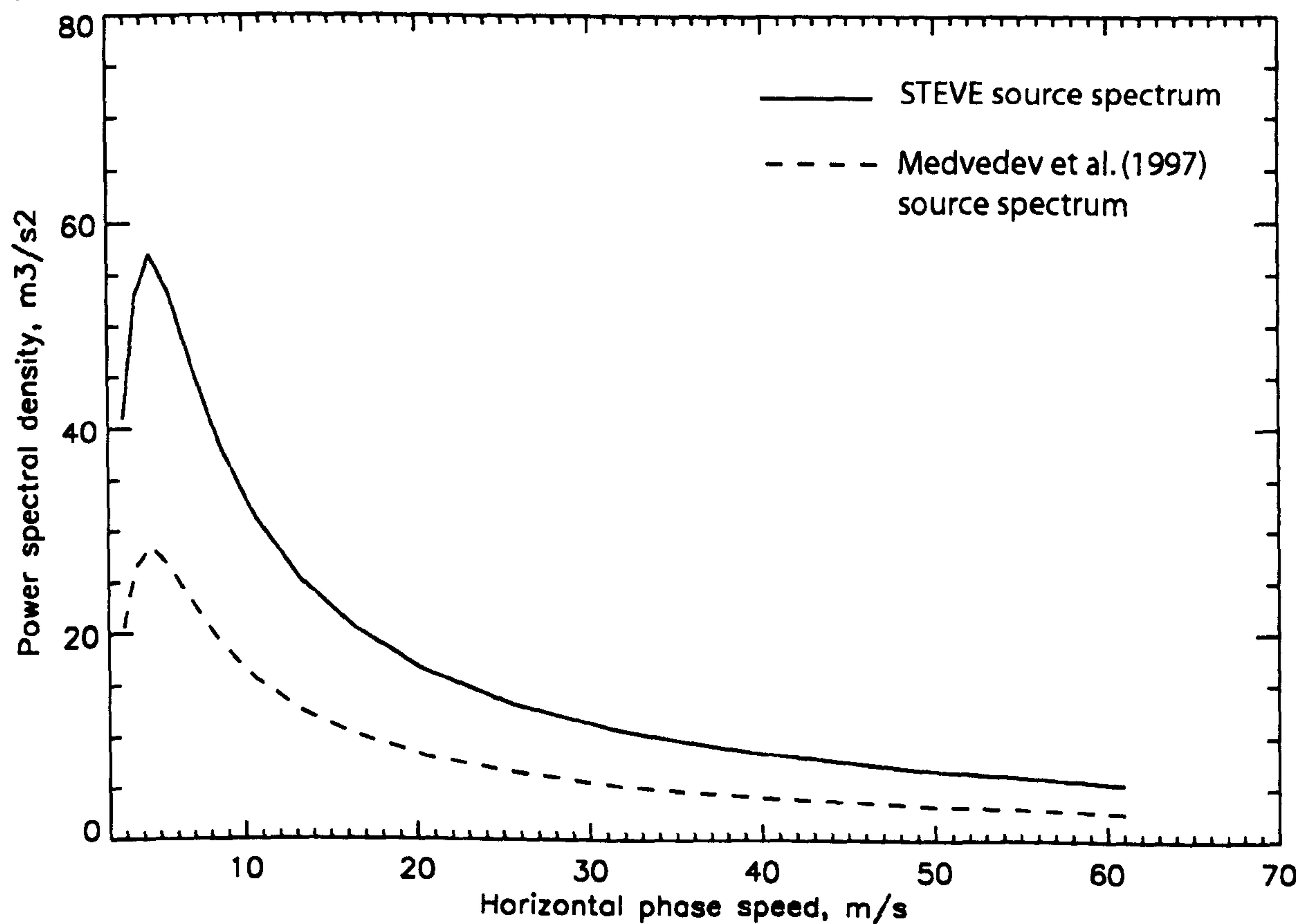


Figure III.9, Power spectral density as a function of horizontal phase speed for the source spectra used in the STEVE model and Medvedev *et al.* (1997).

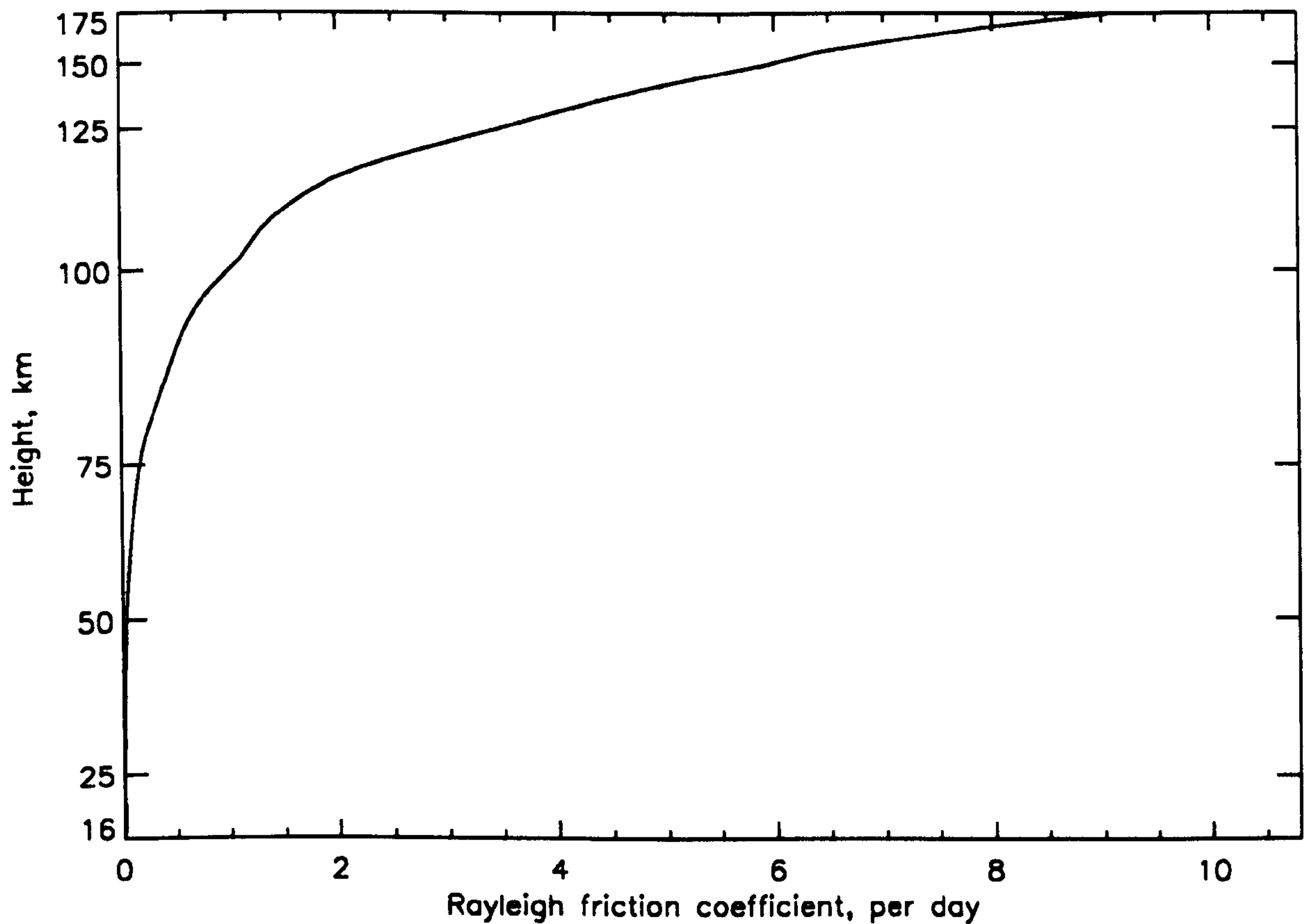


Figure III.10, Rayleigh friction coefficient used in the STEVE model as a function of height.

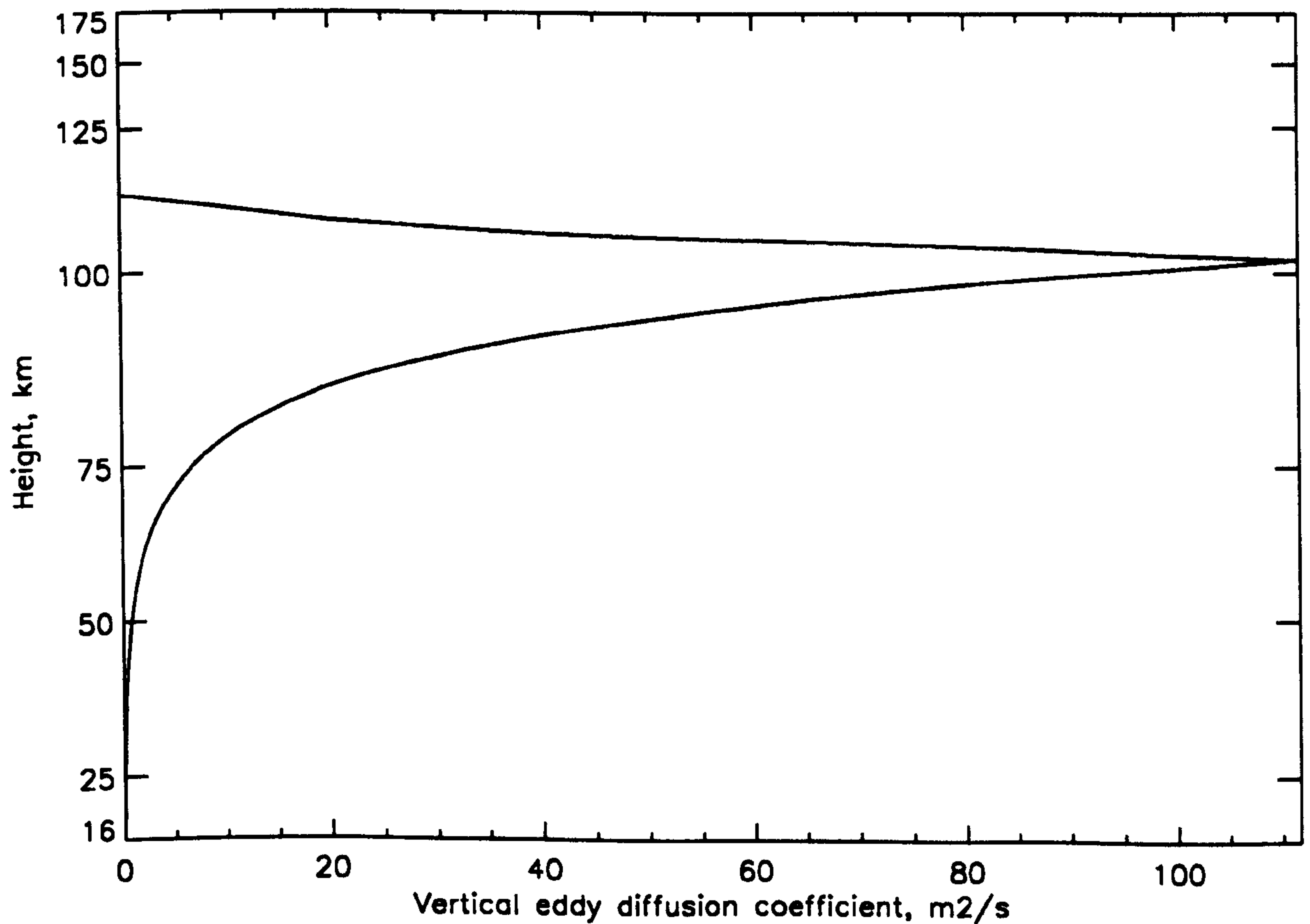


Figure III.11, Globally averaged eddy diffusion coefficient used in the STEVE model, based on Garcia and Solomon (1985).

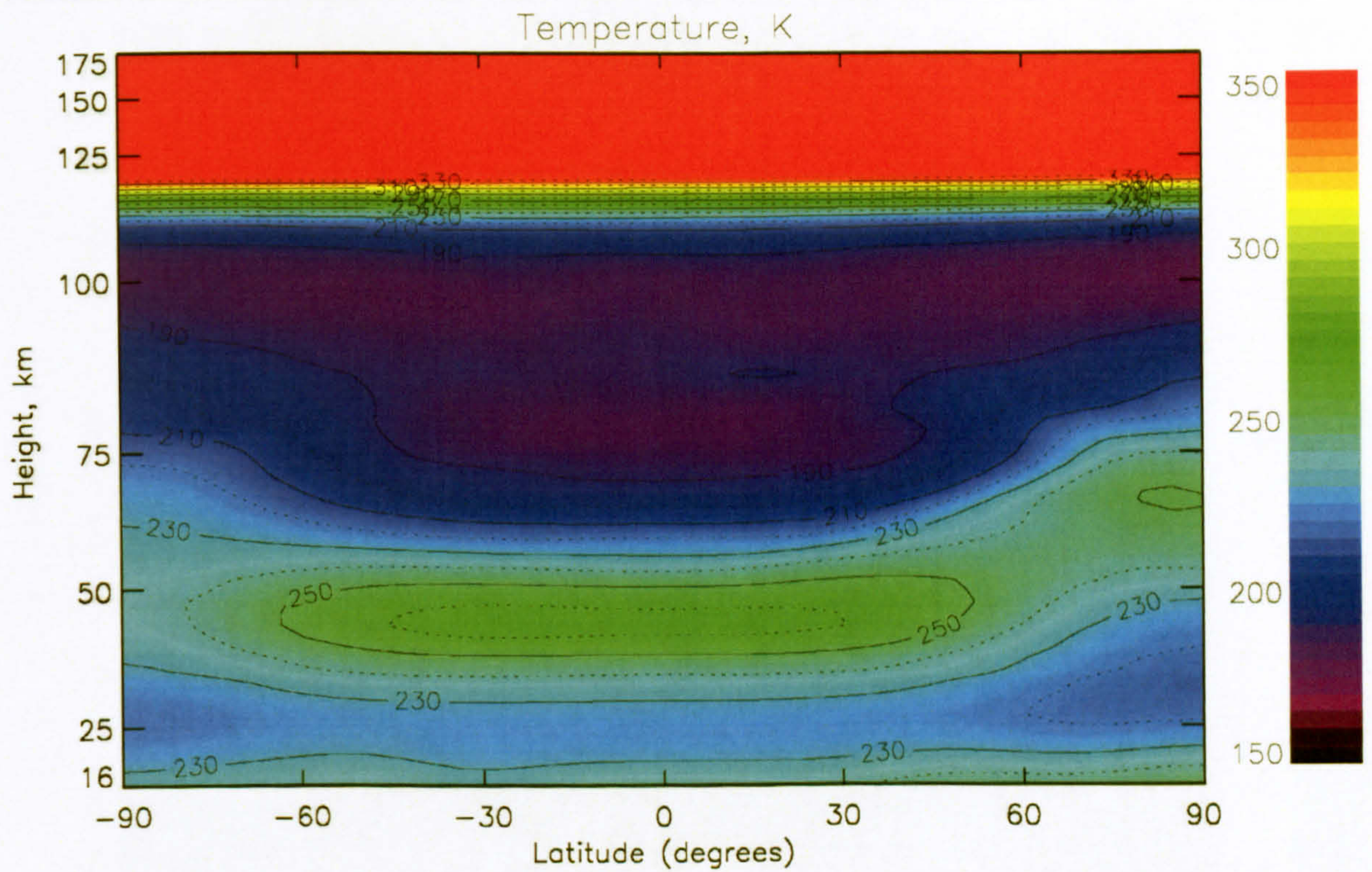


Figure III.12, Zonal mean temperature structure during the March equinox at solar minimum calculated by the STEVE model.

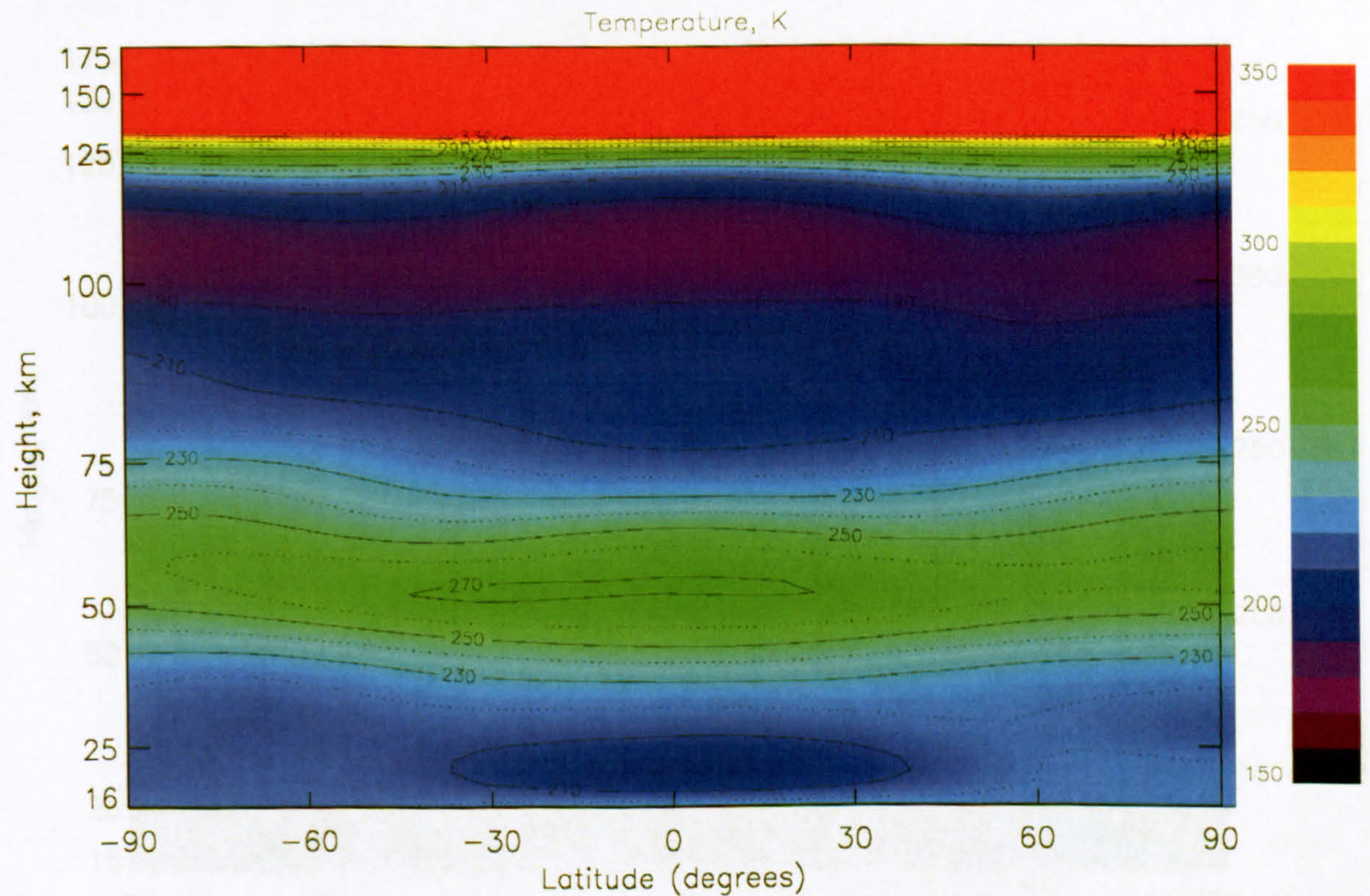


Figure III.13, Zonal mean temperature structure during the March equinox at solar minimum from the MSIS-E90 empirical model.

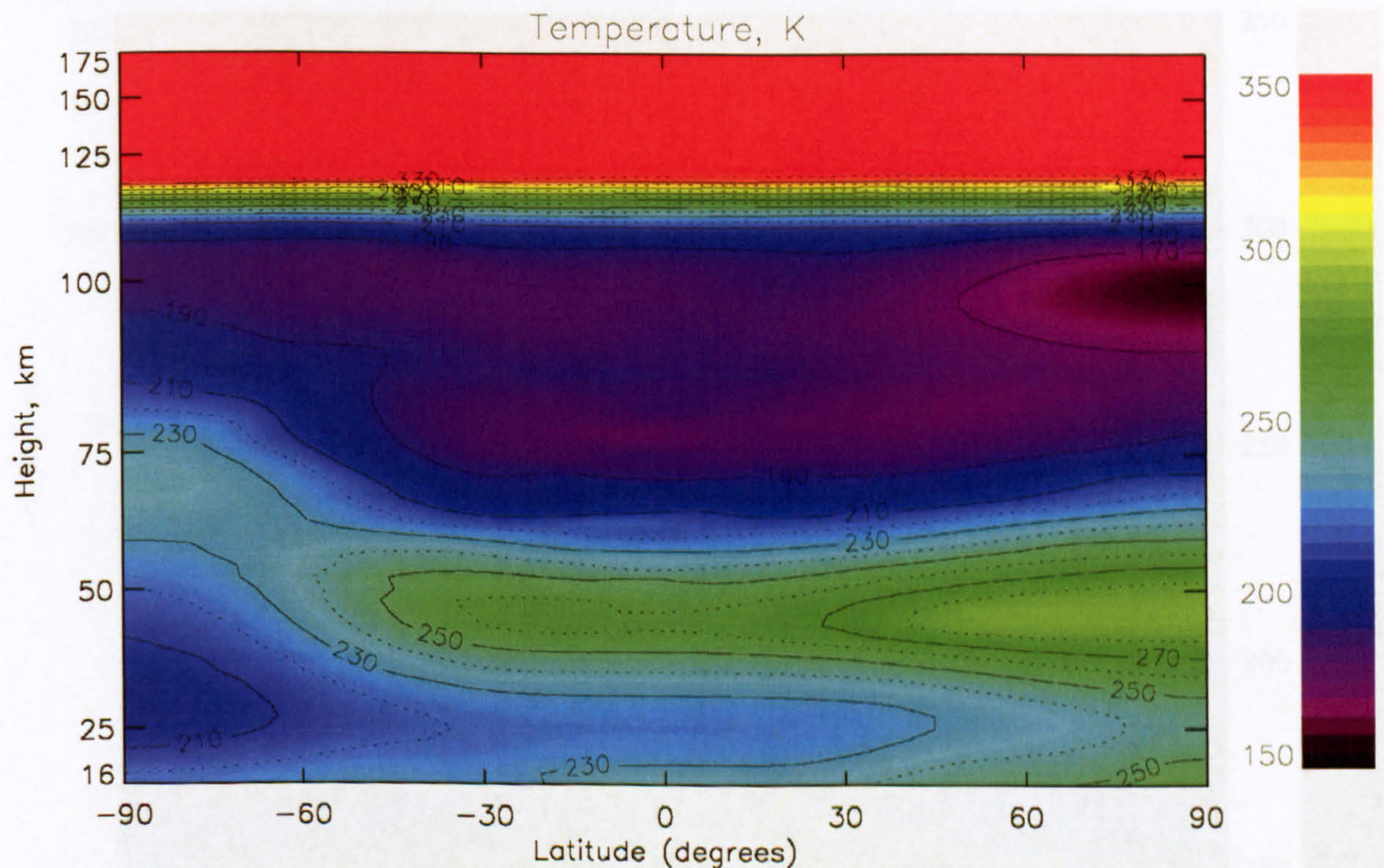


Figure III.14, Zonal mean temperature structure during the June solstice at solar minimum calculated by the STEVE model.

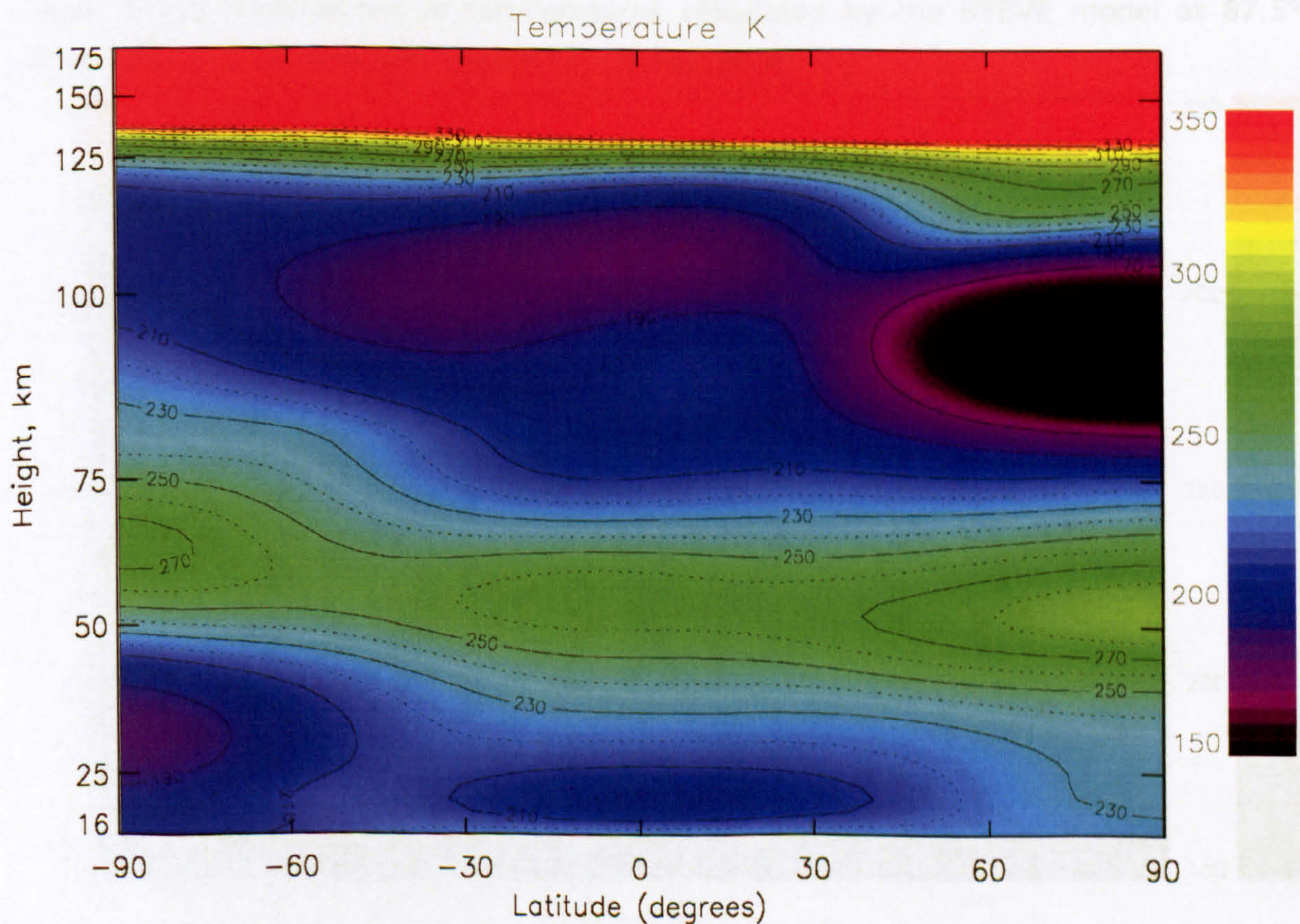


Figure III.15, Zonal mean temperature structure during the June solstice at solar minimum from the MSIS-E90 empirical model.

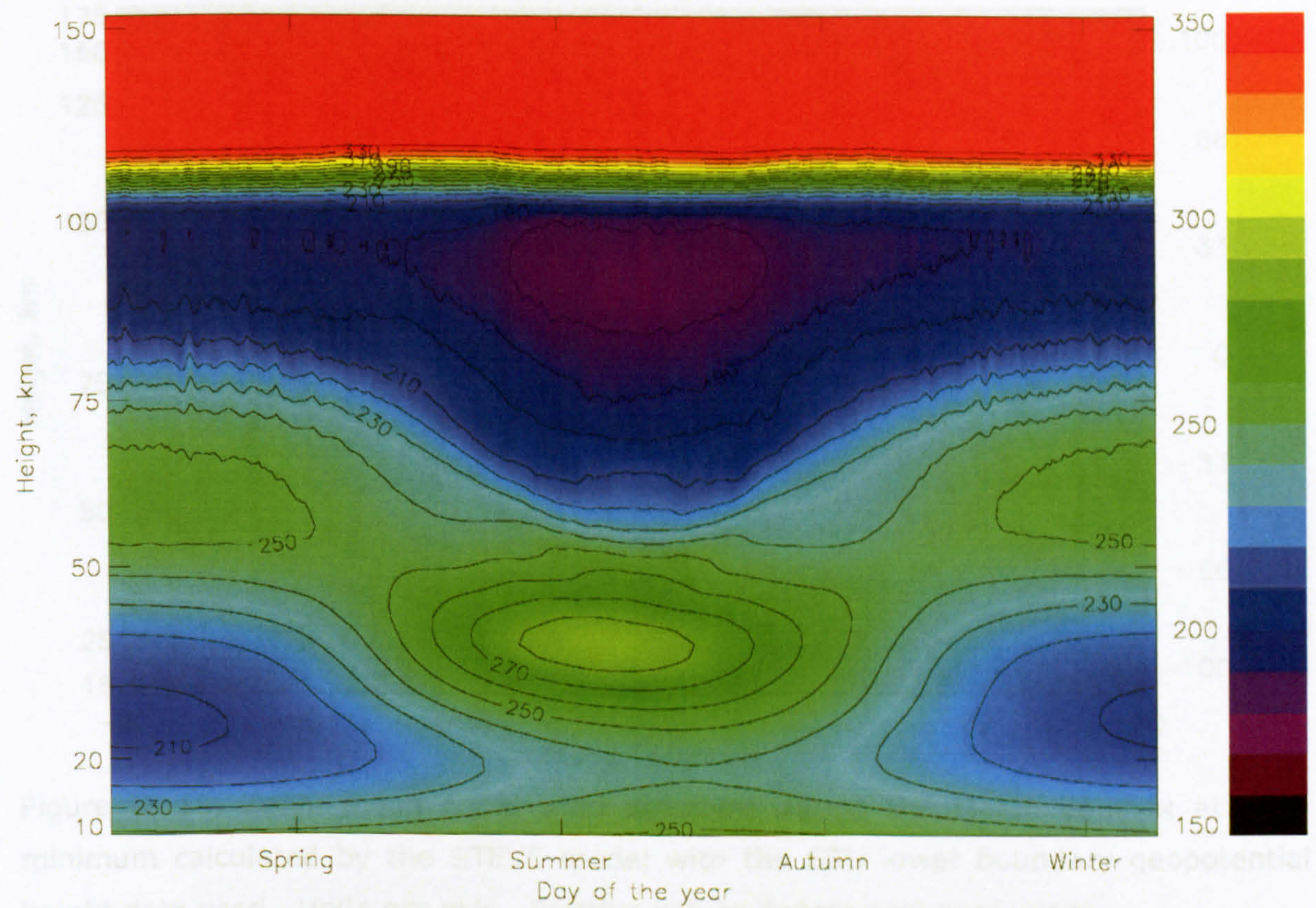


Figure III.16, Time-series of temperatures calculated by the STEVE model at 67.5° North during solar minimum conditions. Units are K.

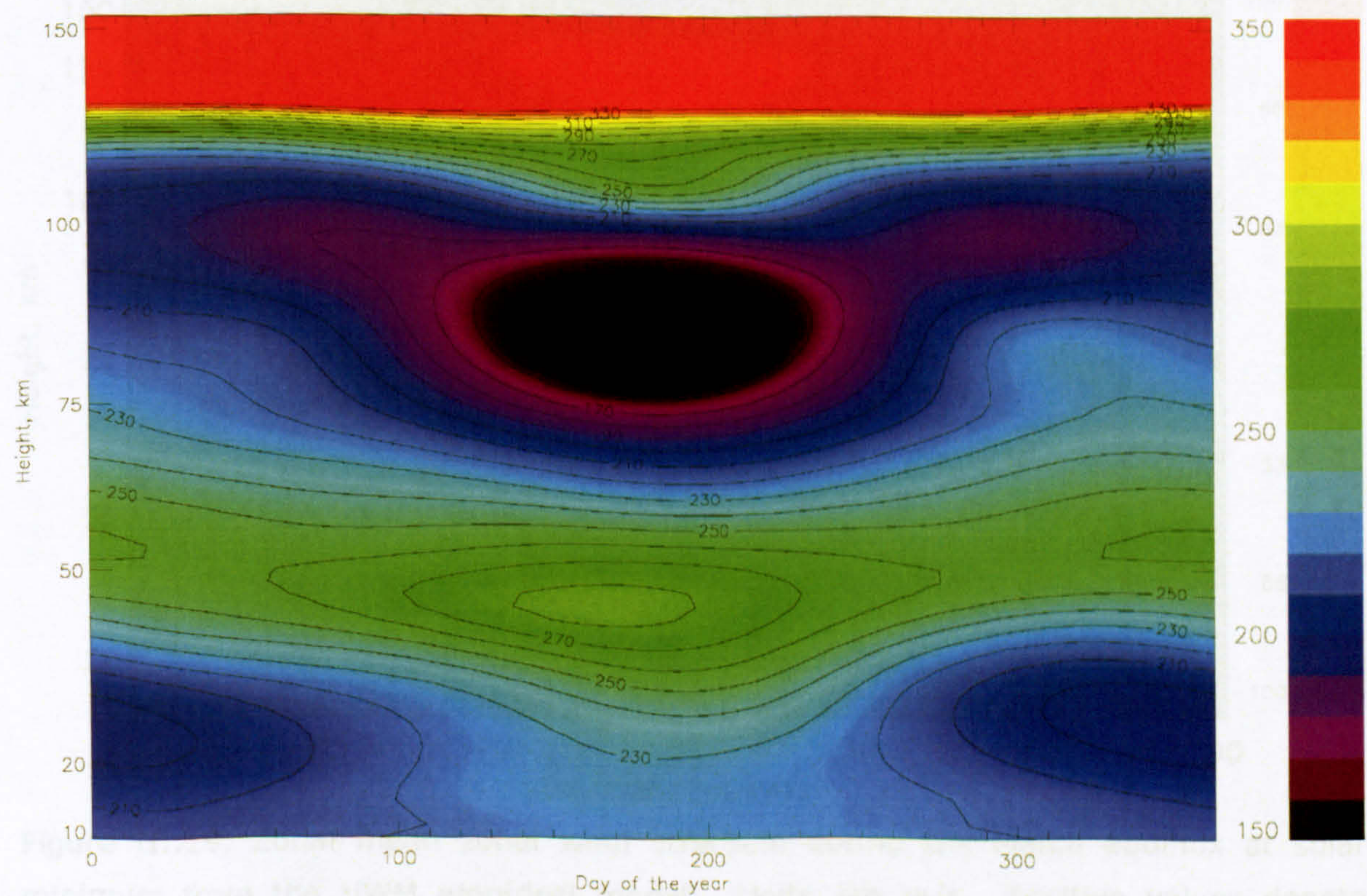


Figure III.17, Time-series of temperatures calculated by the MSIS-E90 model at 67.5° North during solar minimum conditions. Units are K.

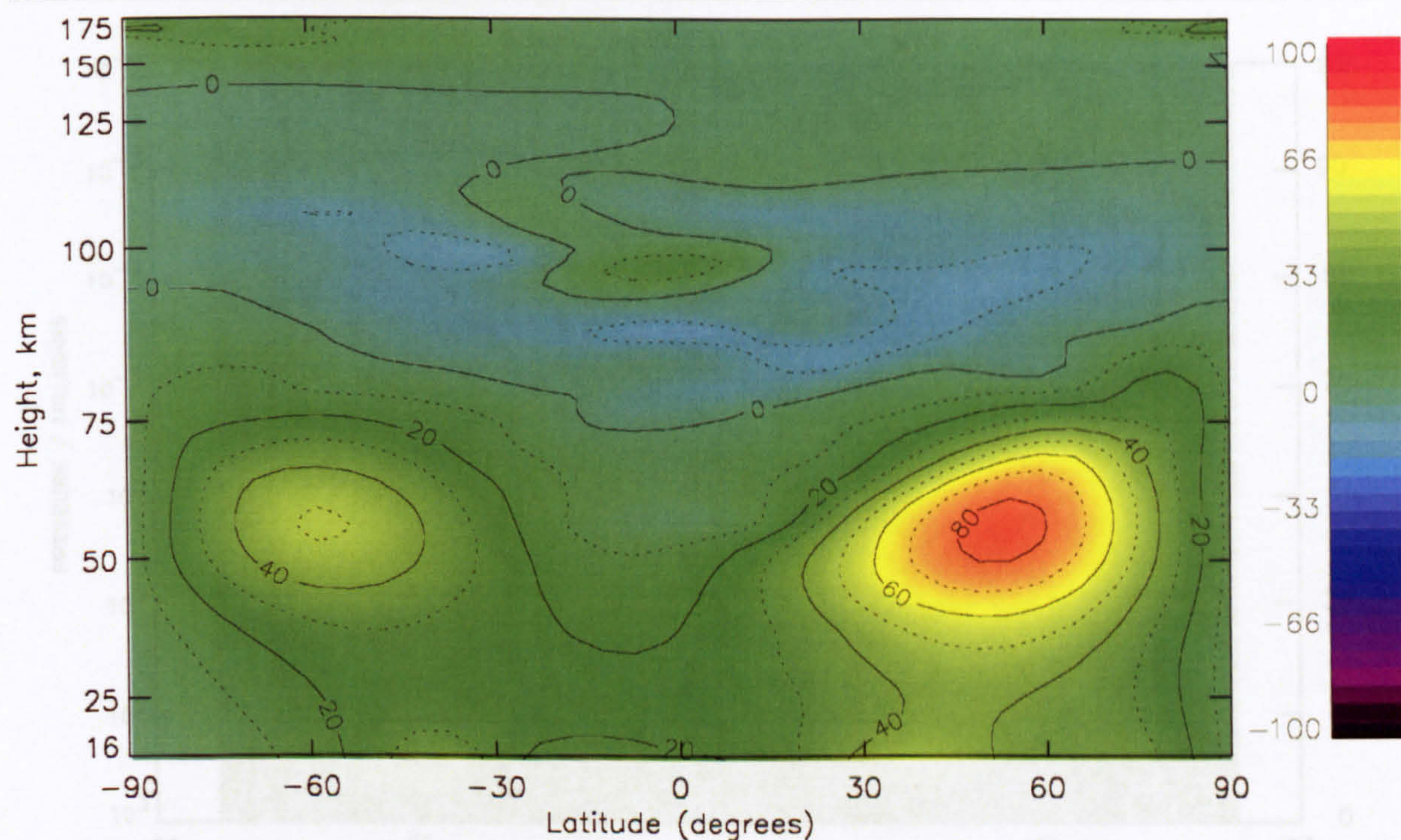


Figure III.18, Zonal mean zonal wind structure during the March equinox at solar minimum calculated by the STEVE model with the SSU lower boundary geopotential height data used. Units are m/s. Positive values denote eastward winds.

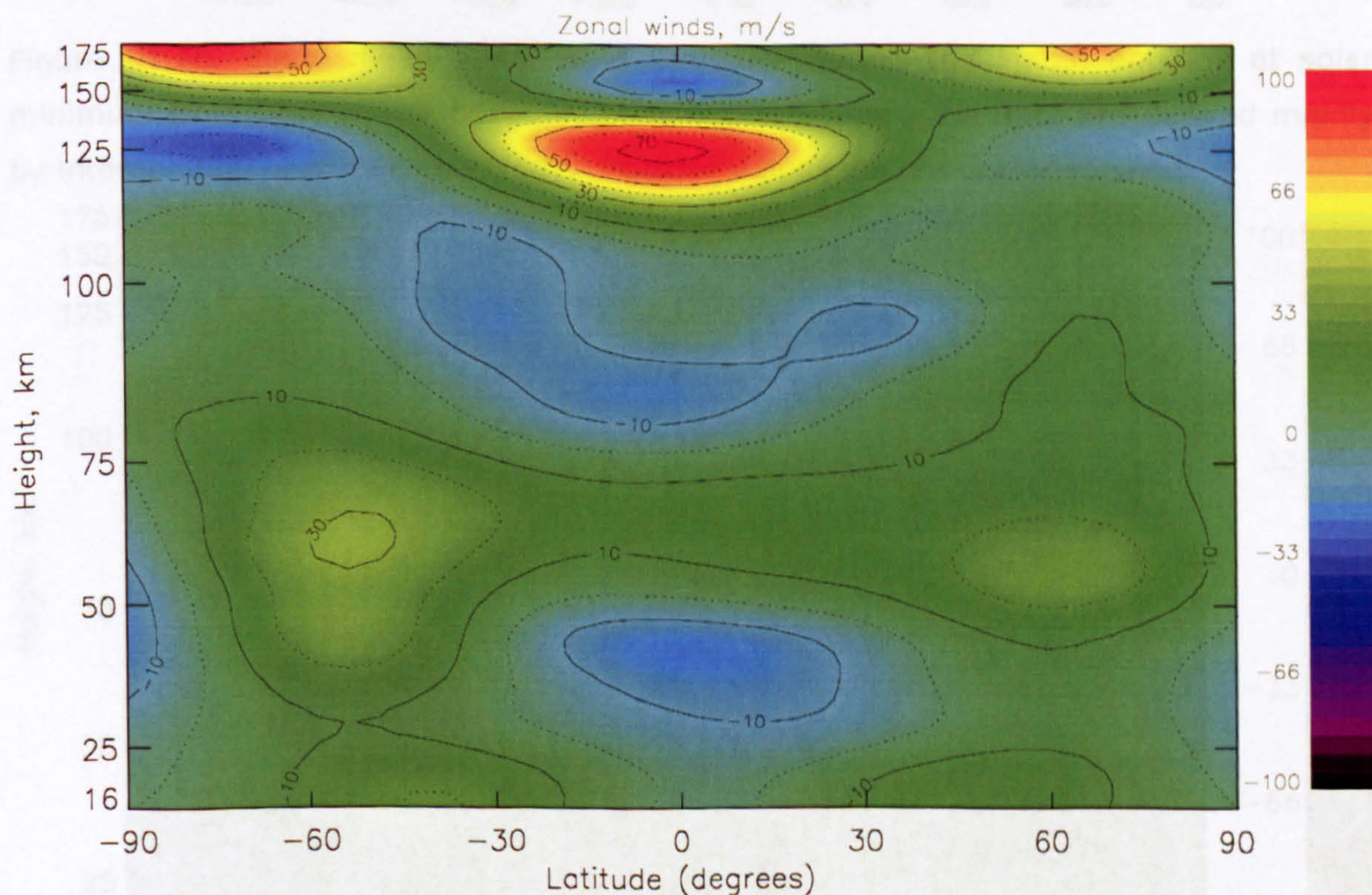


Figure III.19, Zonal mean zonal wind structure during the March equinox at solar minimum from the HWM empirical model. Units are m/s. Positive values denote eastward winds.

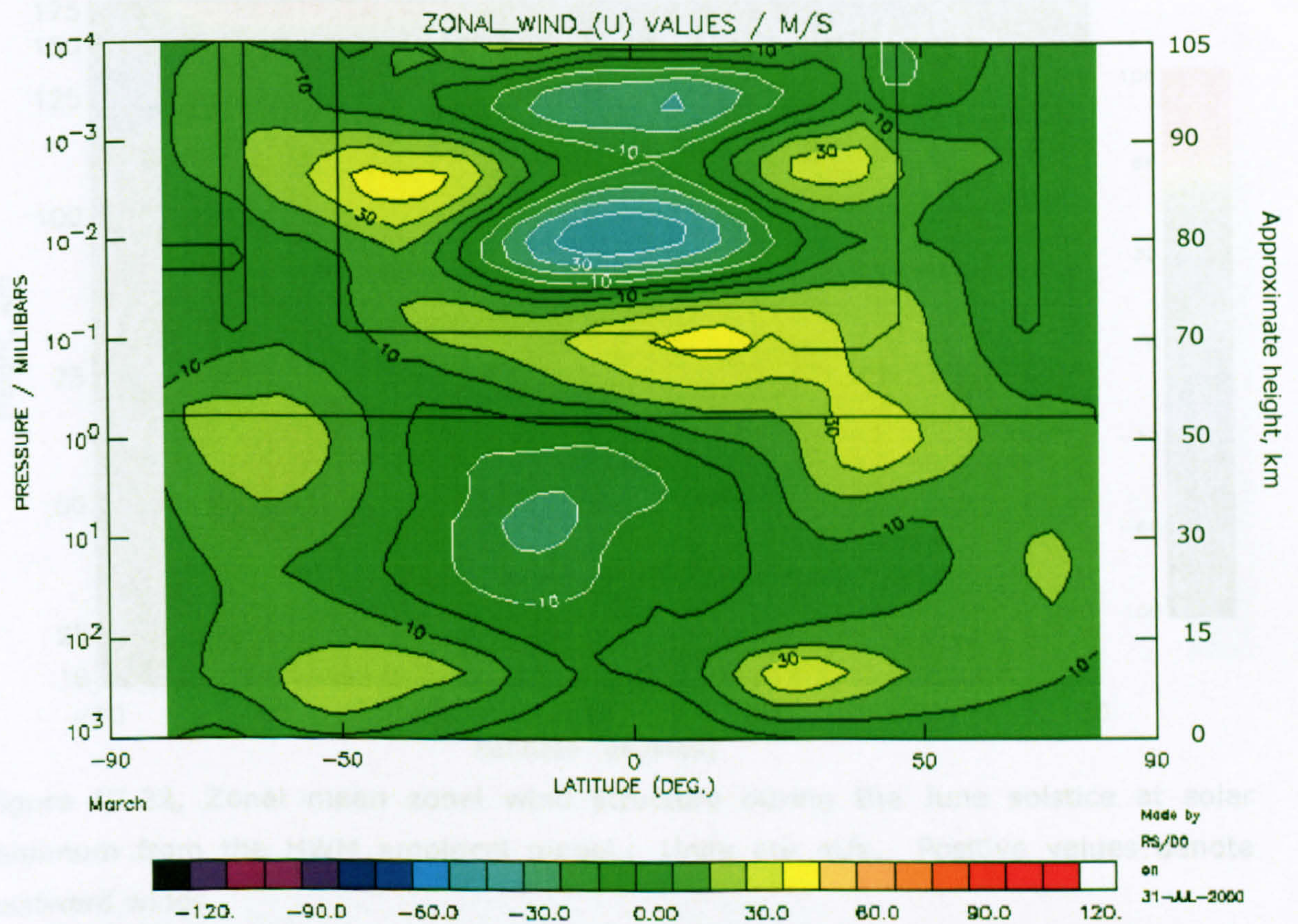


Figure III.20, Zonal mean zonal wind structure during the March equinox at solar minimum from URAP. The boxed regions indicate where the data are derived mainly by interpolation or from climatology. Positive values denote eastward winds.

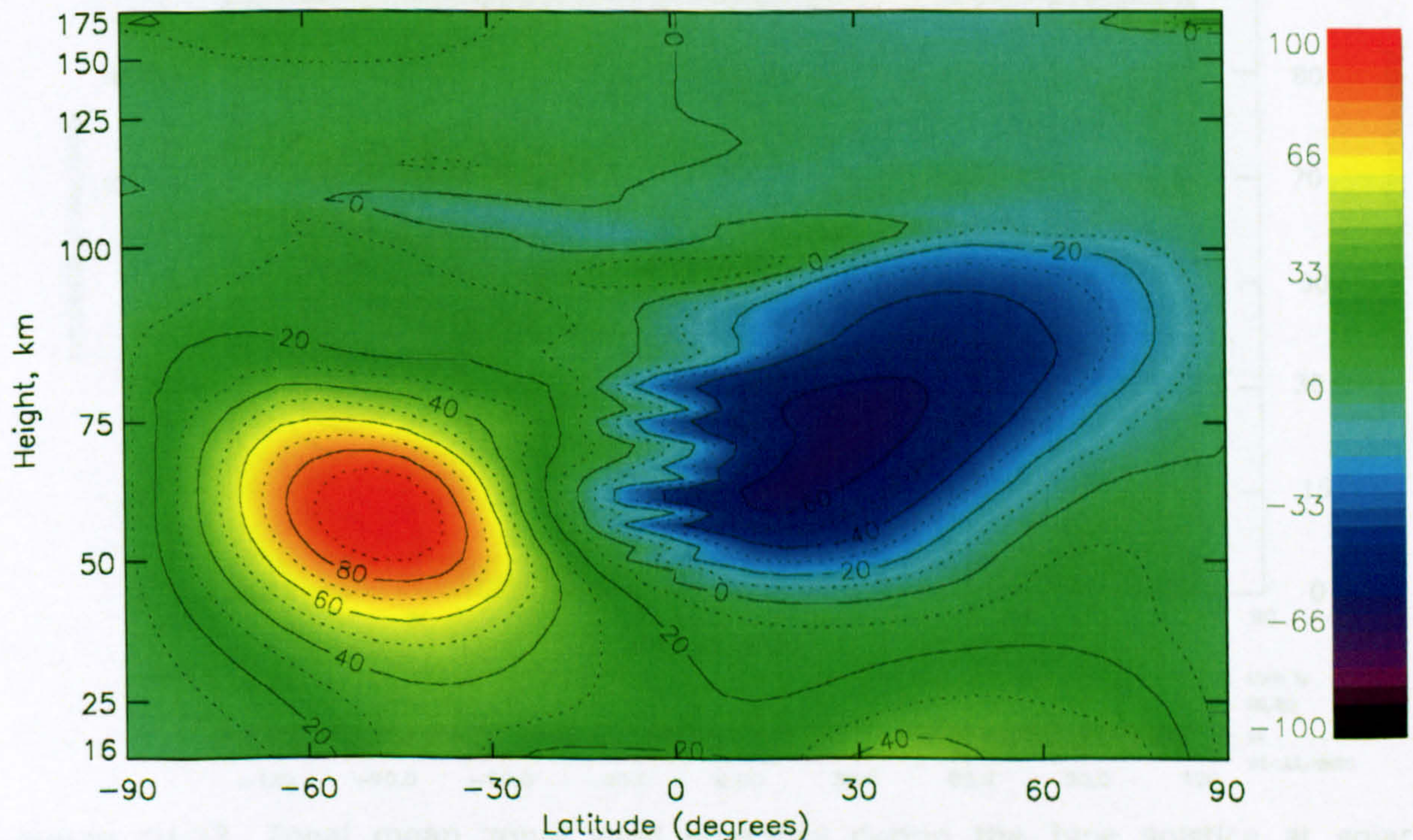


Figure III.21, Zonal mean zonal wind structure during the June solstice at solar minimum calculated by the STEVE model with the SSU lower boundary geopotential height data used. Units are m/s. Positive values denote eastward winds.

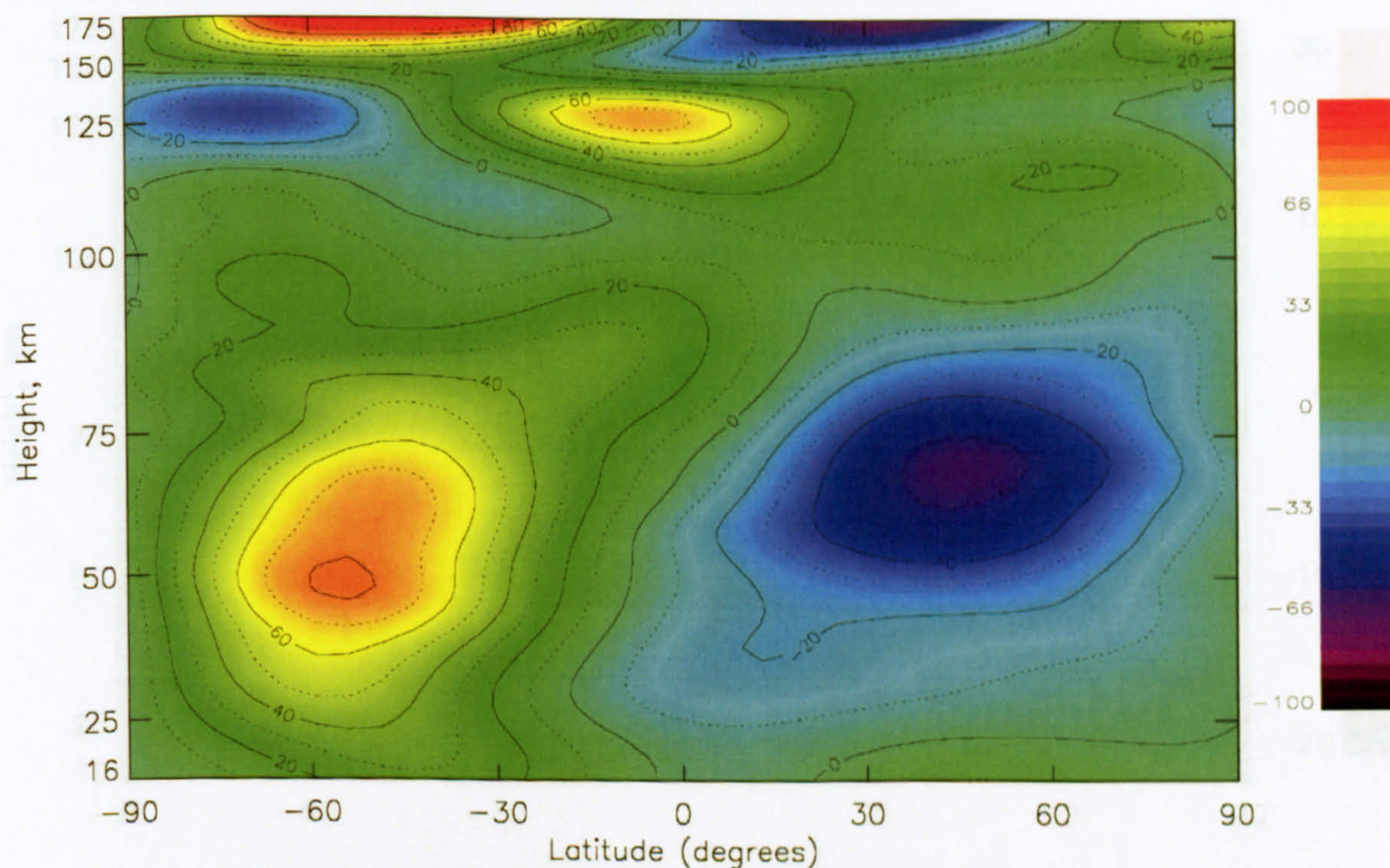


Figure III.22, Zonal mean zonal wind structure during the June solstice at solar minimum from the HWM empirical model. Units are m/s. Positive values denote eastward winds.

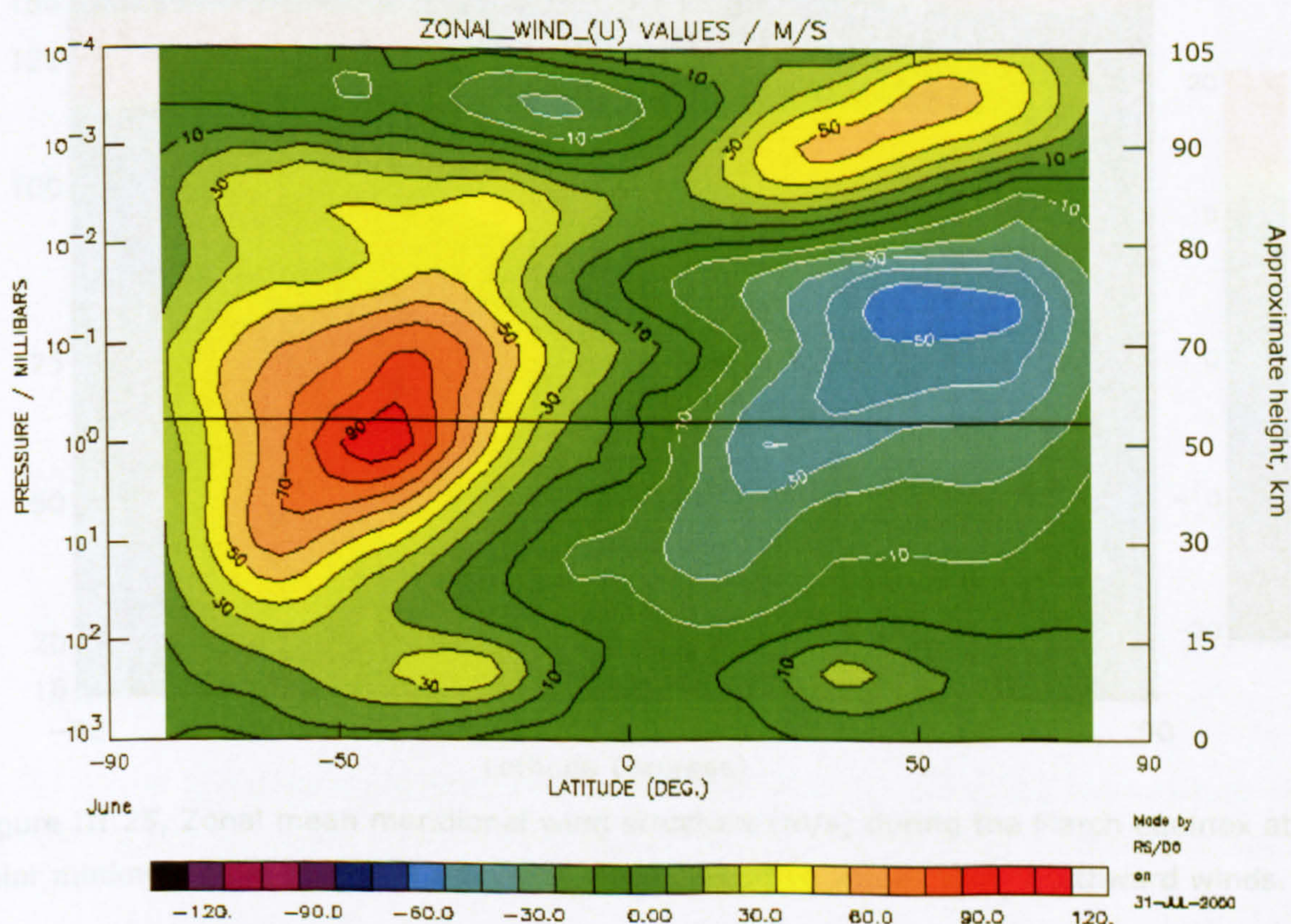


Figure III.23, Zonal mean zonal wind structure during the June solstice at solar minimum from the UARS Reference Atmosphere Project. The boxed regions indicate where the data are derived mainly by interpolation or from climatology.

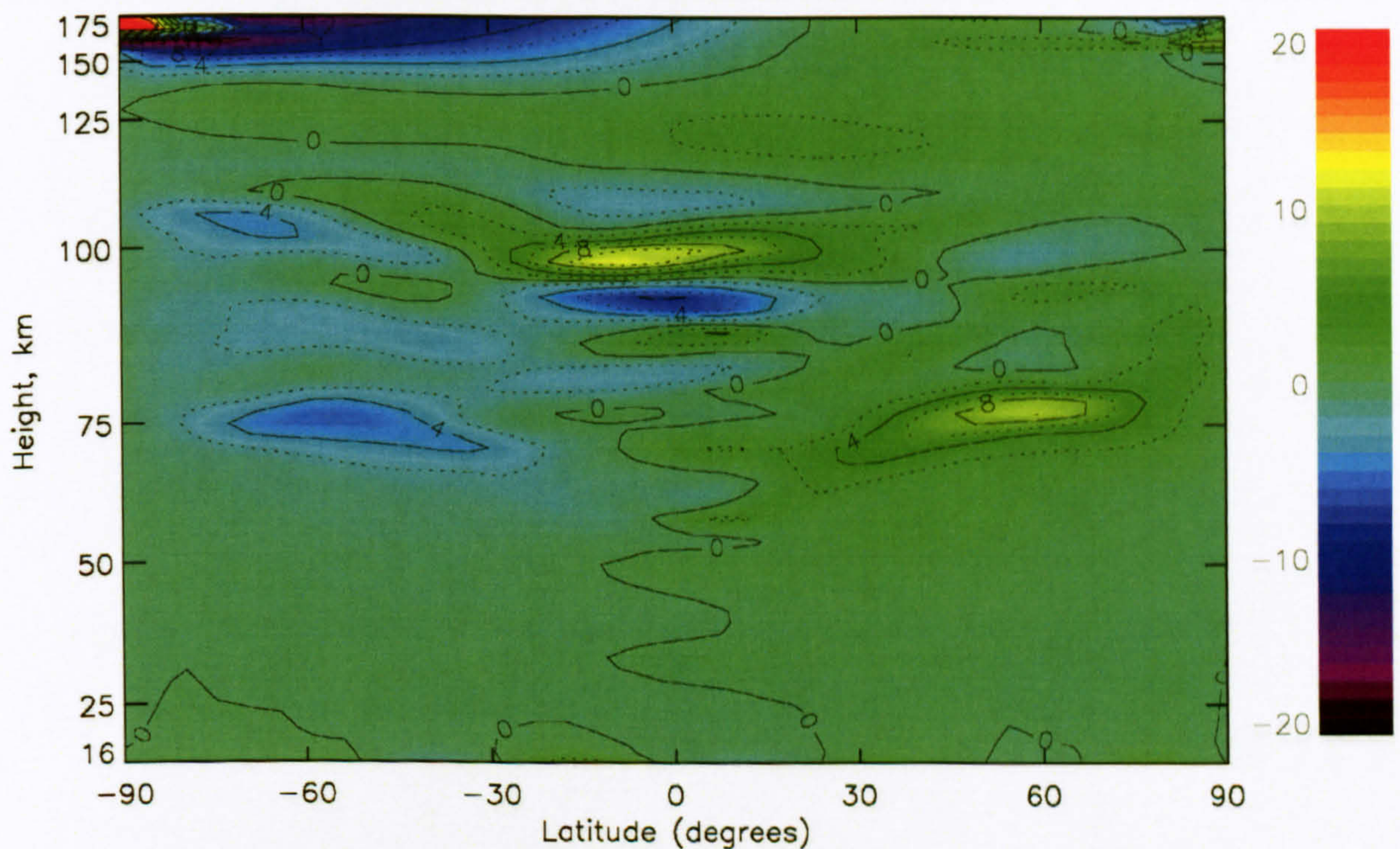


Figure III.24, Zonal mean meridional wind structure (m/s) during the March equinox at solar minimum calculated by the STEVE model. Positive values show northward winds.

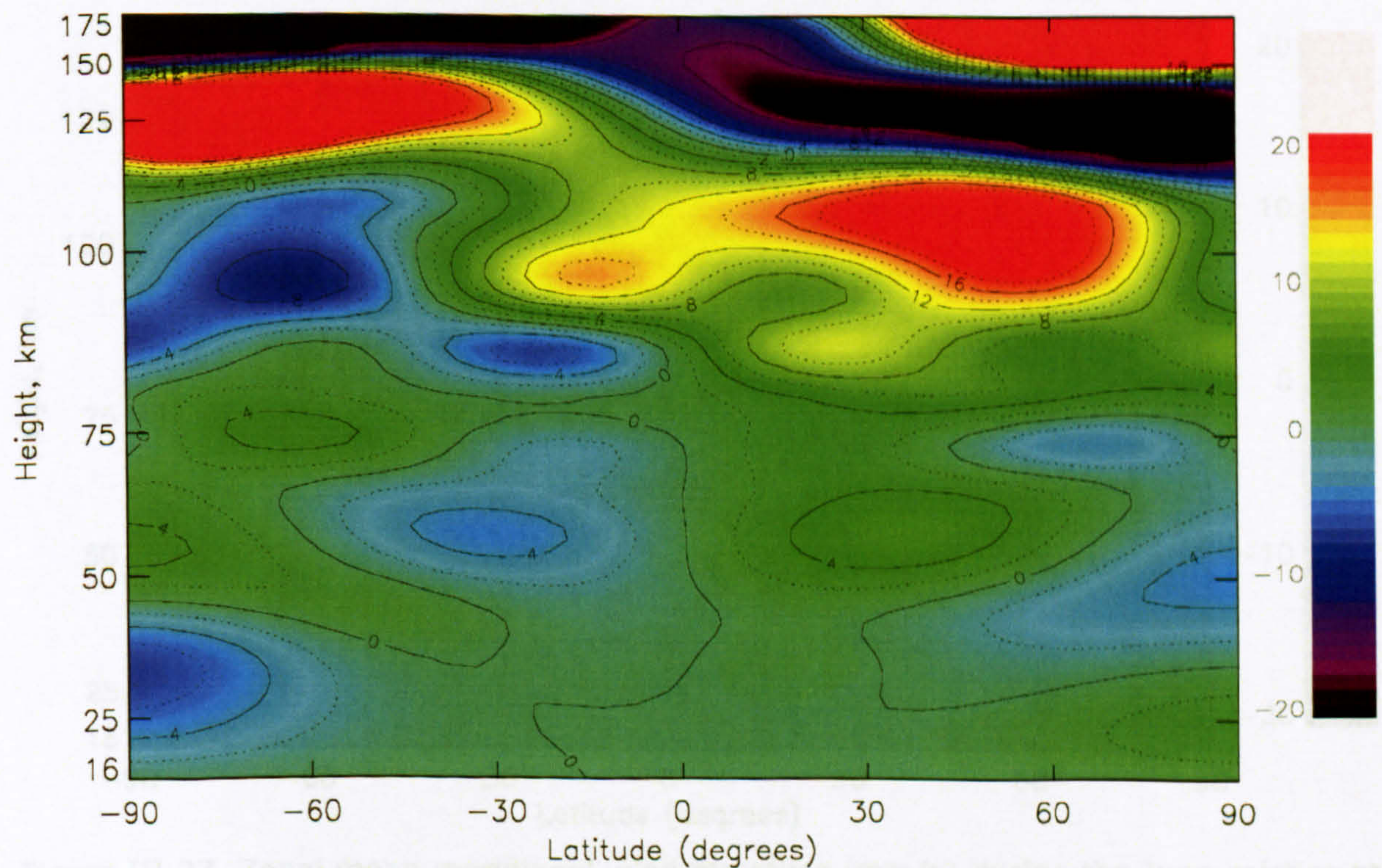


Figure III.25, Zonal mean meridional wind structure (m/s) during the March equinox at solar minimum from the HWM empirical model. Positive values show northward winds.

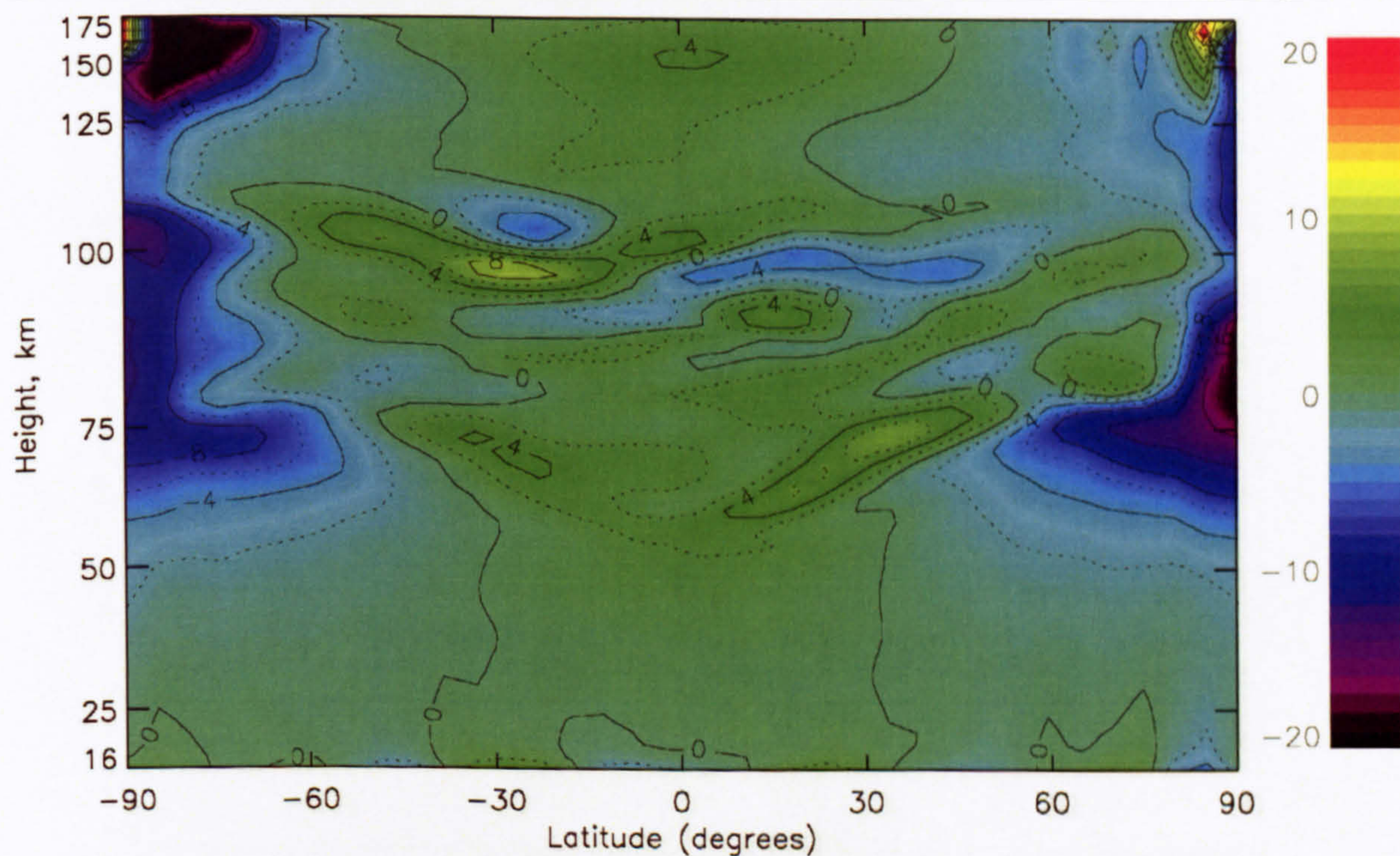


Figure III.26, Zonal mean vertical wind structure (mm/s) during the March equinox at solar minimum calculated by the STEVE model. Positive values show northward winds.

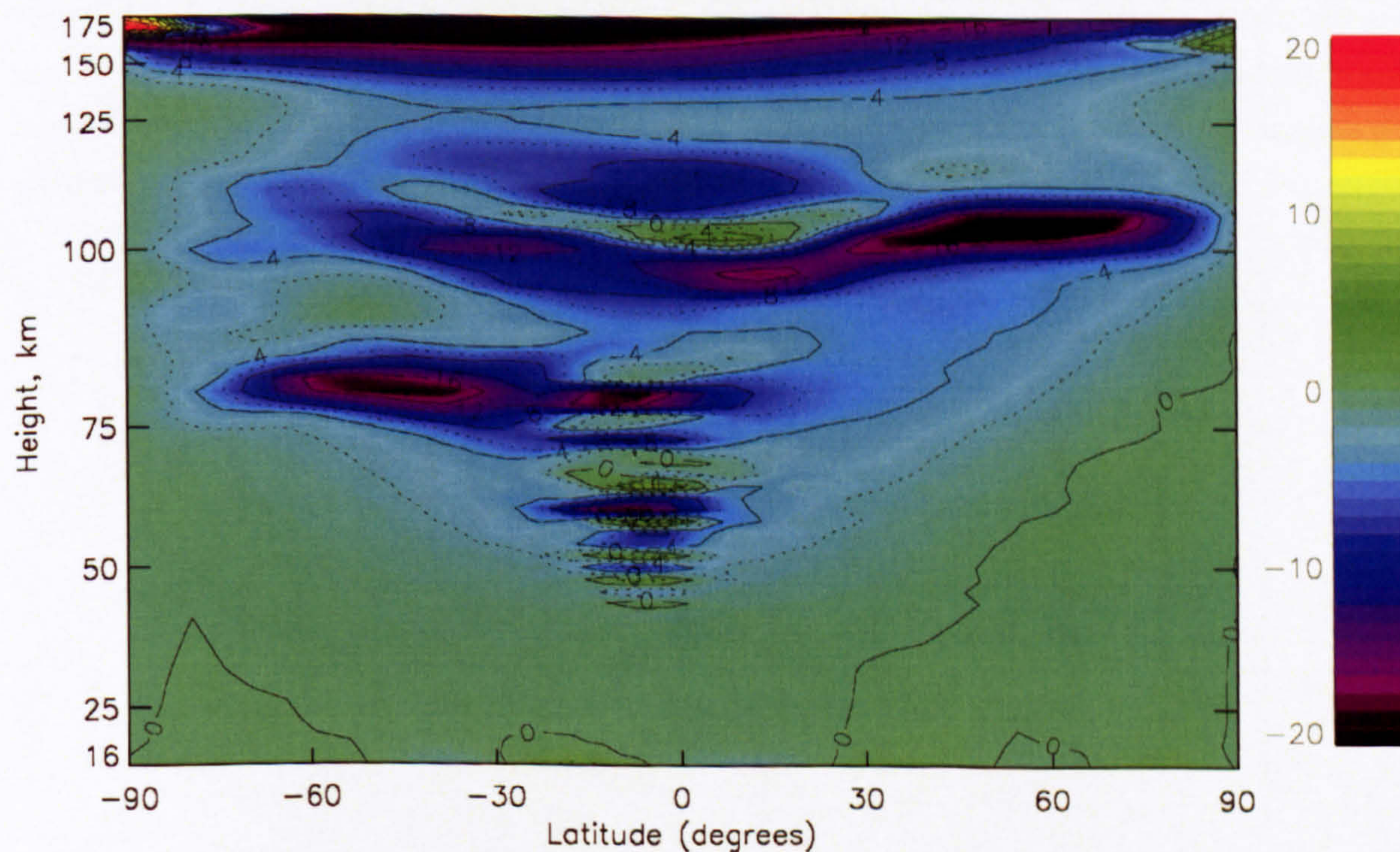


Figure III.27, Zonal mean meridional wind structure (mm/s) during the June solstice at solar minimum calculated by the STEVE model. Positive values show northward winds.

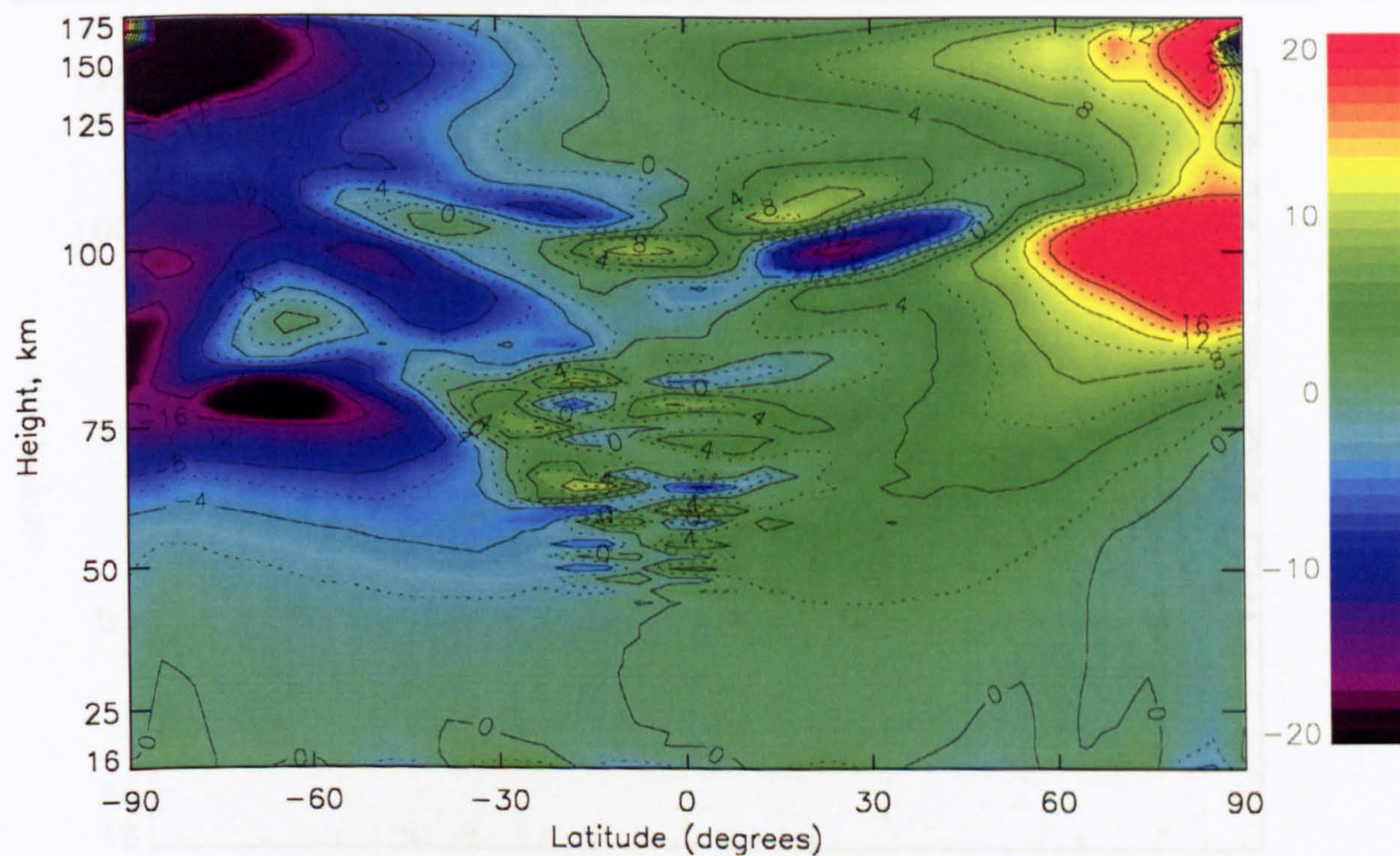


Figure III.28, Zonal mean vertical wind structure (mm/s) during the June solstice at solar minimum calculated by the STEVE model. Positive values show northward winds.

Figure III.29, Zonal mean, daily averaged meridional circulation pattern produced by the STEVE model for the March equinox at solar minimum. The vertical winds have been scaled up by a factor of 500 compared with the horizontal winds. The maximum vertical wind is 20 mm/s and the maximum horizontal wind is 10 m/s.

Zonal Mean Meridional Circulation



Figure III.30, As figure III.29, but for the the June solstice at solar minimum.

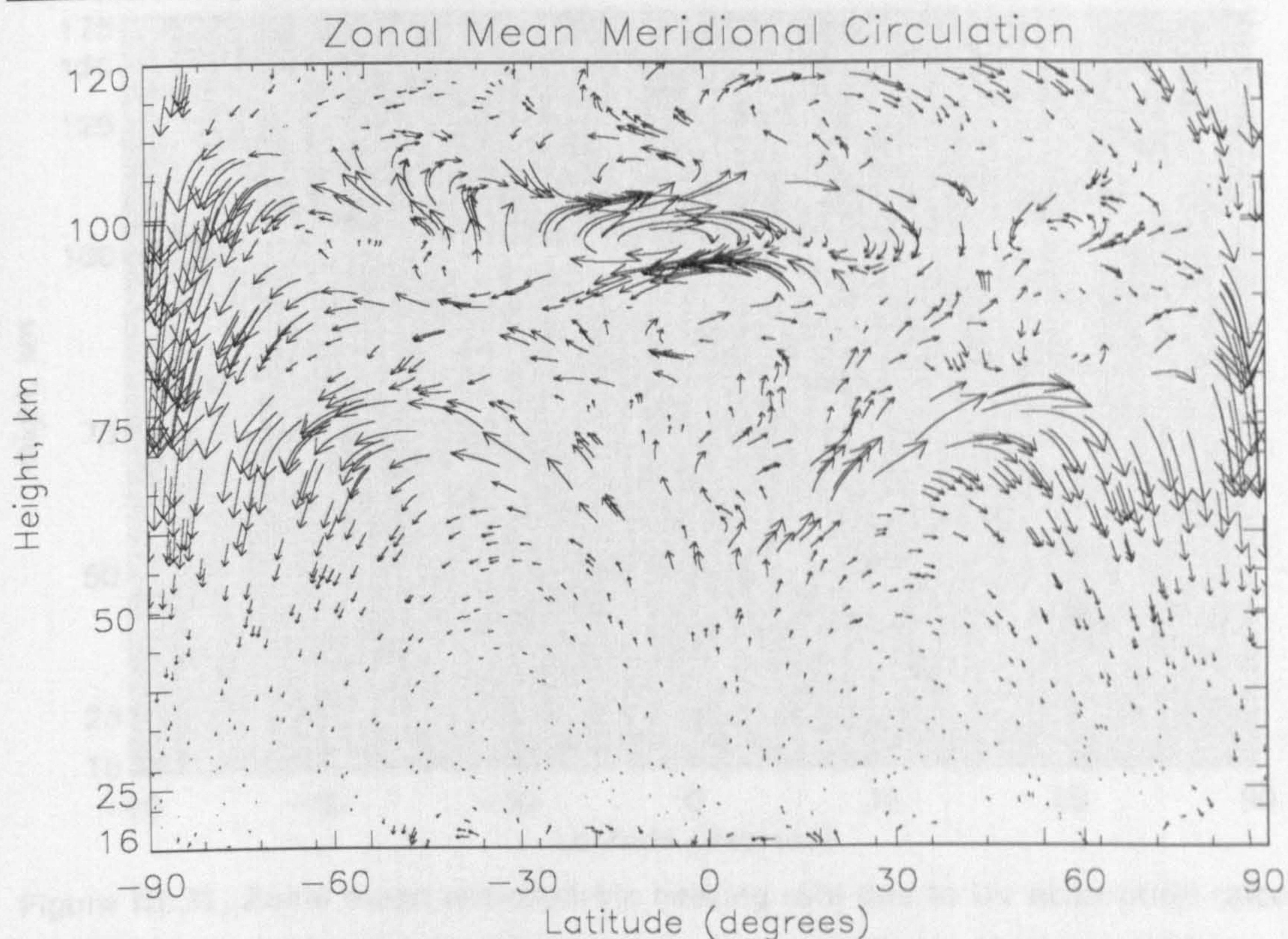


Figure III.29, Zonal mean, daily averaged meridional circulation pattern produced by the STEVE model for the March equinox at solar minimum. The vertical winds have been scaled up by a factor of 500 compared with the horizontal winds. The maximum vertical wind is 20 mm/s and the maximum horizontal wind is 10 m/s.

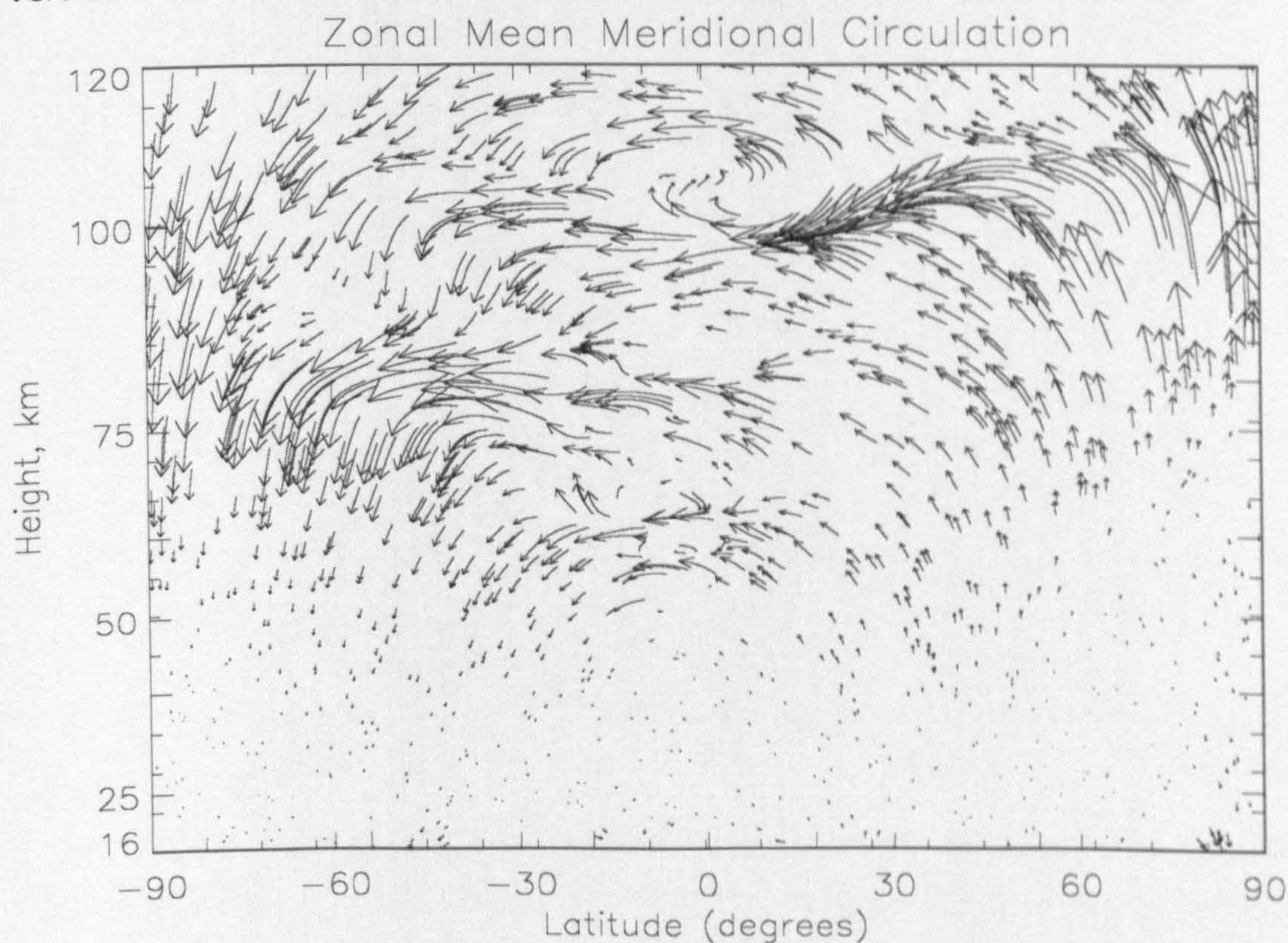


Figure III.30, As figure III.29, but for the the June solstice at solar minimum.

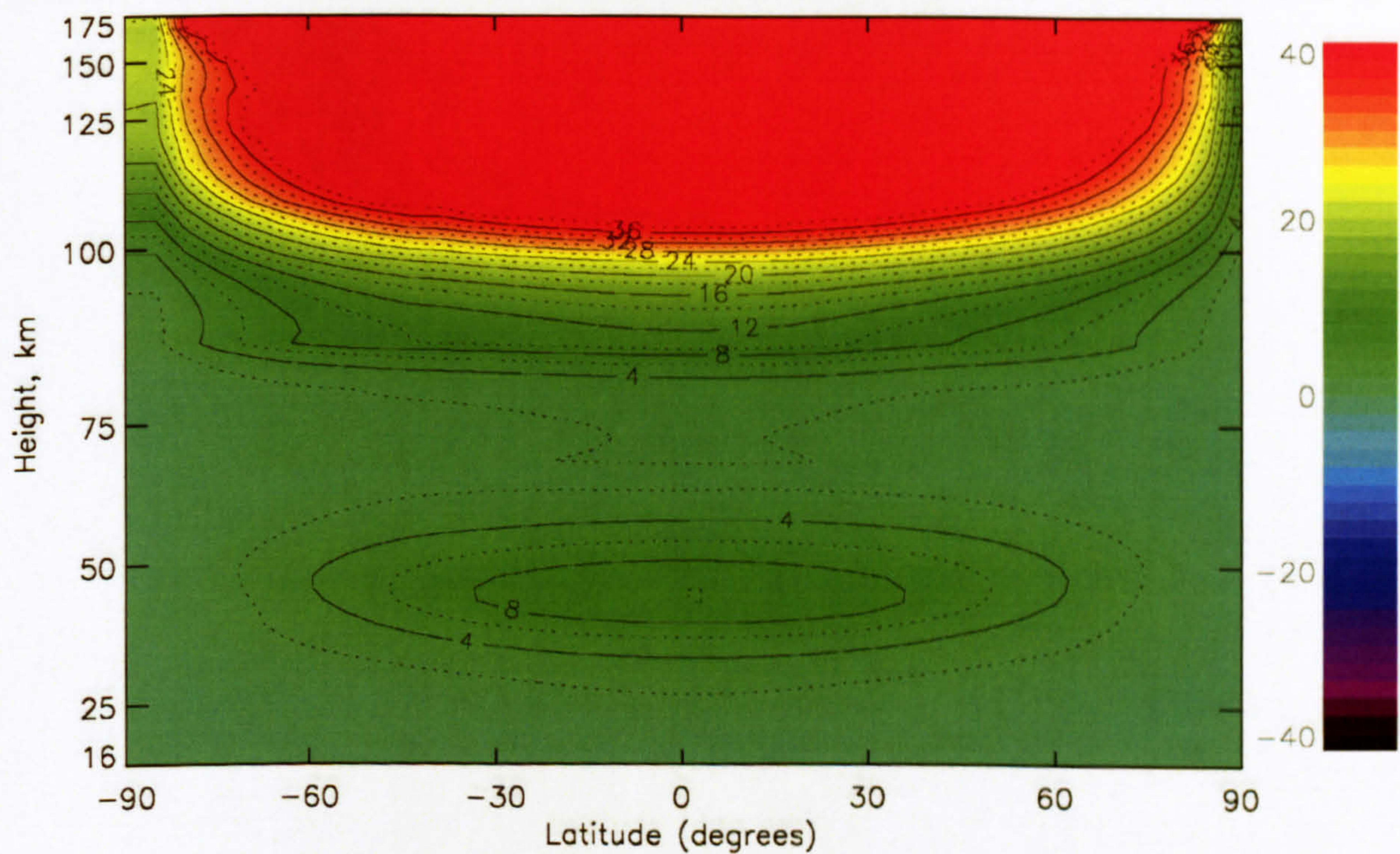


Figure III.31, Zonal mean atmospheric heating rate due to UV absorption calculated by the STEVE model for the March equinox at solar minimum. Units are K/day.

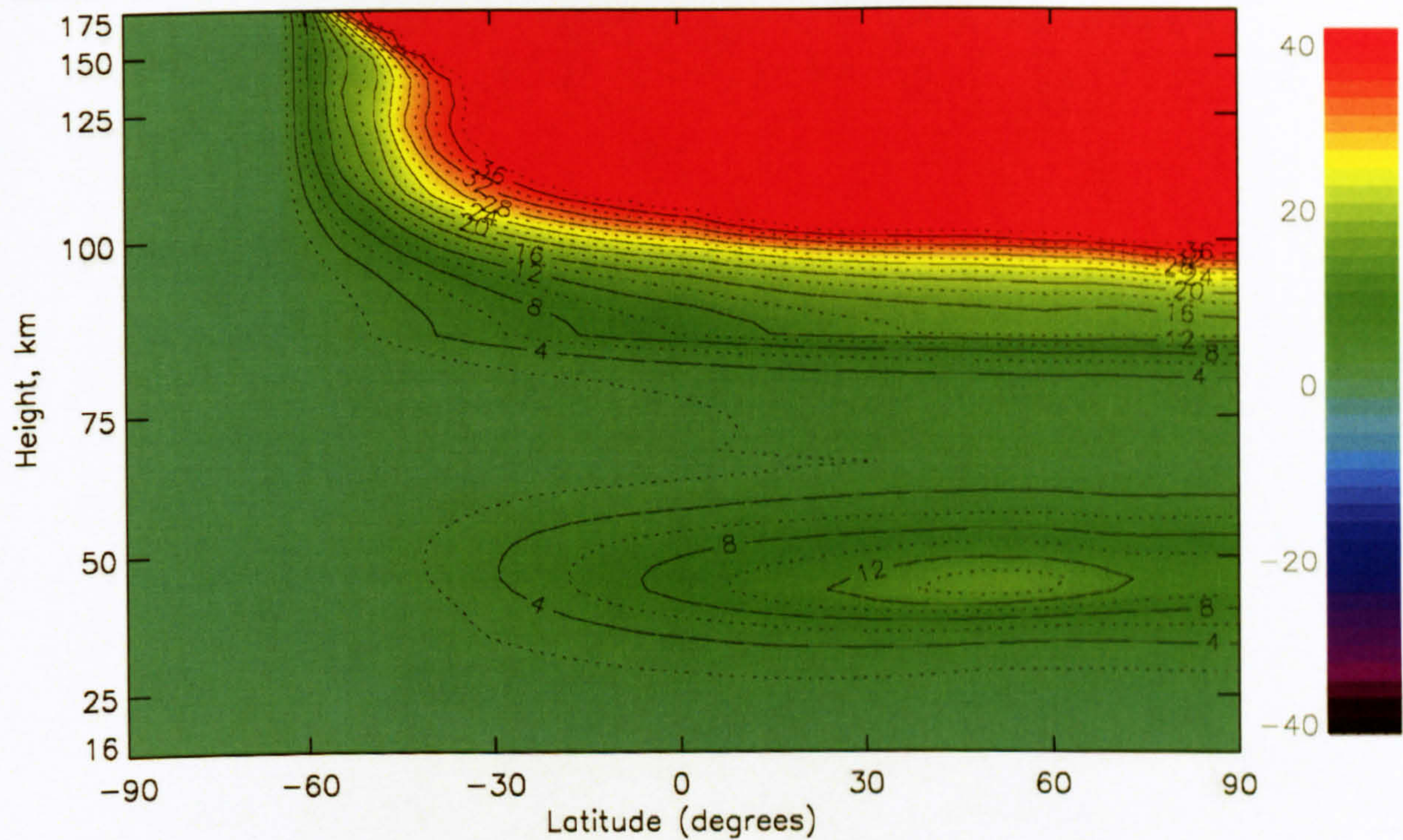


Figure III.32, Zonal mean atmospheric heating rate due to UV absorption calculated by the STEVE model for the June solstice at solar minimum. Units are K/day.

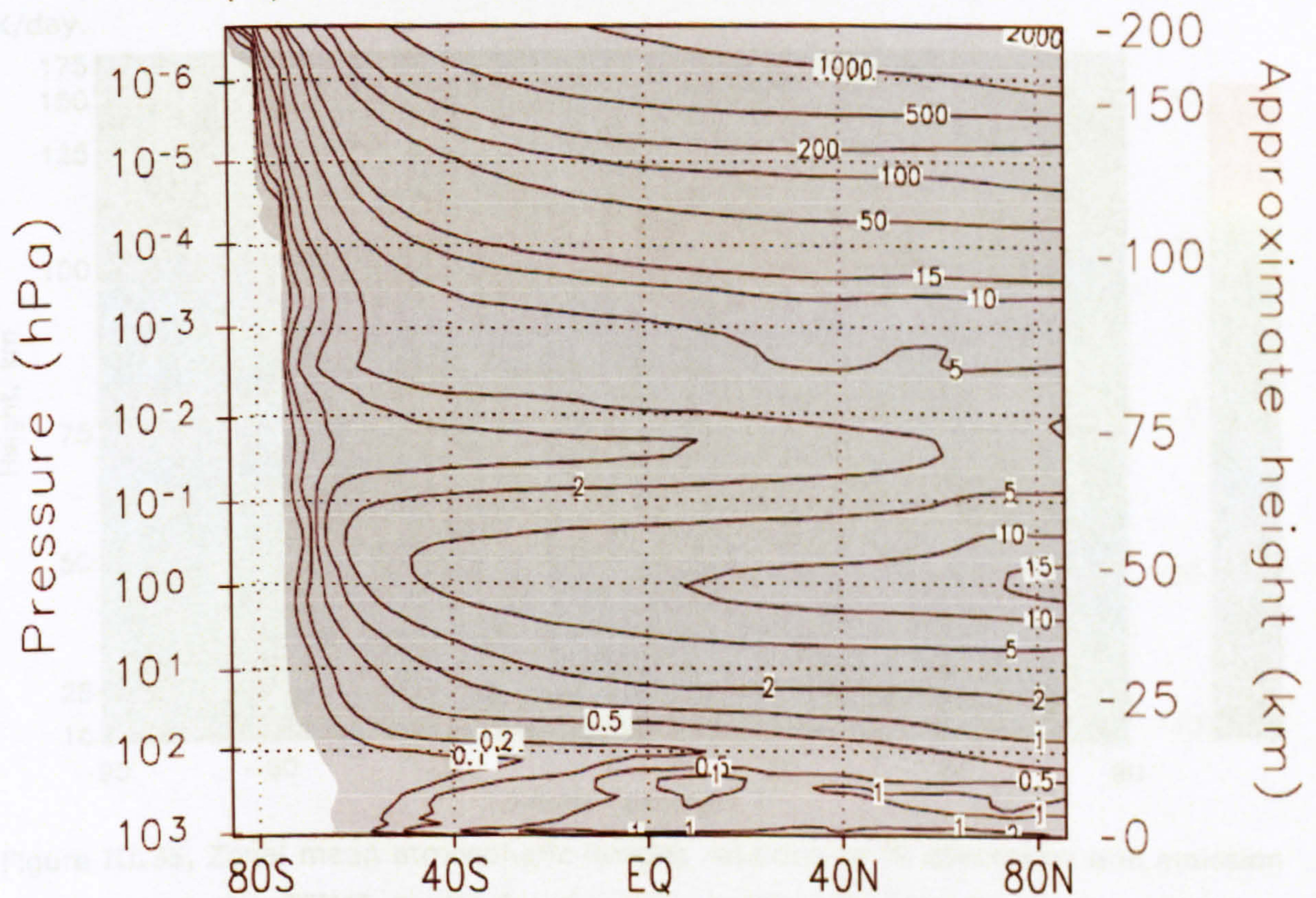


Figure III.33, Zonal mean heating rate due to absorption of solar UV calculated by the CMAM model for the June solstice at solar minimum (Fomichev *et al.*, 2002). Units are K/day.

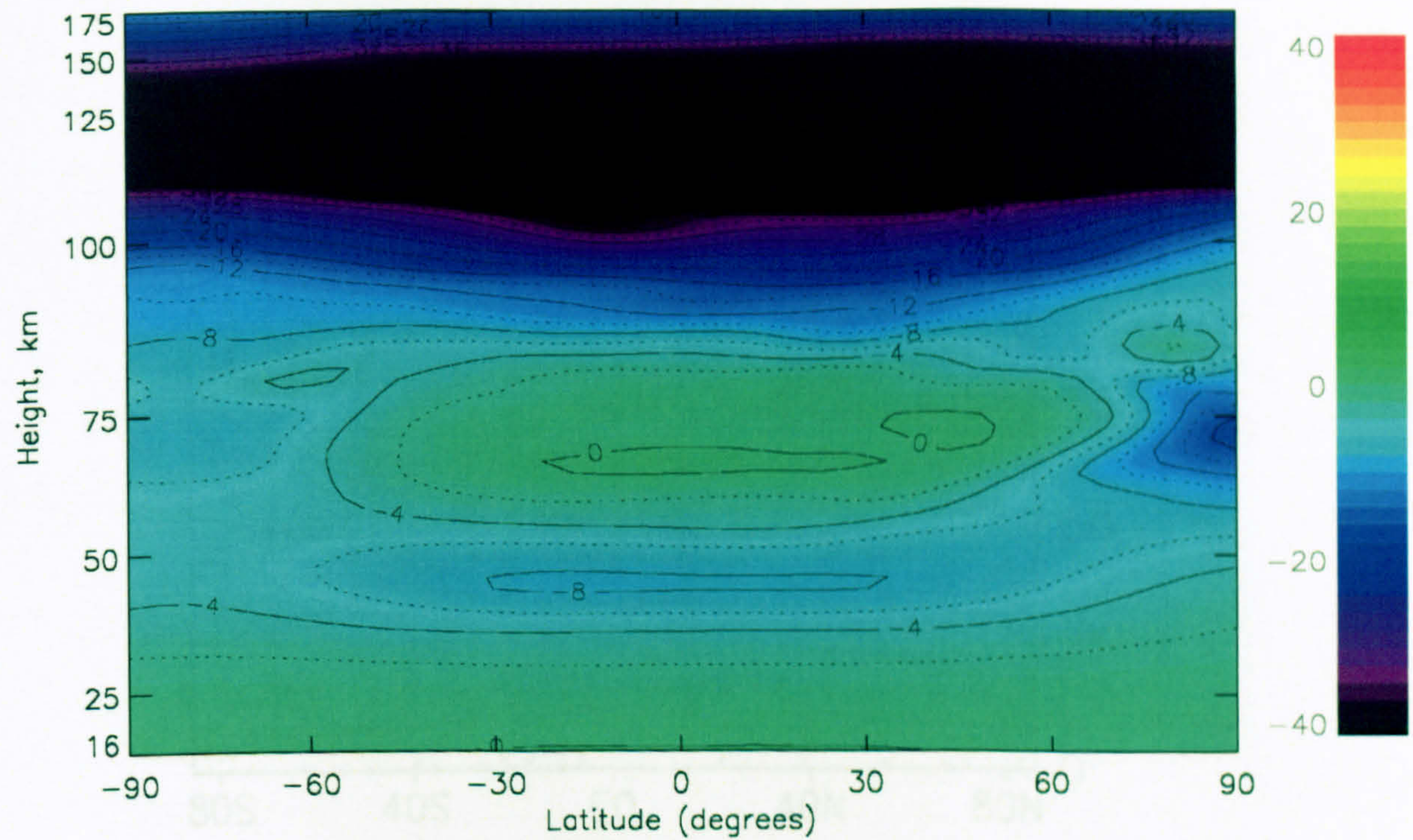


Figure III.34, Zonal mean atmospheric heating rate due to IR absorption and emission calculated by the STEVE model for the March equinox at solar minimum. Units are K/day.

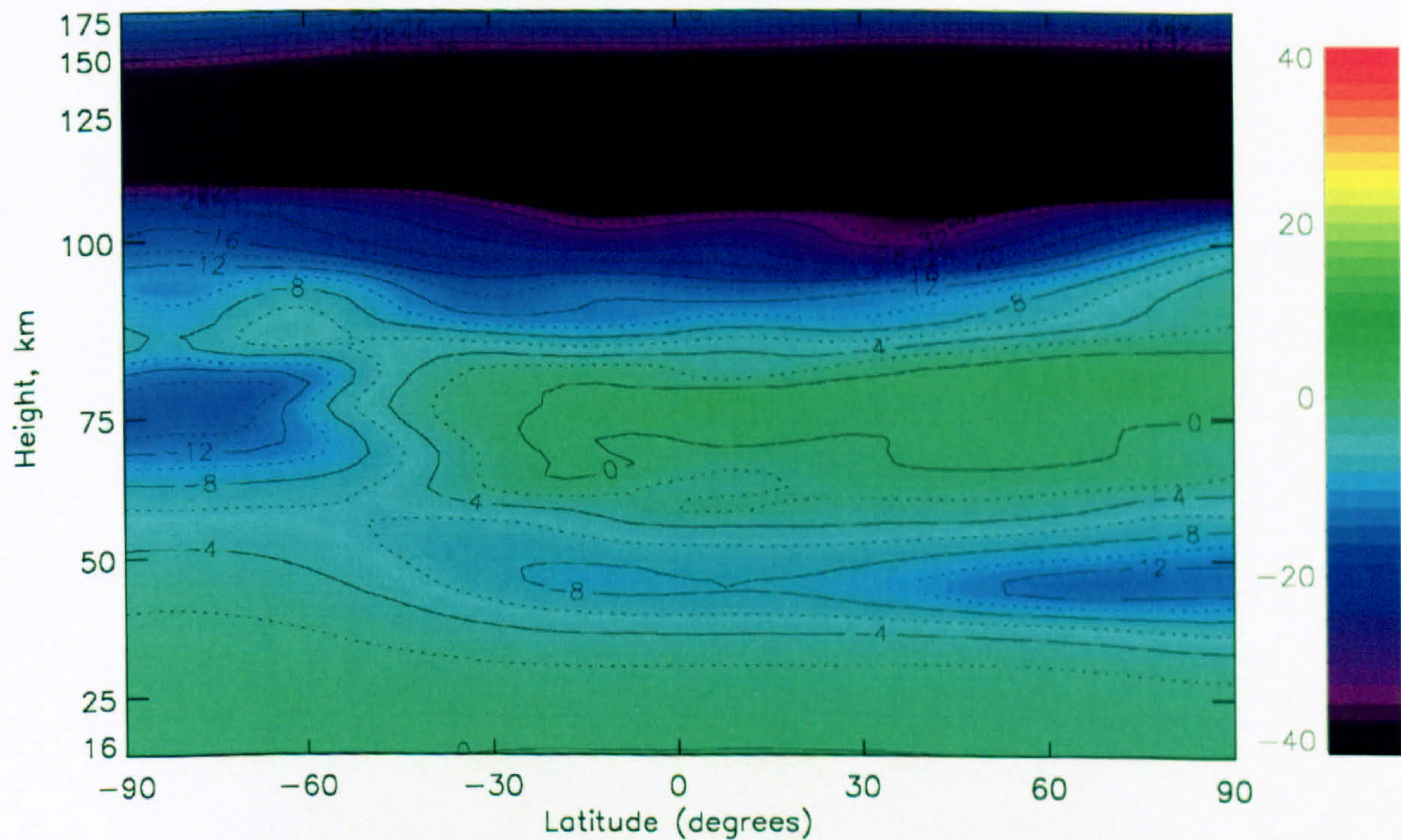


Figure III.35, Zonal mean atmospheric heating rate due to IR absorption and emission calculated by the STEVE model for the June solstice at solar minimum. Units are K/day.

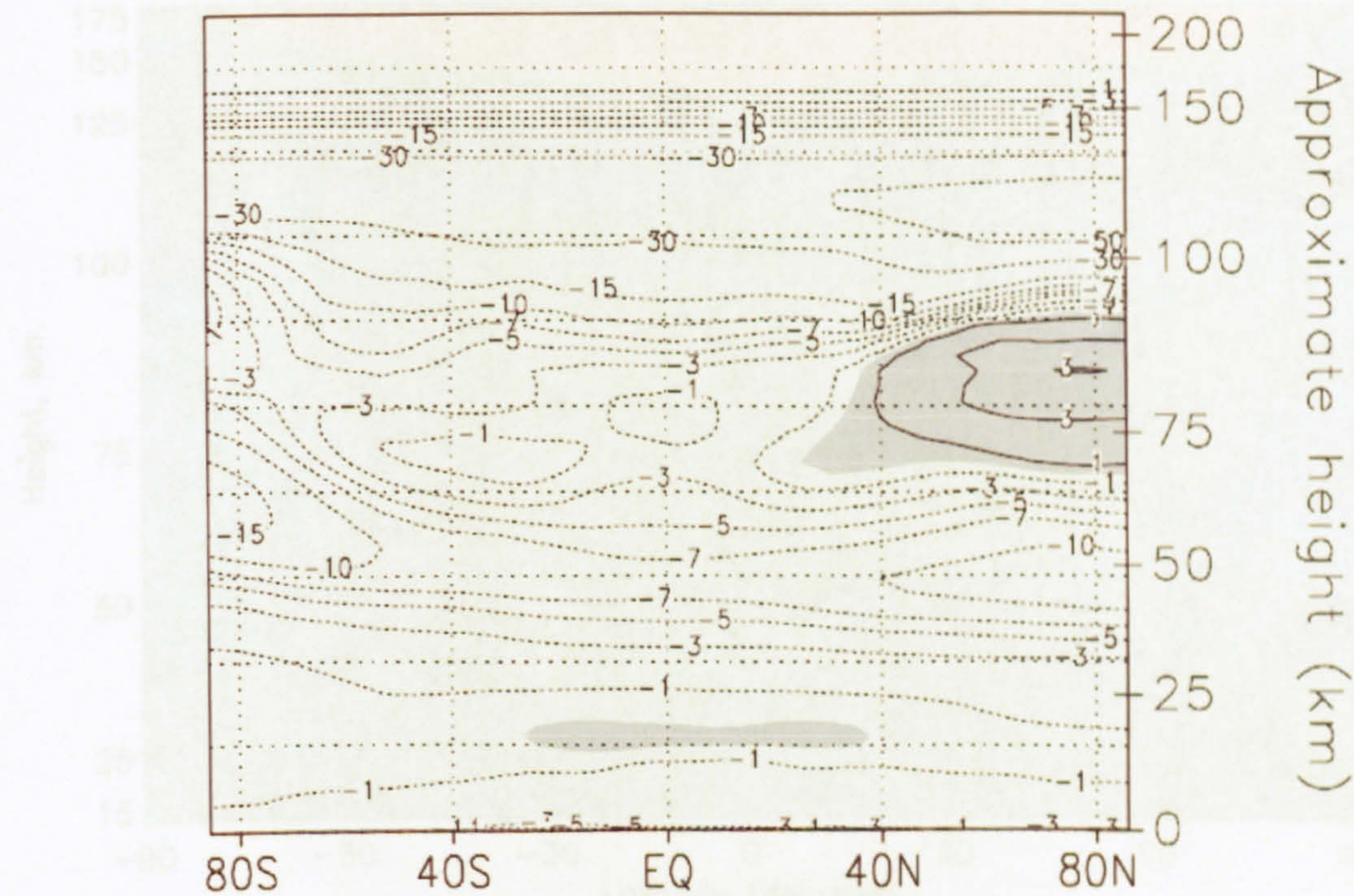


Figure III.36, Zonal mean cooling rate due to IR absorption and emission calculated by the CMAM model for the June solstice at solar minimum (Fomichev *et al.*, 2002). Units are K/day.

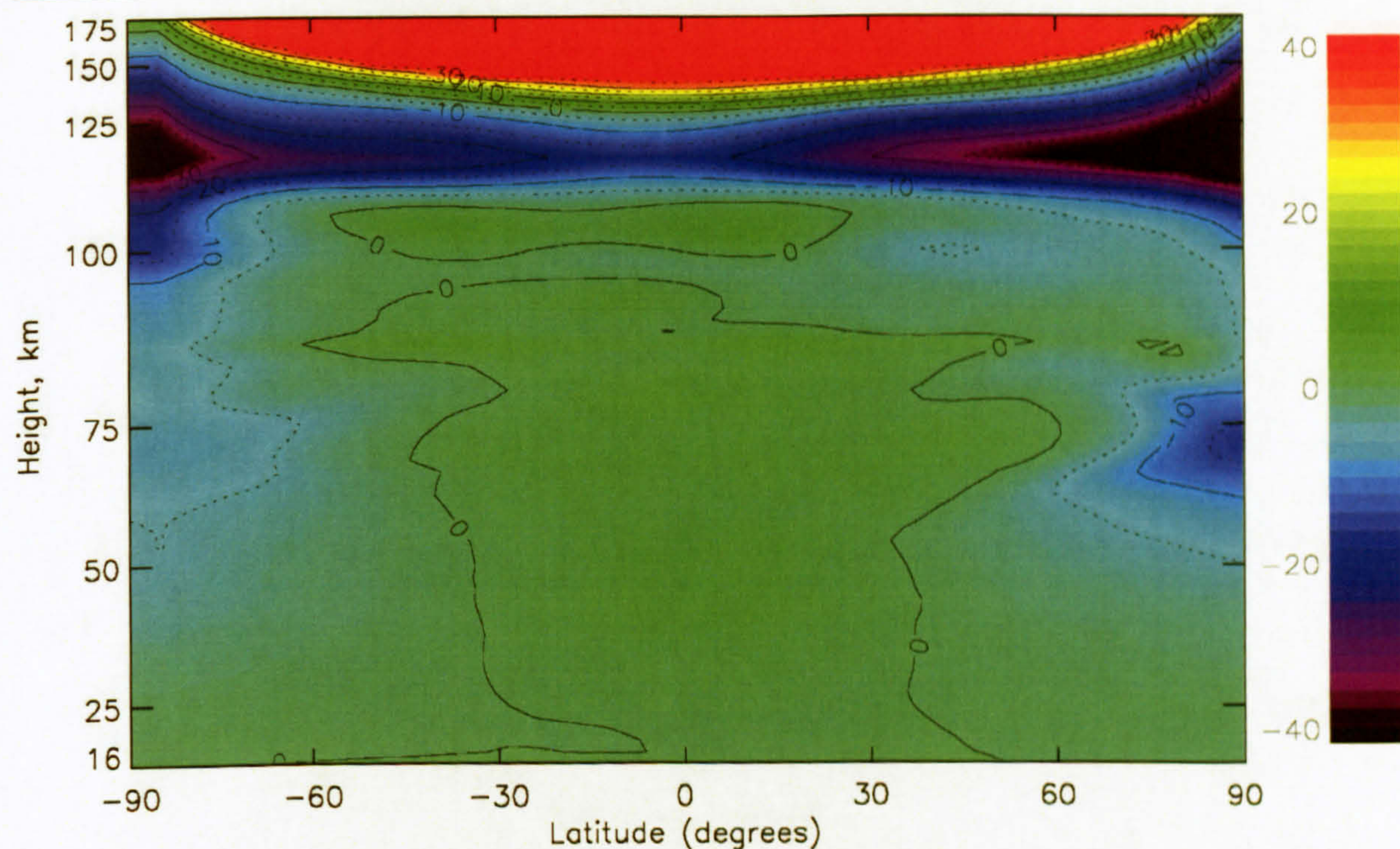


Figure III.37, Zonal mean net diabatic heating rate due to IR absorption and emission UV absorption calculated by the STEVE model for the March equinox at solar minimum. Units are K/day.

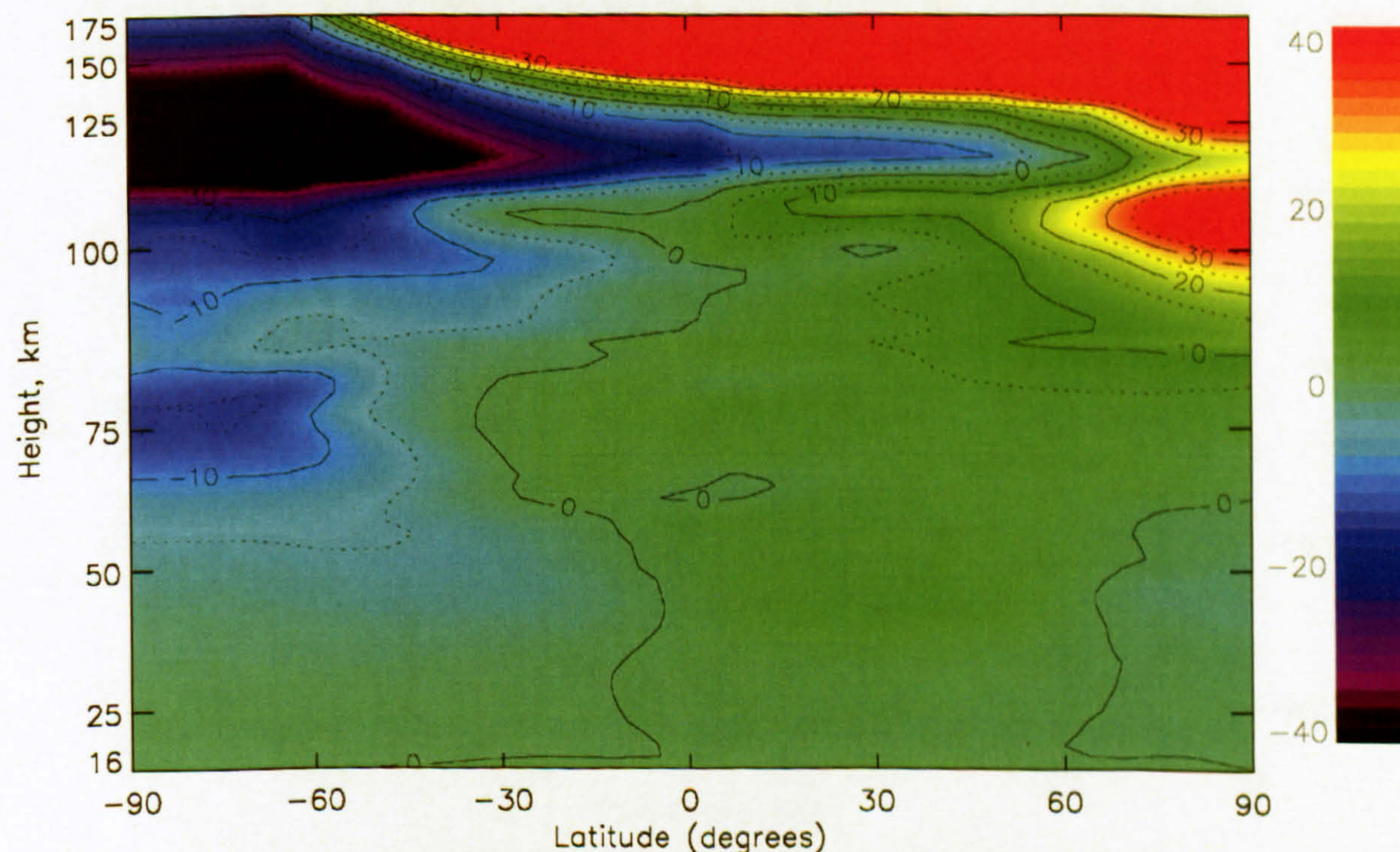


Figure III.38, Zonal mean net diabatic heating rate due to IR absorption and emission UV absorption calculated by the STEVE model for the June solstice at solar minimum. Units are K/day.

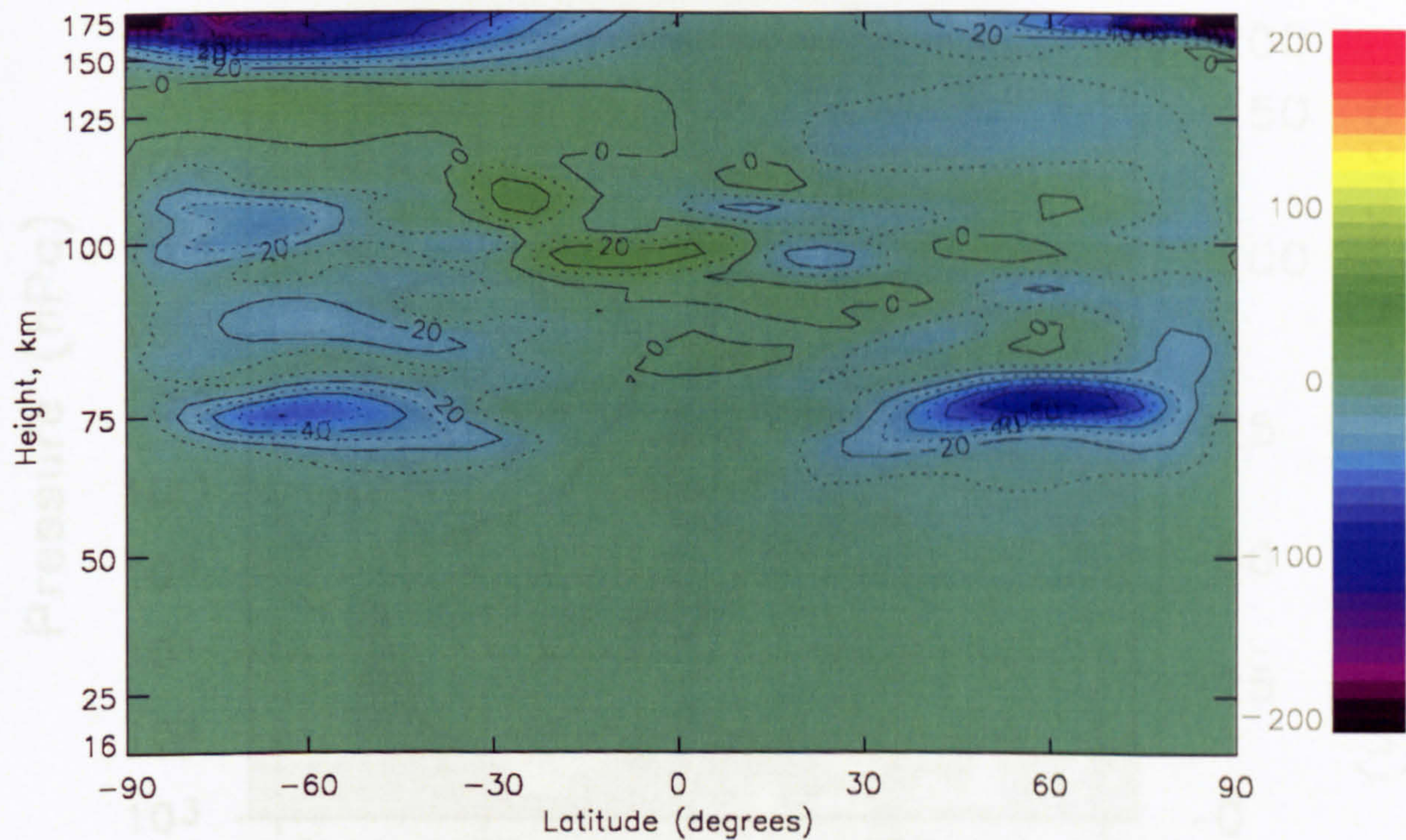


Figure III.39 Zonal mean zonal gravity-wave drag calculated by the STEVE model for the March equinox during solar minimum. Positive values show eastward drag. Units are m/s/day.

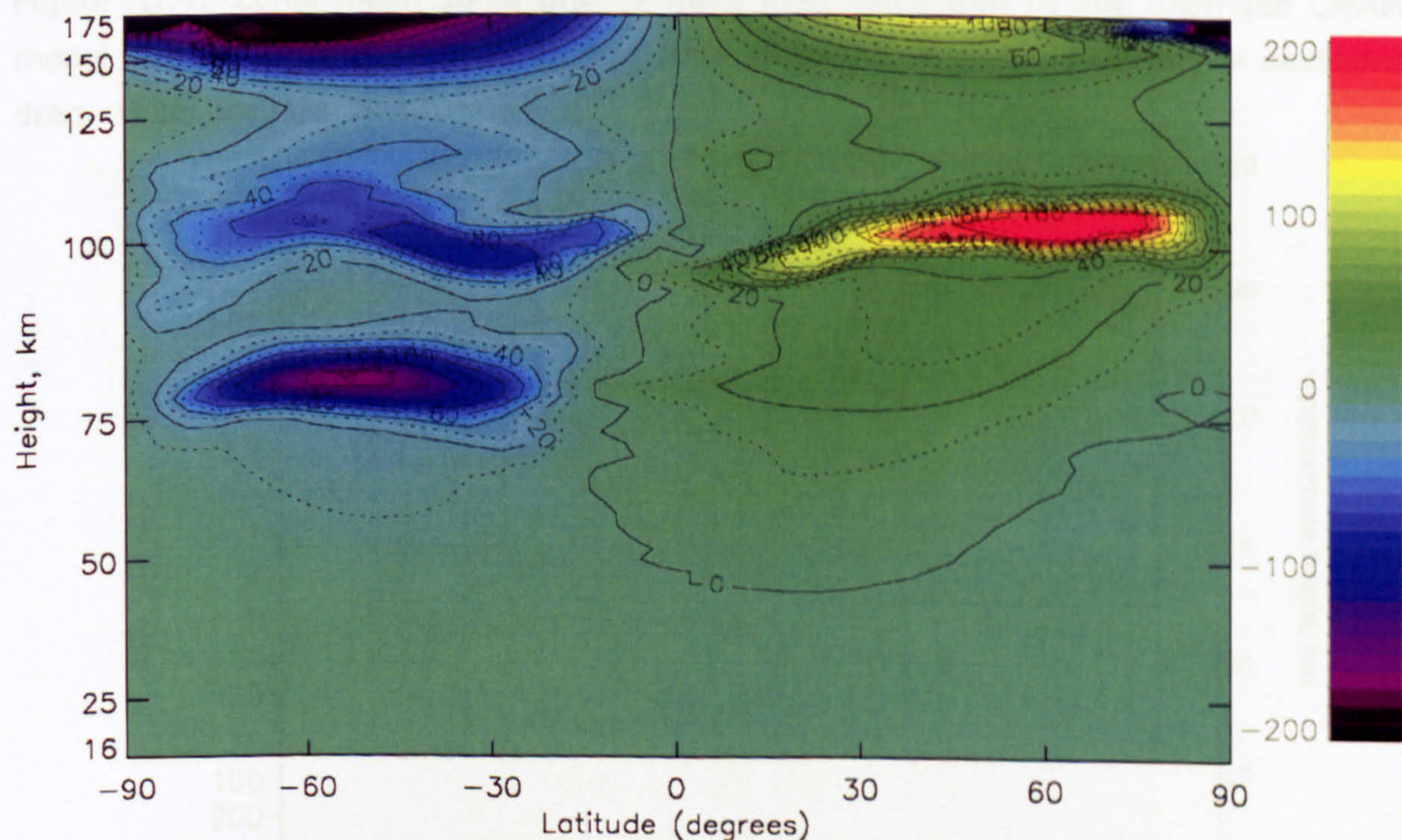


Figure III.40 Zonal mean zonal gravity-wave drag calculated by the STEVE model for the June solstice during solar minimum. Positive values show eastward drag. Units are m/s/day.

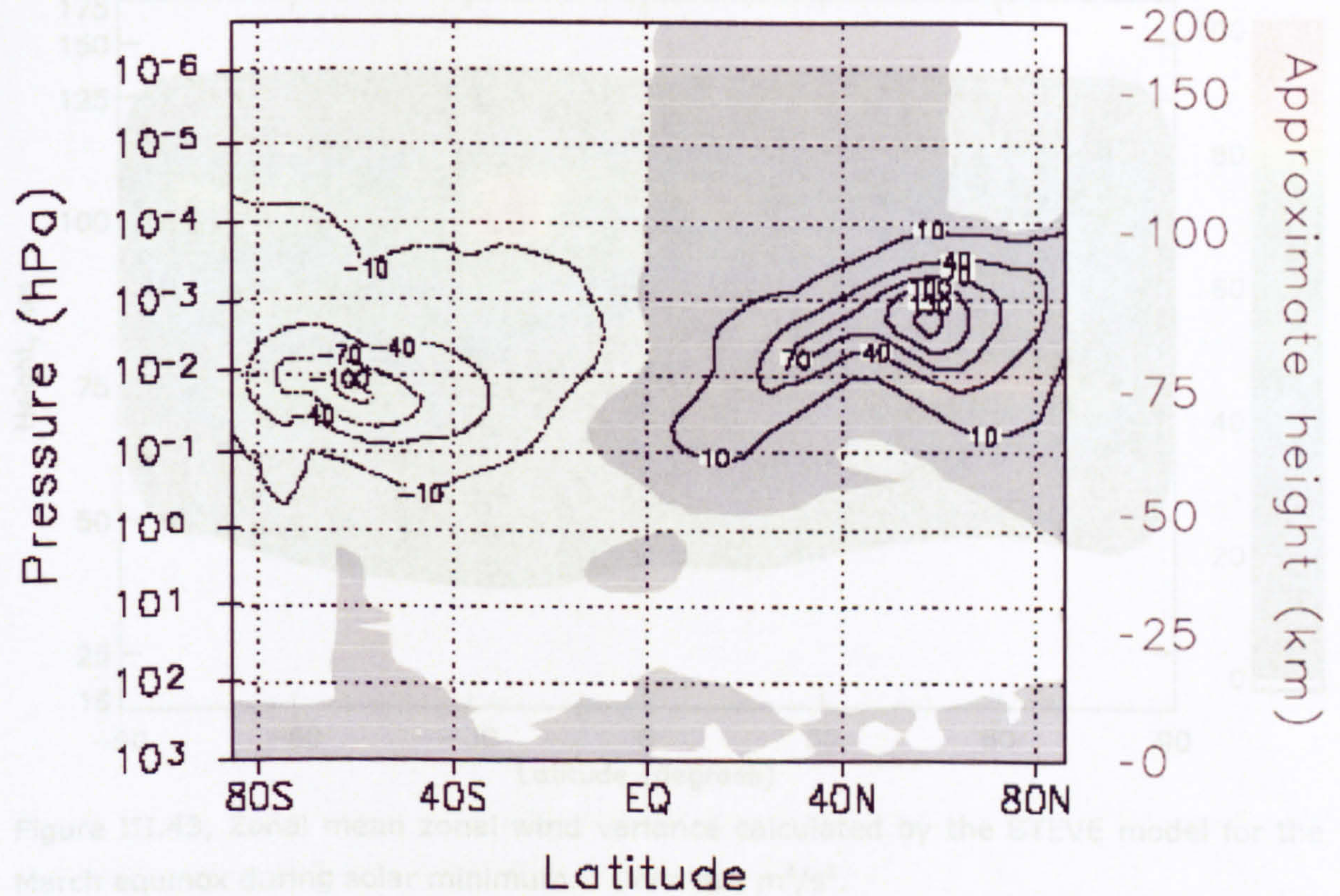


Figure III.41 Zonal mean zonal gravity-wave drag calculated by the extended CMAM model (Fomichev *et al.*, 2002) for the June solstice. Positive values show eastward drag. Units are m/s/day.

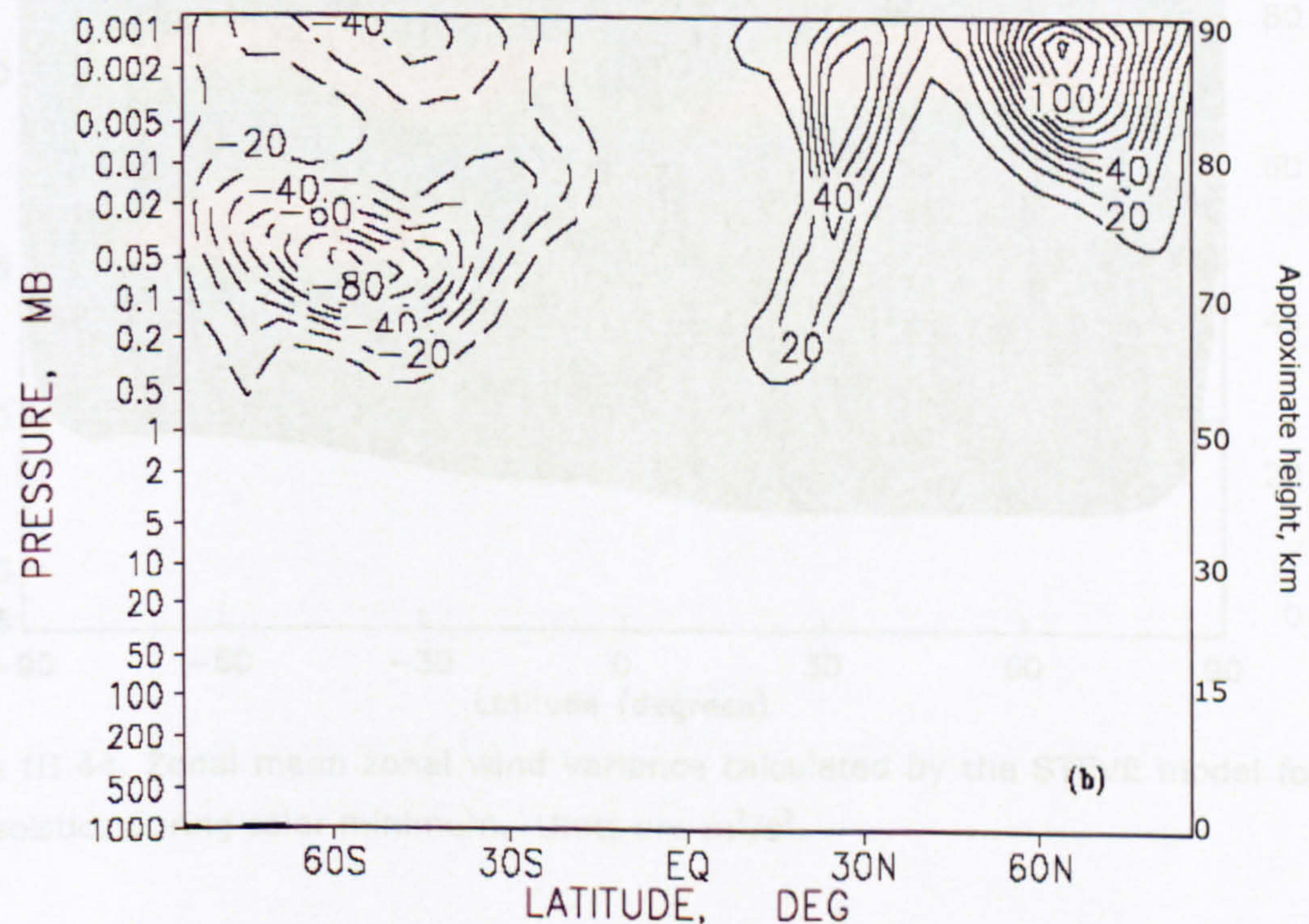


Figure III.42 Zonal mean zonal gravity-wave drag calculated by the CMAM model (Medvedev and Klaassen, 2003) for the June solstice. Positive values show eastward drag. Units are m/s/day.

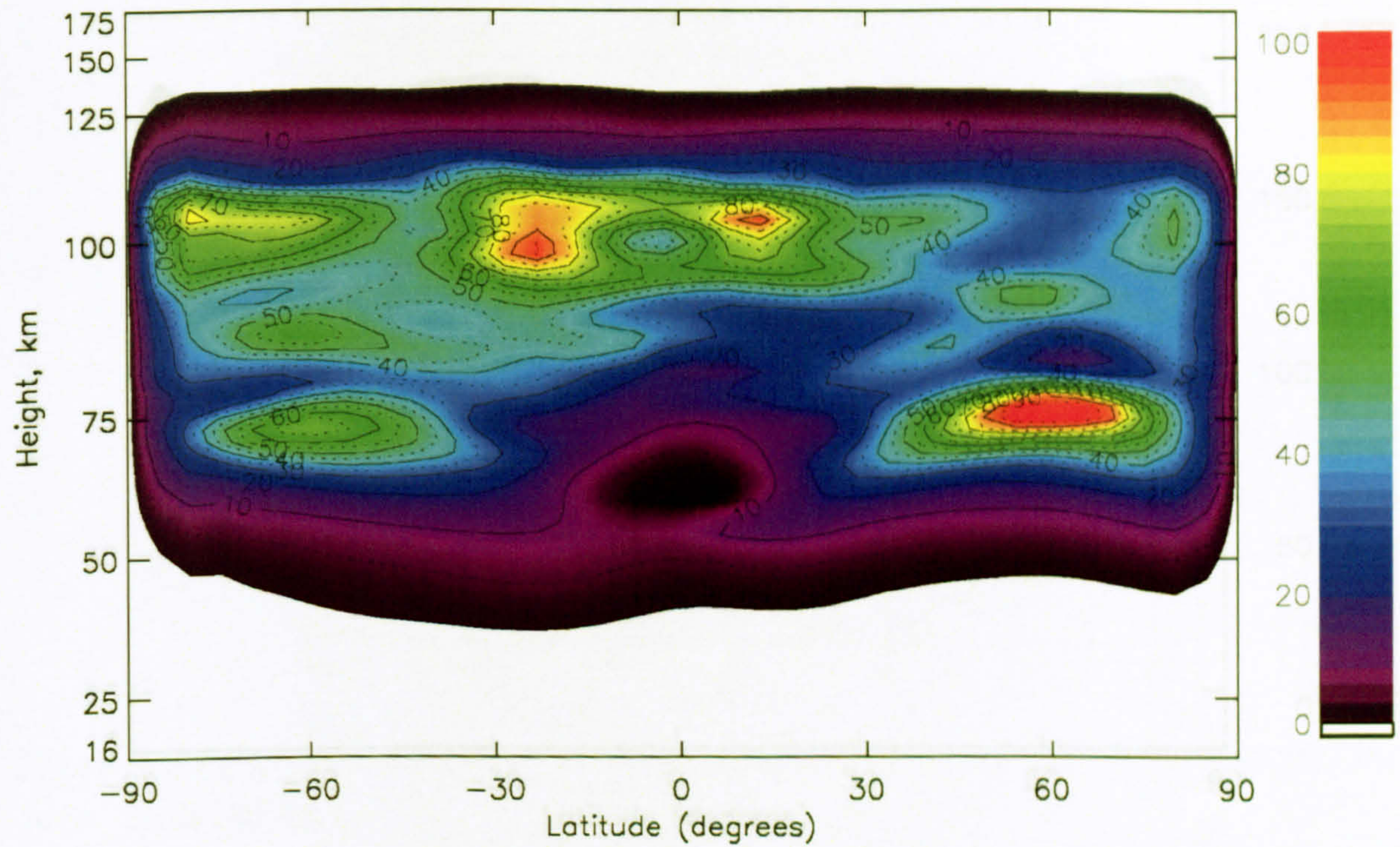


Figure III.43, Zonal mean zonal wind variance calculated by the STEVE model for the March equinox during solar minimum. Units are m^2/s^2 .

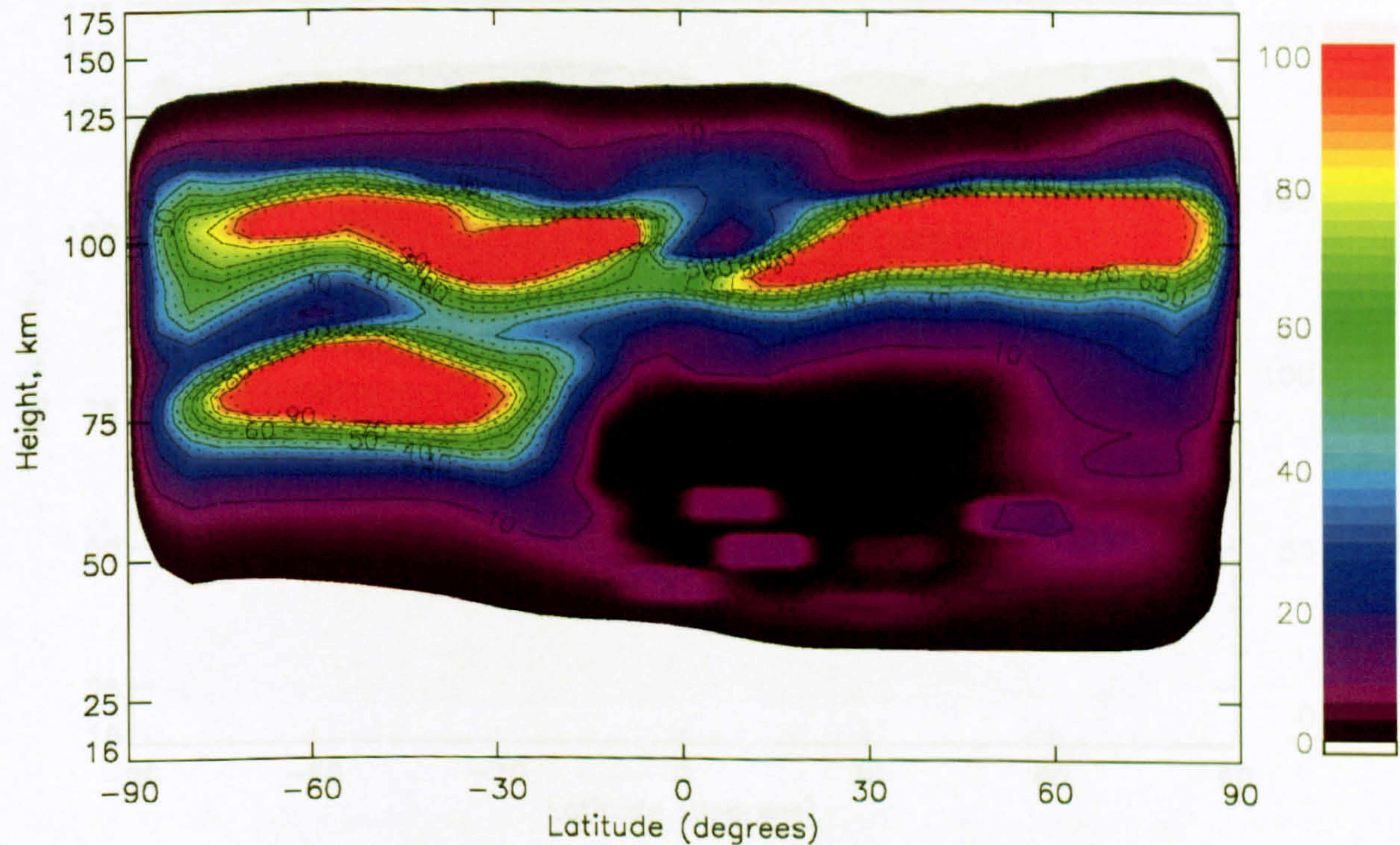


Figure III.44, Zonal mean zonal wind variance calculated by the STEVE model for the June solstice during solar minimum. Units are m^2/s^2 .

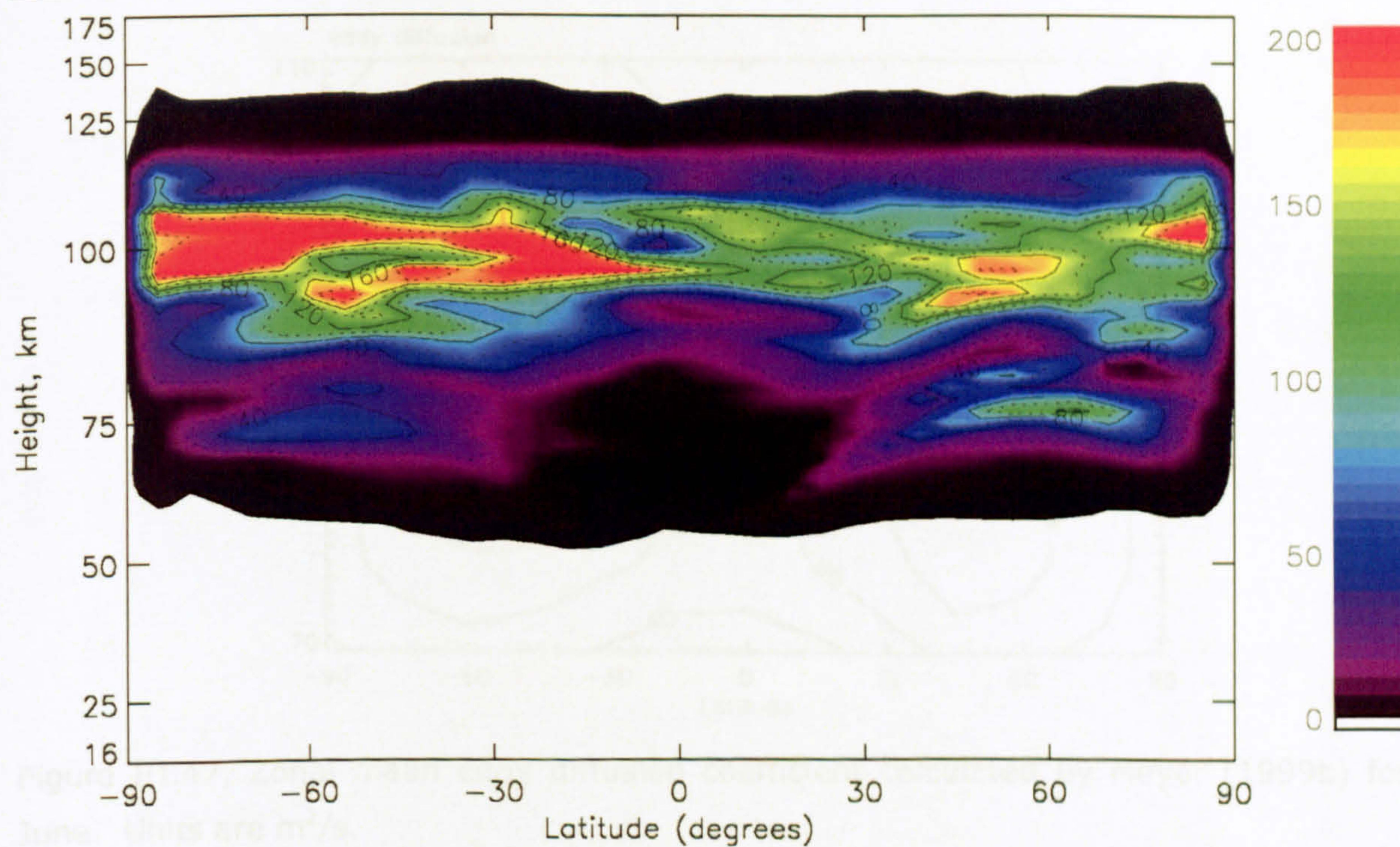


Figure III.45, Zonal mean eddy diffusion coefficient calculated by the STEVE model for the March equinox during solar minimum. Units are m^2/s .

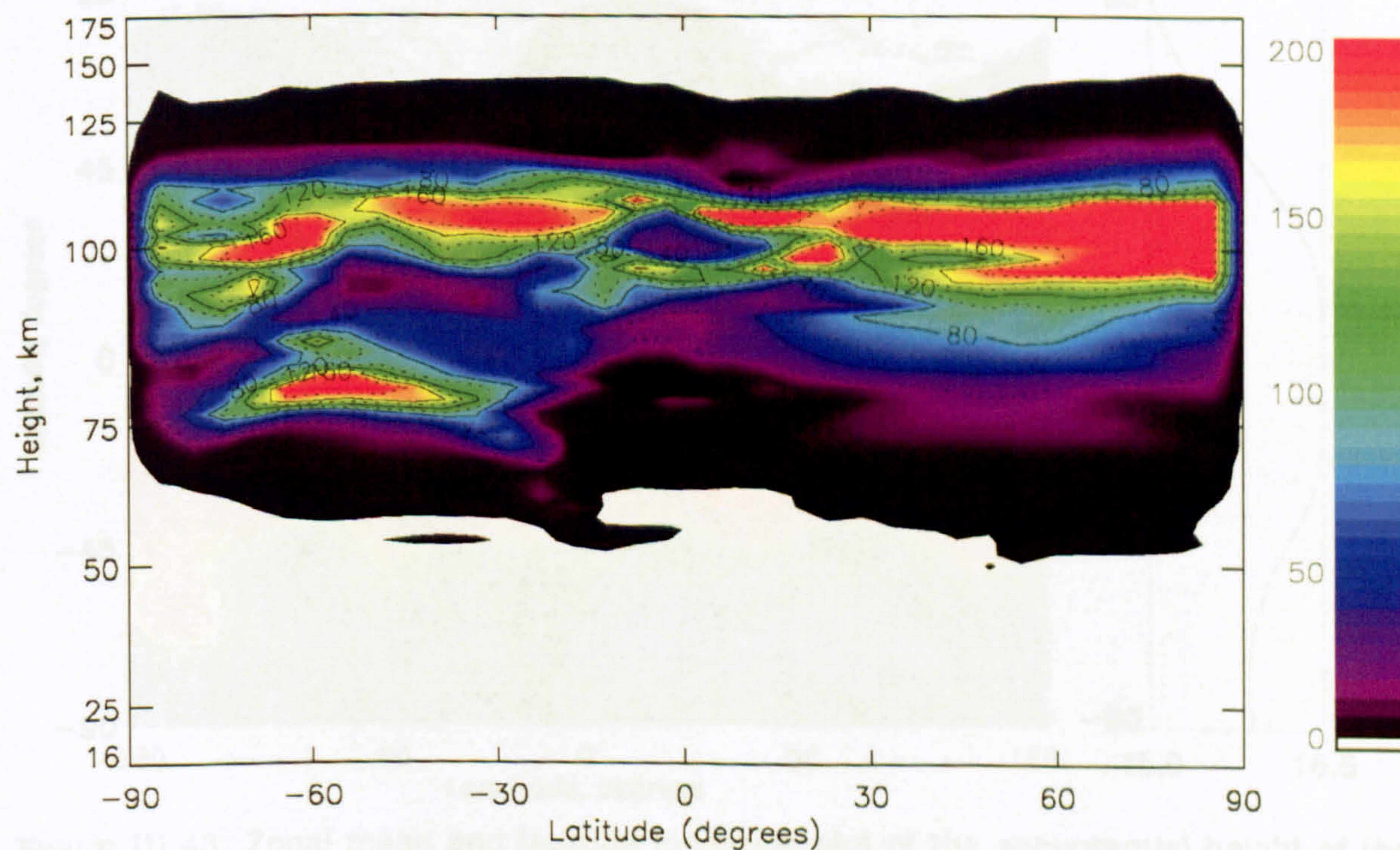


Figure III.46, Zonal mean eddy diffusion coefficient calculated by the STEVE model for the June solstice during solar minimum. Units are m^2/s .

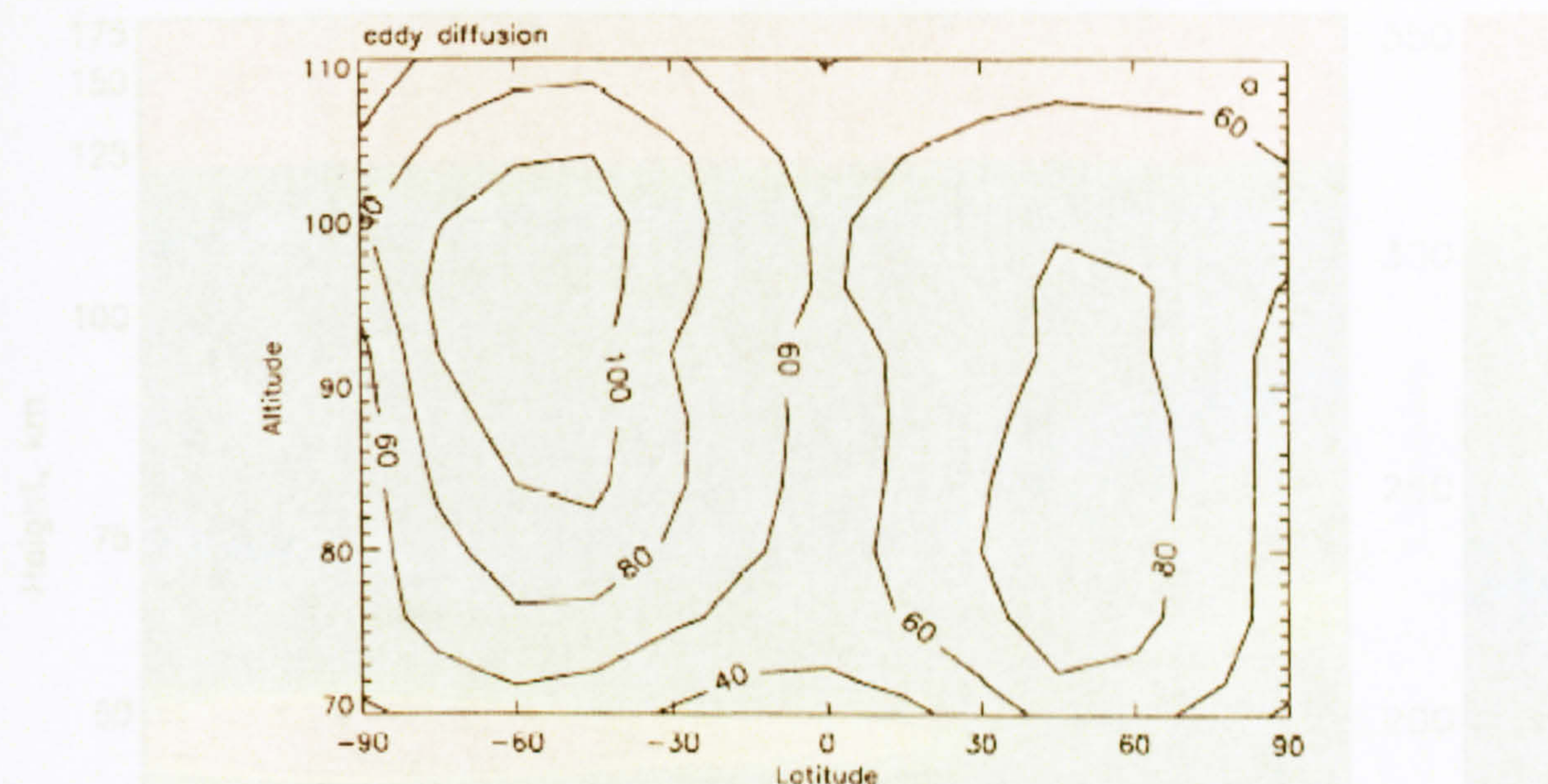


Figure III.47, Zonal mean eddy diffusion coefficient calculated by Meyer (1999b) for June. Units are m^2/s .

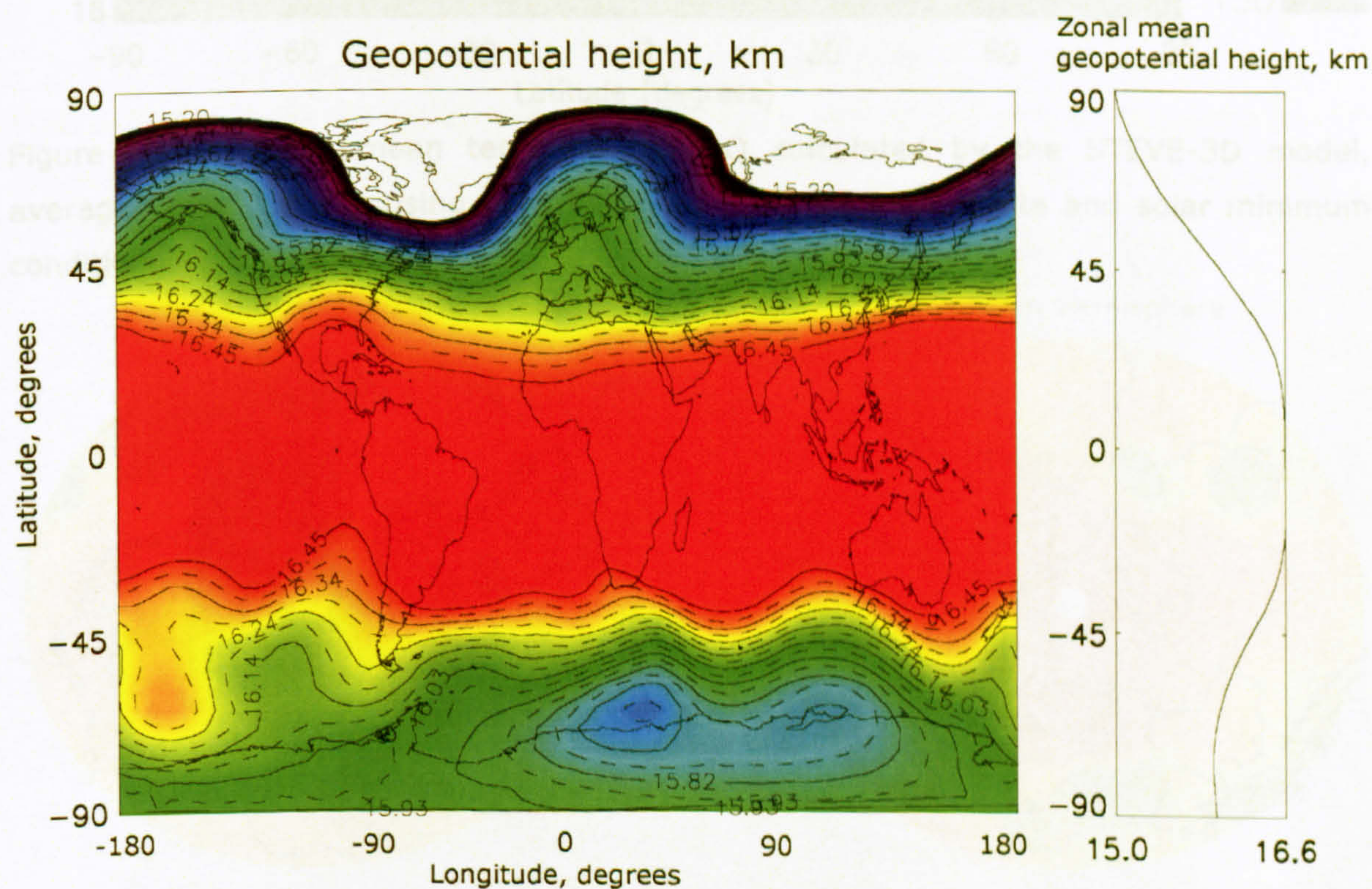


Figure III.48, Zonal mean and latitude longitude plot of the geopotential height of the 100 mb pressure level as a function of latitude for the December solstice during 1992. Data from the assimilated Stratospheric Sounding Unit (Miller *et al.*, 1980).

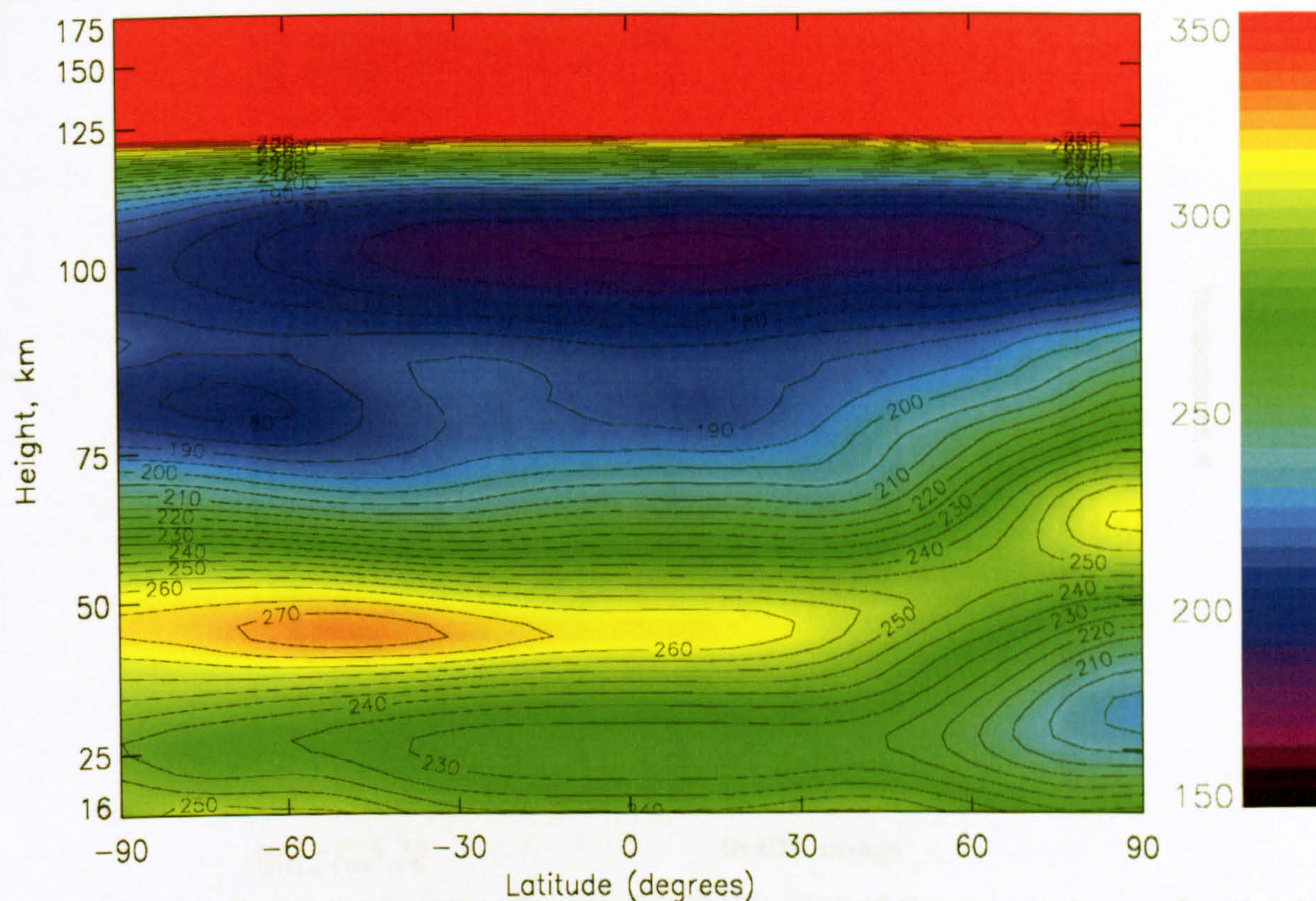


Figure III.49, Zonal mean temperatures (K) calculated by the STEVE-3D model, averaged over January using the 1993 SSU lower boundary data and solar minimum conditions.

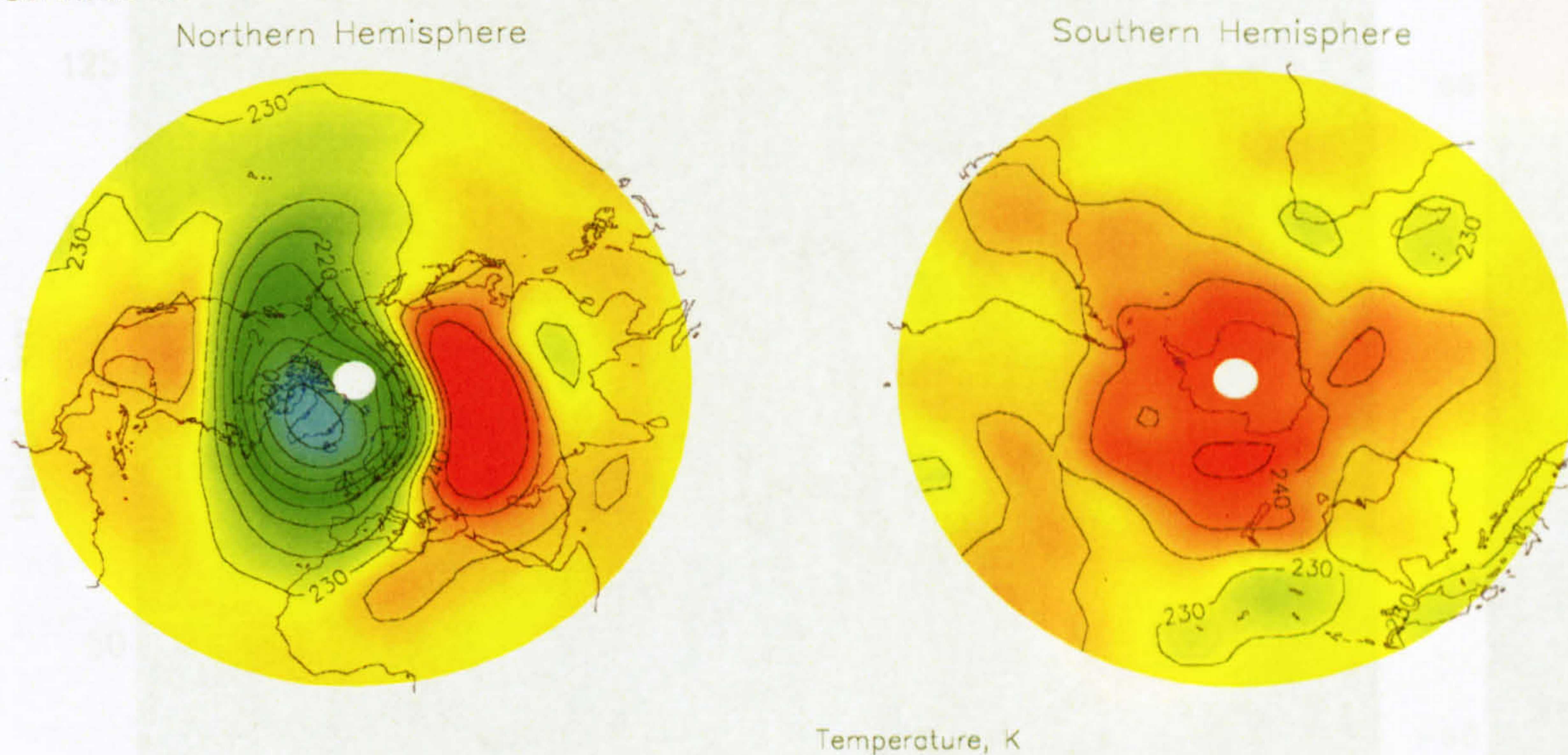


Figure III.50, Daily averaged, polar stereographic plot of the temperatures (K) for the 9th January at 10 mb calculated by the STEVE-3D model using the 1993 SSU lower boundary data and solar minimum conditions.

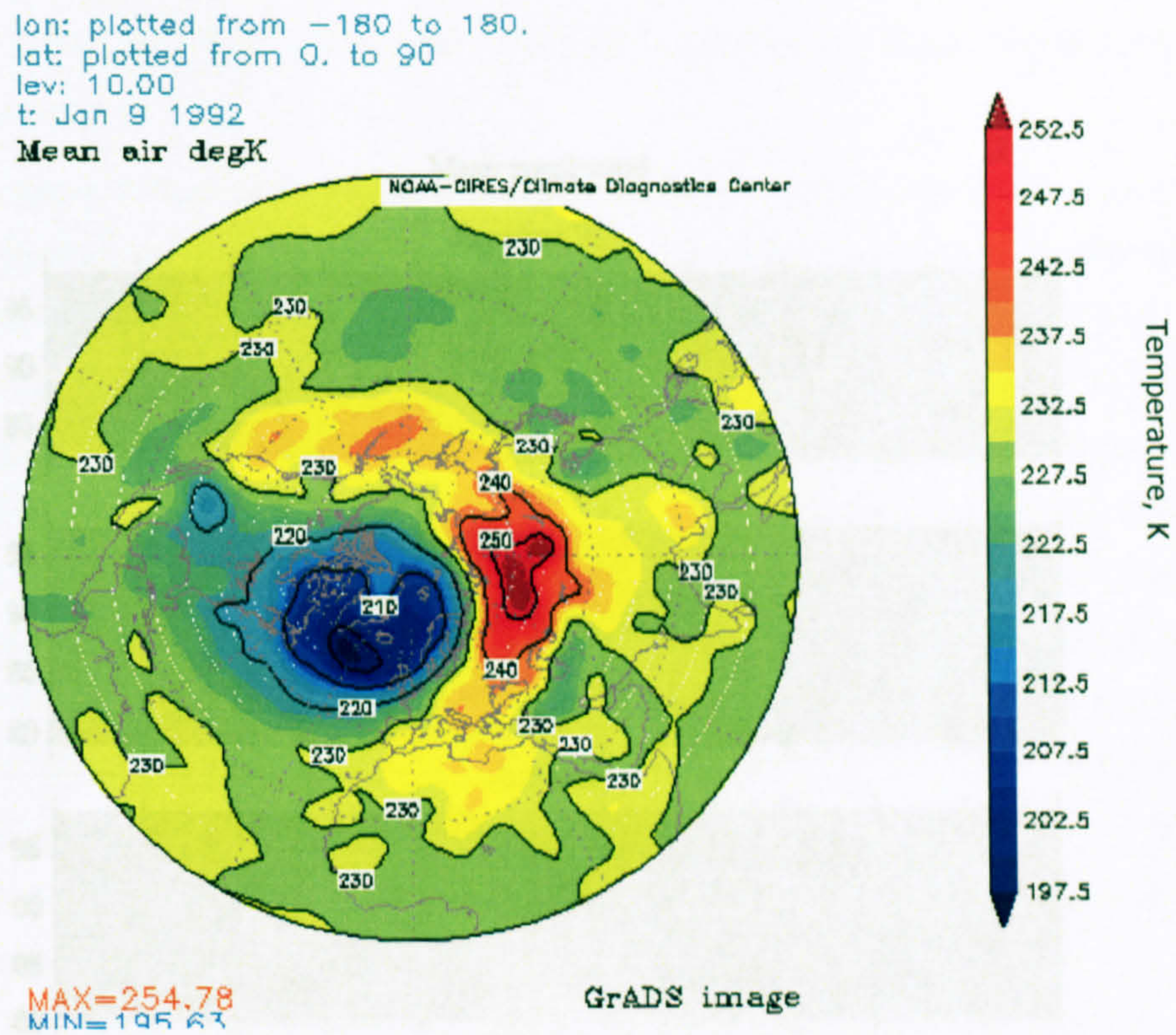


Figure III.51, Daily averaged, polar stereographic plot of the temperatures for the 9th January 1993 at 10 mb in the winter hemisphere from the NCEP reanalysis database.

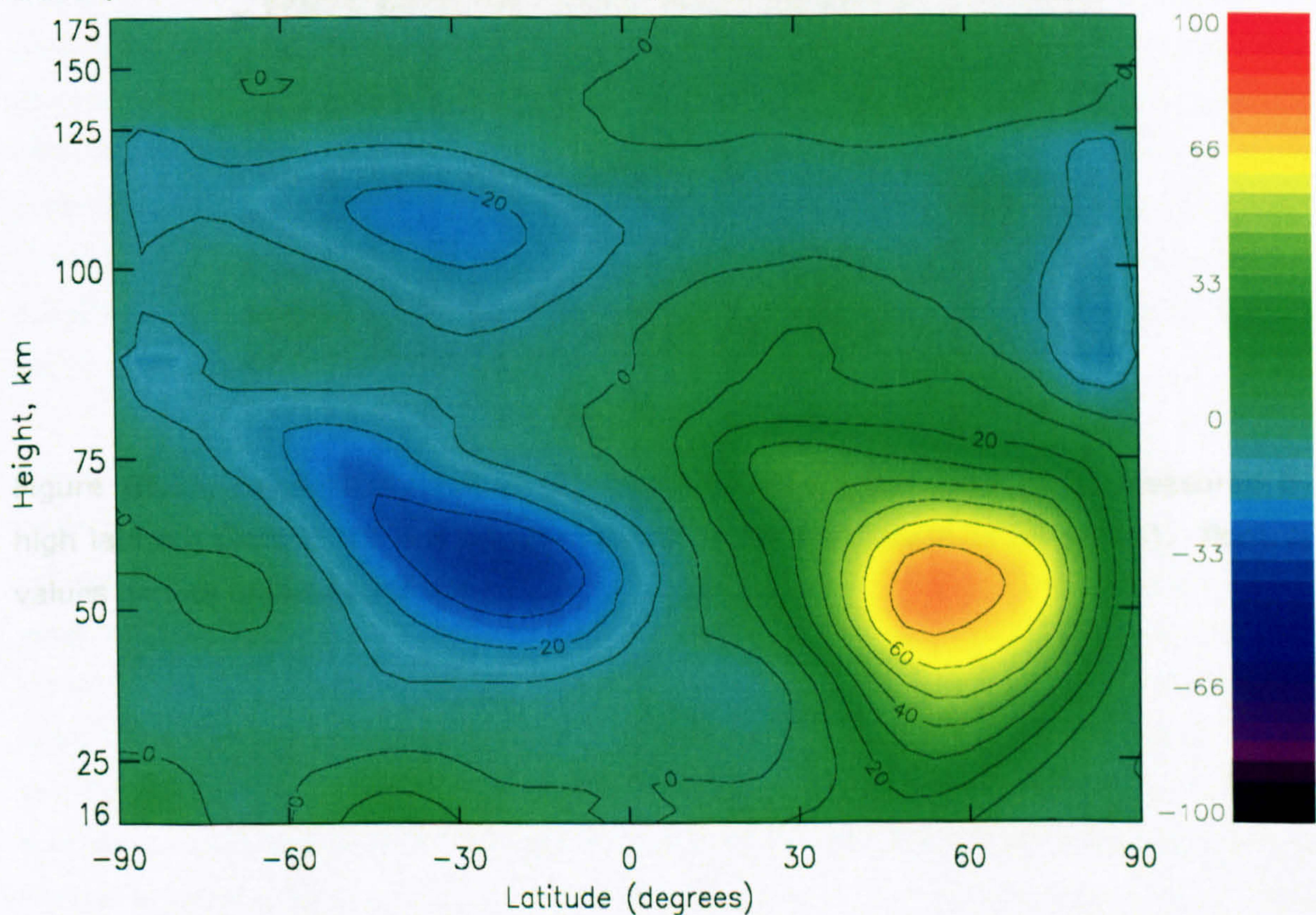


Figure III.52, Zonal mean zonal winds (m/s) calculated by the STEVE-3D model, averaged over January using the 1993 SSU lower boundary data and solar minimum conditions. Positive values denote eastwards winds.

CHAPTER IV. MODELLING THE IMPACT OF SOLAR STORMS IN THE MIDDLE ATMOSPHERE

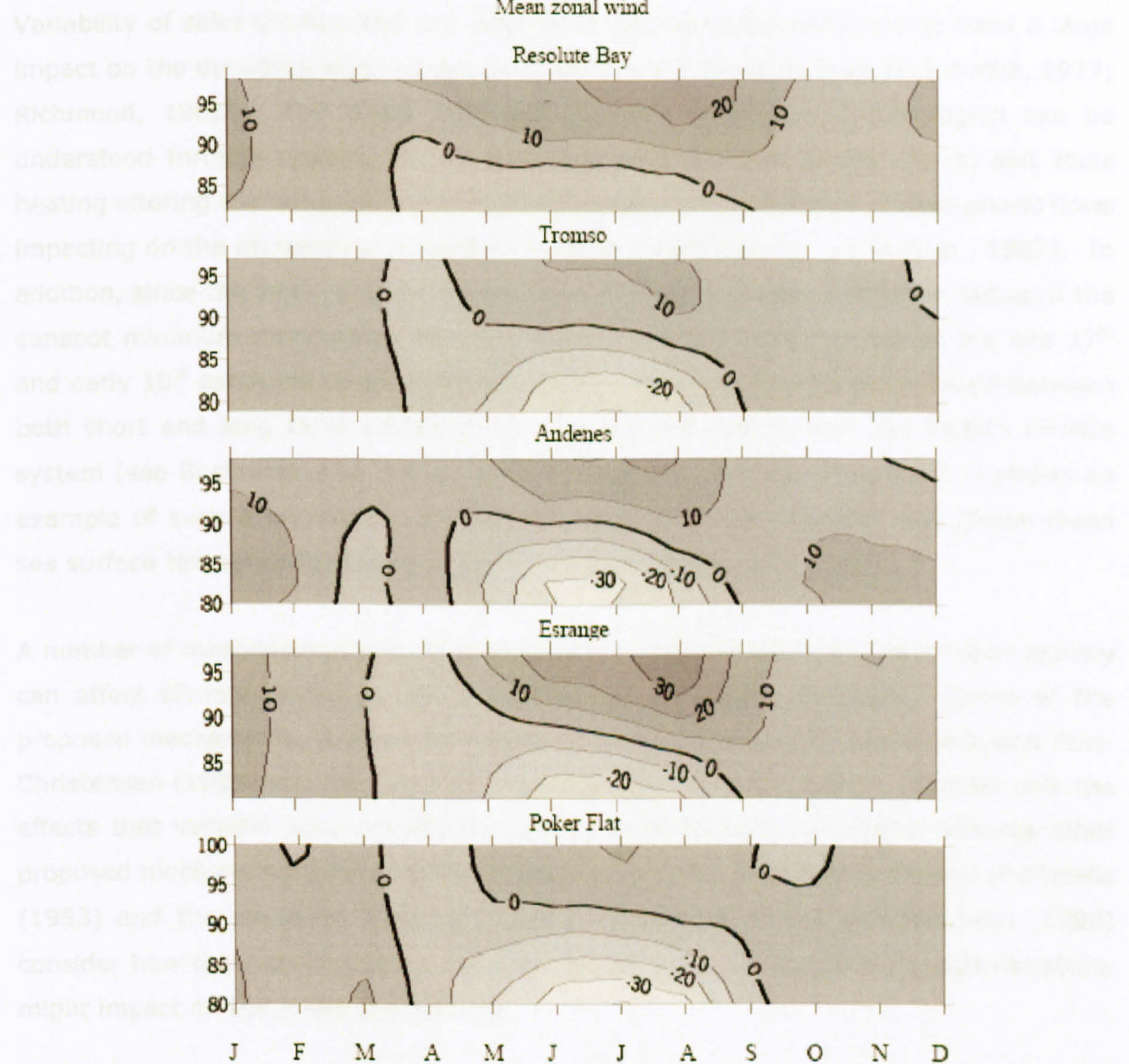


Figure III.53, Height/time plot of the monthly mean zonal wind (m/s) measured by high latitude (Arctic) MF and meteor radars. After Portnyagin *et al.* (2004). Positive values denote eastwards winds.

CHAPTER IV, MODELLING THE EFFECT OF THE SOLAR CYCLE ON GRAVITY WAVES IN THE MIDDLE ATMOSPHERE

Variability of solar UV flux and the solar wind has been demonstrated to have a large impact on the dynamics and energetics of the neutral thermosphere (e.g. Roble, 1977; Richmond, 1983). The direct influence of solar variability in this region can be understood through changes in the high frequency UV (see section I.4.1) and Joule heating altering the temperature of the thermosphere and changes in ionospheric flows impacting on the momentum budget of the thermosphere (e.g. Roble *et al.*, 1987). In addition, since the work of Eddy (1976), who showed a rough correlation between the sunspot minimum reported by Maunder (1894) and the Little Ice Age of the late 17th and early 18th centuries, there have also been a number of correlations found between both short and long term variability of solar activity indices and the Earth's climate system (see Bochníček and Hejda, 2002 for a recent review). Figure (IV.1) shows an example of such a correlation between the annual sunspot number and global mean sea surface temperatures between 1860 and 1985 (after Reid, 2000).

A number of mechanisms have been proposed to explain how changes in solar activity can affect climate, although no one theory is universally accepted. Some of the proposed mechanisms, such as the cosmic ray / cloud theory of Svensmark and Friis-Christensen (1997) and the solar UV / ozone theory of Haigh (1996), consider only the effects that variable solar activity can have on the lower atmosphere, whereas other proposed mechanisms such as global electric circuit / cloud theory of Tinsley and Heelis (1993) and the planetary wave amplification theory of Arnold and Robinson (1998) consider how the relatively large response of the upper atmosphere to solar variability might impact on the lower atmosphere.

This chapter will present the STEVE model simulation of the effects of the 11-year solar cycle on gravity-wave activity in the middle atmosphere. Due to the long data sets required and difficulties associated with removing the effects of other impacts such as large volcanic eruptions, there is relatively little observational evidence of an 11-year signature in gravity-wave activity in the middle atmosphere. No previous attempts have been made to reproduce these observations in atmospheric models, but Arnold and Robinson (2003), using observational results of Gavrilov *et al.* (1995), have shown that even small changes in the gravity-wave drag can increase the response of the Northern Hemisphere winter polar stratospheric response to the solar cycle. This result, along with other modelling studies of the downward propagation of stratospheric anomalies (e.g. Holton and Mass, 1976; Yoden, 1987; Christiansen,

The Effect of the Solar Cycle on Gravity Waves in the Middle Atmosphere Chapter IV
1999), show how changes in gravity-wave activity in the middle atmosphere might affect the Earth's climate system.

In order to demonstrate that the STEVE model may be suitable for this study, section (IV.1) will show the annual cycle in gravity-wave activity in the mid-to-high latitude middle atmosphere in the STEVE model and compare it to the well documented observational evidence and the models proposed by Eckermann (1995; EK95 hereafter) and Fritts and Vincent (1987; FV87 hereafter). Section (IV.2) will show the effect of variable solar UV flux on gravity waves in the STEVE model to see if any of the observed features can be reproduced in a relatively simple atmospheric model. Section (IV.3) will use the STEVE model to investigate the factors affecting this result and discuss the thunderstorm-track mechanism suggested by Bošková and Laštovička (2001).

IV.1 MODELLING THE EFFECT OF THE ANNUAL CYCLE ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

In the absence of extensive observations or any other model results, it is not possible to rigorously validate the solar cycle effects on gravity waves simulated using the STEVE model. However, this section will demonstrate the model's ability to perform long-term simulations of gravity-wave activity and to replicate the well-documented annual cycle effects on gravity-wave activity at mid-to-high latitudes in the middle atmosphere. The observed effects fall into two categories; the annual cycle in gravity-wave activity seen in the stratosphere and lower mesosphere, and the semi-annual cycle in gravity-wave activity seen in the mid and upper mesosphere. The following section will briefly discuss these observations. Section (IV.1.2) will outline the two widely accepted theories which explain these effects; the stratospheric annual cycle theory of EK95 and mesospheric semi-annual cycle of FV87. Section (IV.1.3) will present the STEVE model results and compare these to the observations and theories presented here.

IV.1.1 OBSERVATIONS OF THE EFFECT OF THE ANNUAL CYCLE ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

At mid to high latitudes in the stratosphere and lower mesosphere, rocket (e.g. Hirota, 1984; EK95) and lidar observations (e.g. Mitchell *et al.*, 1991; Wilson *et al.*, 1991) of gravity-wave horizontal wind variances and potential energy per unit mass respectively, show an annual cycle in gravity-wave activity, which maximises during winter and minimises during summer. Figure (IV.2) shows an example of this cycle,

taken from rocket observations of the total horizontal wind variance at 60 km altitude above Primrose Lake (55° N, 110 °W) and Thule (77° N, 69 °W; after EK95). In figure (IV.2), the wave activity observed during winter is 2 – 3 times greater than that observed during summer.

In the middle and upper mesosphere, a semi-annual variation in gravity-wave activity has been observed in radar observations at mid to high latitudes (e.g., Meek *et al.*, 1985; FV87; Manson and Meek, 1993; Gavrilov *et al.*, 1995; Nakamura *et al.*, 1996). Figure (IV.3) shows the gravity-wave spectral densities for horizontal and vertical wind components from the phase coherent medium frequency radar at Saskatoon (52° N, 107 °W) from Manson and Meek (1993). The observations are divided into for day and night and cover gravity waves with periods 10 – 100 minutes observed during 1989. The semi-annual oscillation which maximises during winter and summer and minimises during autumn and spring can be clearly seen in both the day and night-time observations below the mesopause (~90 km). Taking the square of the values shown, it can be seen that both components of the horizontal wind variance vary by a factor of 2 between solstice and equinox. These observations also show a difference between the gravity-wave activity during the day and night, which the authors attributed to gravity wave / diurnal tide coupling. This coupling will be discussed in detail in chapter (VI).

IV.1.2 THEORIES OF THE EFFECT OF THE ANNUAL CYCLE ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

The theories proposed by EK95 and FV87 have been used to successfully explain the annual cycle in gravity-wave activity in the stratosphere and the semi-annual cycle in gravity-wave activity in the mesosphere respectively. Neither of these theories requires the tropospheric source to be time-dependant, but both require changes in the middle atmosphere conditions to alter the wave activity. Both theories start from the Lindzen (1981) linear saturation theory for vertically propagating, medium-frequency gravity waves. This theory shows that, in the absence of dissipation, the horizontal wind variance associated with a single gravity wave varies as a function of altitude, given by,

$$\overline{u'^2} = \overline{u_0'^2} \frac{m}{m_0} \exp \frac{(z - z_0)}{H_\rho} \quad \text{IV.1}$$

where $\overline{u'^2}$ is the horizontal wind variance, H_ρ is the density scale height and the subscript 0 denotes the initial value of each parameter. From this equation it can be seen that, in the absence of dissipation or background wind shear, the horizontal wind

variance for a single gravity wave grows exponentially with altitude throughout the middle atmosphere.

Using equation (IV.1), the mid-frequency dispersion relation (equation II.31) and neglecting wind shear, EK95 showed that the horizontal wind variance was proportional to the buoyancy frequency of the middle atmosphere,

$$\overline{u'^2} \propto \frac{N(z)}{N_0} \exp\left(\int_{z_0}^z \frac{dz}{H_\rho(z)}\right) \quad \text{IV.2}$$

and therefore related to the background temperature of the atmosphere. At mid to high latitudes, there is an annual cycle in temperatures in the stratosphere and lower mesosphere, shown in figure (III.16) at 67.5° North using data from the MSIS-E90 empirical model. The low temperatures in the winter lead to smaller scale heights and larger $N(z)$ and therefore a faster growth of wave activity with height than during summer. Figure (IV.2) shows normalised values of the observed horizontal wind variance and those predicted by EK95. As can be seen, despite using many simplifications, the EK95 theory is capable of reproducing the annual cycle in the observations with reasonable accuracy.

EK95 also showed that a similar relationship to (IV.2) exists for the temperature variances. By assuming that $f/\omega \ll 1$, and using the relation between the temperature and wind perturbations created by a gravity wave, given by,

$$\frac{\overline{T'^2}}{T_0^2} \approx \frac{N^2}{g^2} \left[1 - \left(\frac{f}{\omega}\right)^2\right] \overline{u'^2} \quad \text{IV.3}$$

EK95 showed that the height evolution of the temperature perturbations for a gravity wave in the stratosphere could be approximated to,

$$\frac{\overline{T'^2}}{T_0^2} \propto \frac{N^3(z')}{N_0} \exp\left(\int_{z_0}^z \frac{dz}{H_\rho(z)}\right). \quad \text{IV.4}$$

From equation (IV.4) it can be seen that this theory predicts that the temperature variances obey a similar annual cycle behaviour to the horizontal wind variances, but scaled by an additional factor of N^2 which leads to a stronger observed annual cycle, which is in agreement with the observations (e.g. EK95).

In the middle and upper mesosphere, FV87 argued that gravity wave amplitudes are limited by convective instability (linear saturation condition, Lindzen, 1981). Under this theory, the maximum variance which a single medium frequency gravity wave can have before it becomes convectively unstable is given by,

$$\overline{u'^2} = \frac{(c_{ph} - u \cos \phi)^2}{2}$$

IV.5

+where c_{ph} is the horizontal phase speed in the Eulerian frame, ϕ is the angle, in the horizontal plane, between the background wind velocity and the direction of phase propagation (e.g. Fritts, 1984). From equation (IV.5) it can be seen that, given no preferential phase propagation direction of the gravity-wave source, the maximum achievable horizontal wind variance occurs when the background wind maximises in the opposite direction to the wave phase propagation. Figure (IV.4) shows the zonal mean zonal winds at 67.5° North throughout the year, from the HWM empirical model (as only winds in the mesosphere are considered here, the HWM results are adequate for the discussion here). Between ~70 – 100 km, the largest zonal mean zonal winds at this latitude are seen during the two solstices, corresponding to the larger horizontal gradient in temperature seen at solstice than at equinox (figures III.12 and III.14). This, along with equation (IV.5) explains the semi-annual cycle seen in radar observations of wind variance at these altitudes.

It is worth noting that the two theories stated here rely on approximations which apply only to certain altitude ranges, although neither theory explicitly states the altitude range over which these approximations hold. Indeed, the regions over which the saturation condition used by FV87 and the conservative propagation condition used by EK95 are valid, are both functions of the background atmosphere, and therefore depend on season and location. This can be highlighted by contrasting the annual cycle in gravity-wave activity seen at 60 km altitude above Primrose Lake and Thule (figure IV.2) with the semi-annual cycle observed at heights down to 61 km altitude above Saskatoon (figure IV.3). The following section will discuss the STEVE model results for the effect of the annual cycle on gravity-wave activity throughout the middle atmosphere, and highlight the region of transition from the annual to semi-annual oscillation seen in the model results.

IV.1.3 STEVE MODEL RESULTS OF THE EFFECT OF THE ANNUAL CYCLE ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

The STEVE model was run for 19 months, starting from the June solstice, under perpetual solar minimum conditions. The annual cycle in solar heating is included, but no time dependence in the gravity-wave source is included in the model simulation. The MK2000 gravity wave parameterisation used in the model calculates the sub-grid-scale horizontal wind variance created by gravity waves (see section III.4.5). The first 5.5 months of data are not used, so that the starting conditions of the simulation have

as little effect on the results as possible. Figure (IV.6) shows the zonal mean total horizontal wind variance calculated by the STEVE model as a function of time at 67.5° North and at approximately 64 km altitude (the data shown is for the 0.042 mbar pressure surface). The STEVE model reproduces the key features of the observed annual cycle in total gravity wave wind variance (figure IV.2), with a maximum during winter and a minimum during summer. The model simulation shows a factor of 4 difference in the total variance between summer and winter, which is comparable to the factor of 4 observed and the factor of 3 modelled by EK95 at 55° N and 77° N. However, the model results show a more rapid change between the maximum and minimum values and a more prolonged winter maximum than the observations shown here. This is consistent with the more pronounced annual cycle in temperatures seen in the STEVE model (figure III.15) than, for example the MSIS-E90 empirical model (figure III.16) at the same position.

In order to demonstrate that the STEVE model is able to reproduce the semi-annual oscillation in gravity-wave activity in the middle and upper mesosphere, in accordance with the theory of FV87, the model must be able to reproduce the annual cycle in zonal mean zonal winds, shown in figure (IV.4) (as the STEVE model has been shown to produce zonal winds which are more consistent with observation and theory than the simulated meridional winds, discussed in section III.4.2, only the zonal winds and zonal variance will be considered in the middle and upper mesosphere region, where the background wind is believed to be crucial). Figure (IV.5) shows the time series of the zonal mean zonal winds in the STEVE model at 67.5° N during solar minimum conditions. In the middle and high mesosphere region, the STEVE model closely reproduces the amplitudes and the annual cycle in zonal mean zonal winds produced by the HWM model. According to the FV87 theory described in the previous section, the relatively high zonal mean winds in the STEVE model at lower altitudes (discussed in section III.4.2) should not interfere with the model's ability to reproduce the semi-annual cycle in zonal wind variance in the middle and upper mesosphere, as this theory depends solely upon the *in situ* winds (it is worth noting that the zonal winds in the URAP, shown in figure I.5, are also stronger than those produced by the HWM).

Figure (IV.7) shows the time series of the zonal mean zonal wind variance calculated by the STEVE model at ~88 km altitude (the 7.5×10^{-4} mb pressure surface) at 67.5° N. The STEVE model reproduces the semi-annual oscillation, with the solstitial maximum and equinoctial minimum, in agreement with the observations shown in figure (IV.3). The slight lag of the spring minimum and summer maximum is larger in the model than in the observations presented here, but is consistent with the lag of the peak zonal mean winds in the model, which occurs after the summer and winter solstices.

This lag is associated with the response time of the atmosphere to changes in the solar forcing. A factor of 3 variation is seen in the modelled zonal mean zonal wind variance between equinox and solstice, which is comparable to the factor of ~ 2 variation seen in figure (IV.3).

The day-to-day variation in the zonal mean zonal wind variance seen in figures (IV.6 and IV.7) highlights the nonlinear behaviour of the MK2000 gravity wave parameterisation used in the STEVE model. From relatively small day-to-day changes in the background wind (IV.6) and temperature (figure III.16) fields, very large changes in the gravity-wave activity can sometimes be seen. This feature is a consequence of the nonlinear theory on which this parameterisation is based (section II.4.6) and is seen in the strong latitudinal variations in the gravity-wave drag calculated by this parameterisation (figure III.40). It is worth noting that such behaviour is not unique to this parameterisation (e.g. McLandress, 1998), and similar day-to-day variations are reported in observations (e.g. FV87; Thayaparan *et al.*, 1995; Bošková and Laštovička, 2001), although these include effects of variable gravity-wave sources. This result highlights the difficulty associated with analysing short-term gravity wave variability in global-scale models. Throughout this work, this will be addressed by using multiple data points, either in space or time, to show each effect.

As noted in the previous section, the EK95 and FV87 theories which describe the variability of gravity-wave activity in the middle atmosphere, do not tightly constrain the altitude range over which these theories apply. Figures (IV.8a and IV.8b) show the height-time structure of the zonal mean total horizontal wind variance and the zonal mean zonal wind variance at 67.5° North calculated by the STEVE model. Both of these figures show an annual cycle in activity in the upper stratosphere and lower mesosphere, up to ~ 70 km, above which a semi-annual cycle is seen in the mesosphere. A further feature seen in the model is an annual cycle in the wind variances which is seen in the lower thermosphere, which is in agreement with the observations of gravity-wave activity in the lower ionosphere (~ 95 km) reported by Bošková and Laštovička (2001), who used the low-frequency D1 method (see section II.3.1.3) at Průhonice (50° N, 15° E), but is contradictory to the results of Gavrilov *et al.* (1994) who used an incoherent scatter radar at Saint Santin (45° N, 2° E) to measure variations in ion temperature, ion drift velocity and electron density in the ionosphere and saw an annual cycle with a minimum during summer. Bošková and Laštovička (2001) suggested that the discrepancy between the observations may be explained by differences in the gravity wave frequencies observed, but so far no theory has been presented which explains either of these observations. As the validity of the

STEVE model above 100 km altitude is questionable, this issue will not be addressed in any further detail here.

As a final test of the model, the variability of gravity-wave activity at the equator is investigated. Observational studies such as EK95 and Manson *et al.* (2002) have shown that the clear annual and semi-annual cycles seen at mid to high latitudes are not present at equatorial latitudes (although tidal variations with season may be important here – see chapter VI). This can be explained by the lack of strong seasonal variation in the zonal mean temperatures and winds. Figures (IV.9 and IV.10) show the time-height series of the zonal mean winds and temperatures calculated by the STEVE model at 2.5° North. While an annual cycle is seen in the zonal wind plot, it is of much smaller amplitude than at higher latitudes, especially in the middle to upper mesosphere where the FV87 theory holds. There is virtually no seasonal change in the temperatures at this latitude in the model. Figures (IV.11 and IV.12) show the height-time series of the zonal mean total horizontal wind variance and zonal mean zonal wind variance produced by parameterised gravity waves in the STEVE model at 2.5° North, respectively. Both plots show that the wind variances associated with gravity-wave activity are generally lower than at mid latitudes, in agreement with the lower gravity-wave source amplitude at low latitudes used in the STEVE model (section III.4.5) and the lower gravity-wave drag in the model at low latitudes, compared with mid latitudes (figures III.39 and III.40). Neither the total wind variance, or the zonal wind variance plots show any clear seasonal cycle, in agreement with the observations and theory.

In summary, the STEVE model has been used to replicate the observed seasonal variation in gravity-wave activity in the middle atmosphere. This seasonal variation was produced using a time invariant gravity-wave source, showing that the variation is produced by *in situ* temperature and wind variations, in agreement with the theories detailed in this chapter. These tests have shown the STEVE model to be suitable to study long-term variations in gravity-wave activity in the upper stratosphere and mesosphere. This region will be considered in the following section, when the STEVE model is used to simulate the effects of the solar cycle in UV flux on gravity-wave activity in the middle atmosphere.

IV.2 MODELLING THE EFFECT OF VARIABLE SOLAR UV FLUX ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

The following section will introduce the observational evidence of the effect of the solar cycle on gravity-wave activity in the middle atmosphere. Common features of and

differences between these observations will be highlighted along with theories that have been presented to explain these observations in section (IV.2.2). In section (IV.2.3), the STEVE model will be used to simulate the effects of the solar cycle in UV flux on gravity waves in the middle atmosphere, to see if any of the observed features are reproducible in a relatively simple dynamical model. In section (IV.3), the theories presented in section (IV.2.2) will be tested using the STEVE model.

IV.2.1 OBSERVATIONS OF THE EFFECT OF THE SOLAR CYCLE ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

There are relatively few data sets of continuous gravity wave observations in the middle atmosphere over the whole, or a significant portion of the 11-year solar cycle, but several observations by middle atmosphere radars have been published over the past decade. Gavrilov *et al.* (1995), Laštovička (1999), Bošková and Laštovička (2001) and Gavrilov *et al.* (2002) have all reported radar observations of gravity-wave activity varying over the solar cycle. The results of each of these studies will be detailed here.

Using the MF radar at Saskatoon (52° N, 107 °W) between 1979 and 1993, Gavrilov *et al.* (1995) provided the first observational evidence of an effect of the solar cycle on gravity waves in the middle atmosphere. Figure (IV.13) shows the monthly differences between total wind variance and the mean wind variance measured by the Saskatoon MF radar as a function of solar 10.7 cm flux using the three different temporal filters (WH denotes periods less than 1 hour, HD denotes periods between 1 – 6 hours and RS denotes periods between 2 – 10 hours after diurnal and semi-diurnal harmonics have been removed from the wind data). While there is considerable spread in the monthly mean values, all three data sets show a negative, or near zero, correlation between gravity-wave activity and solar 10.7 cm flux. The amplitude of this change in gravity-wave activity varies with height between 64 and 99 km.

Laštovička (1999) and Bošková and Laštovička (2001) both used a 7 year night time data set (1988 – 1995) from the LF D1 drifts at Průhonice (50° N, 15 °E) to study the effects of the solar cycle on gravity-wave activity in the lower ionosphere (~95 km). Figure (IV.14) shows the gravity-wave activity during winter for gravity waves with periods 10 – 90 minutes, along with the sunspot number / 100 for each 3 month bin. Laštovička (1999) showed that, for these periods of gravity waves, there was a good correlation (over 95 % significance) between the sunspot number and gravity wave amplitude during winter (for longer period gravity waves, this significance was shown to be lower than 95 %, and the effects of the Mt. Pinatubo eruption were shown to be

more important than for shorter period gravity waves). This data from Průhonice appears to show the opposite effect of the solar cycle on gravity-wave activity than that observed at Saskatoon, indicating that geographical position may be an important factor.

Bošková and Laštovička (2001) also showed that the data from Průhonice indicated that the effect of solar activity on gravity-wave activity in the middle atmosphere may depend upon season. Figure (IV.15) shows the monthly average gravity-wave activity measured by the radar for gravity wave periods 10 – 30 minutes (gravity wave periods up to 90 minutes show a similar response). The pronounced summer maximum during medium solar activity (1992 – 1993) does not appear in the data for high solar activity (1989 and 91), suggesting that during summer there is a negative correlation between solar activity and gravity-wave activity in amplitudes during summer, which is opposite to the effect seen during winter.

So far, only Gavrilov *et al.* (2002) have explicitly investigated the geographical dependence of the effect which solar activity has on gravity waves in the middle atmosphere. The authors used long-term measurements from the MU radar at Shigaraki (35° N, 136 °E) from 1986 – 1999 between 65 – 80 km, the LF D1 drifts at Collm (52° N, 15 °E) from 1983 – 1999 between 80 – 110 km and the MF radar at Saskatoon (52° N, 107 °W) from 1983 – 1999 between 60 – 100 km, which are shown in figure (IV.16). The quadratic fits to the horizontal wind data show the apparent effect of the solar cycle on gravity-wave activity at each height and frequency range. Changes of up to 100 % between solar minimum and maximum are seen at certain locations. The quadratic fits are dependant upon height, geographical location, frequency range and cardinal direction (meridional or zonal), showing a complicated picture of the response of gravity waves in the middle atmosphere to the solar cycle.

It is worth noting that other studies, which have specifically looked for a correlation between solar activity and gravity-wave activity measured by middle atmosphere radars, have failed to find such a correlation. Eckermann *et al.* (1997) used an MF radar at Christmas Island (2° N, 157 °W) to study gravity-wave activity between 60 – 98 km during 1990 – 1994, but saw no correlation between solar activity and gravity-wave activity, although this might be explained by the comparatively short data set or the geographical dependence on such a correlation noted above.

Further, it is worth noting that multiple year observations (e.g. Bošková and Laštovička, 2001; Eckermann *et al.*, 1997) show that there is no detectable response of gravity-wave activity to the 27 day solar rotation period in the middle atmosphere,

despite this having an effect on stratospheric ozone (e.g. Chandra *et al.*, 1994). Pancheva *et al.* (1991) showed that in the case of particularly pronounced solar 27 day variations, a 27 day effect on radio wave absorption could be detected (through varying the ionisation), but in other circumstances 27 day variations are usually associated with planetary wave activity of similar periods.

In summary, the observational evidence of the influence of solar activity on gravity-wave activity indicates that, at least at mid-latitudes, gravity-wave activity may be modulated by solar activity, depending on height, geographical position, gravity-wave propagation direction and the frequency of the gravity waves. Changes in gravity-wave horizontal wind variance up to ~100 % have been seen in the mesosphere. Evidence of such effects in the STEVE model will be investigated in section (IV.2.3).

IV.2.2 THEORIES OF THE OBSERVED EFFECT OF THE SOLAR CYCLE ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

Several theories have been proposed which suggest that the observed changes in gravity-wave activity detailed in the previous section can be explained by either solar activity effects, or other atmospheric effecters, such as volcanic eruptions. Gavrilov *et al.* (1994) used a simple model of the atmosphere to show that solar impacts on the neutral and electron densities (and therefore ion-drag) could produce a negative correlation between solar activity and gravity-wave activity in the thermosphere. This model result can explain the thermospheric gravity wave observations (100 – 300 km altitude) of Gavrilov *et al.* (1994), but is apparently at odds with the observations of Laštovička (1999).

Gavrilov *et al.* (1995) argued that the observations at mesospheric altitudes could not be easily explained by changes in neutral and electron densities. Gavrilov *et al.* (1995) suggested that solar cycle influences on the lower atmosphere might alter the gravity-wave activity seen in the mesosphere, presumably by altering the wave filtering and wave sources at these latitudes.

Gavrilov *et al.* (2002) proposed that other changes in the lower atmosphere, created by the eruption of Mount Pinatubo in June 1991 and El Niño events (which maximised during 1990 – 1994), could be responsible for inter-annual changes in gravity-wave sources. The effect of El Niño on tropospheric winds was observed by Gage *et al.* (1996), who noted that the smallest amplitude seasonal variation in zonal winds were observed during 1991 – 1992, which roughly coincides with maximum in solar activity.

Laštovička *et al.* (1998) showed that the Mount Pinatubo volcanic eruption significantly enhanced the activity of gravity waves with periods over 90 minutes at Průhonice. However, the authors showed that the eruption had very little, if any, effect on gravity waves with periods less than 90 minutes.

Bošková and Laštovička (2001) proposed that changes in the position of the northern hemisphere winter storm track around the solar cycle (Brown and John, 1979), could lead to an increased gravity-wave source around 50° N at solar maximum, compared with solar minimum. This mechanism may explain the observations at Průhonice during winter, although this alone would not explain either the summer-time observations at Průhonice, or the observations at more distant locations.

Both Laštovička (1999) and Bošková and Laštovička (2001) also cited possible changes in gravity wave filtering in the stratosphere with solar activity as possible explanations for their observations. Observational studies such as Labitzke and van Loon (1988, 1995) and Hood *et al.* (1993) and modelling studies such as Larkin *et al.* (2000) and Arnold and Robinson (1998, 2000, 2002) have shown seasonally and geographically dependant changes in the stratospheric winds correlated with solar activity. Such changes may account for the geographical dependence of the observations detailed in the previous section.

In section (IV.3) the STEVE model will be used to test quantitatively these theories as well as other possible explanations for the solar-cycle signature in gravity-wave activity seen in the model.

IV.2.3 STEVE MODEL RESULTS OF THE EFFECT OF VARIABLE SOLAR UV FLUX ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

In order to study the effects of the solar cycle on gravity waves in the STEVE model, a continuous 21 year simulation was performed. An 11-year cycle in solar UV flux was included in the simulation, using data from the UARS SOLSTICE instrument (Rottman, 1985) for the changes in the solar flux at each frequency band used in the model. Changes in the chemical constituents of the atmosphere provided by the MSIS-E90 model, which are used by the radiative heating, cooling and diffusion calculations are also included in the simulation although no change in the ozone distribution is included in the model. As the STEVE model is a purely neutral atmosphere model and contains no detailed representation of ion-neutral collisions, changes in geomagnetic conditions and ionospheric parameters are not included in the model. It is important to note that no interannual change in the lower boundary geopotential height data is included in the

model and no changes in the gravity-wave source are included. To ensure that the initial conditions used in the simulation have as little impact on the model results as possible, only data from the second year onwards will be presented here.

Before investigating the changes in gravity-wave activity in the STEVE model, it is important to see that the changes in solar UV with solar activity are impacting on the model. As the highest frequency solar UV radiation is absorbed in the thermosphere, the largest direct effect of the change in solar UV flux on the STEVE model should occur in the thermospheric diabatic heating rate. Figure (IV.17a) shows the difference in the zonal mean diabatic heating rate due to UV absorption between solar minimum and maximum calculated by the model at the June solstice. A strong increase in the heating rate is seen throughout the sun-lit portion of the thermosphere, peaking above 150 km altitude. At mesospheric altitudes, only a small increase in the diabatic heating rate is seen, which corresponds to the relatively small changes in solar UV at the frequencies absorbed at these altitudes.

This increase in the heating rate in the thermosphere leads to an increase in temperature in this region, shown in figure (IV.17b). Temperature increases of over 170 K are seen at ~ 160 km over the summer pole. The temperature increases in the upper-most levels of the model are strongly affected by the estimate of the thermal diffusion through the top of the model and so are not fully self-consistent (see section III.3.5). The temperature increase seen in the thermosphere below ~ 160 km is consistent with that calculated by the MSIS-E90 empirical model, shown in figure (I.1).

In the mesosphere and stratosphere, relatively small changes in temperature are seen in the STEVE model, in agreement with figure (I.1). The small decreases in temperature seen in the mesosphere are associated with changes in the net diabatic heating rate (figure IV.17c) and diffusion, which are both dependant on the temperature and composition of the atmosphere and changes in the dynamics, which are discussed below.

In order to compare the effects of the solar cycle on gravity-wave activity in the STEVE model with the observations presented in section (IV.2.1), the zonal wind variances at various heights throughout the MLT region at 52.5° North are shown in figure (IV.18). This height range and latitude is comparable with the measurements of the LF D1 drifts at Průhonice and Collm and the MF radar at Saskatoon and the wind variances calculated by the STEVE model (calculated at each grid cell by the MK2000 scheme – see Section III.4.5 for further details) in this height range have been shown to be realistic. In order to compensate for the short-period fluctuations discussed earlier, all

data points are averaged over thirty days and the standard deviation of each is shown. Both 10 year data sets are shown. Solar maximum occurs around year 5 of the first cycle and around year 6 of the second.

The least squares quadratic fits to the data shown in figure (IV.18) show changes in the zonal wind variance of up to $\sim 10\%$ between solar minimum and maximum, depending on height and season. At certain heights, such as at 82 km during both summer data sets, almost no variation is seen throughout the solar cycle, whereas at other heights such as 95 km during both summer data sets, stronger changes are seen. Also, at different heights, the opposite change is seen (e.g. 76 km and 95 km during summer). The standard deviation of the data is generally small compared with the absolute values, except at low values of zonal variance. During both summer and winter data sets, the standard deviation of individual data points is often comparable to the size of the solar-cycle signature seen in the model, although some clear trends are visible, especially during the summer data sets.

The height dependence of the solar-cycle signature seen in the model reflects that observed by Gavrilov *et al.* (2002) at Collm and Saskatoon, and the relative size of the standard deviations to the solar-cycle signature is consistent with the observations of Laštovička (1999) at Průhonice, although a smaller number of days were included in the analysis used here. However, the model results show a smaller response to the solar cycle than the observations presented above. One possible explanation for this is that the model only includes changes in UV and certain chemical species, and omits any changes in ozone, ion-neutral collisions and gravity-wave sources. Another possible explanation is that, as the model results are zonally averaged, it might be expected that this would tend to average out any longitudinal variations, so that this value would be smaller than that observed at one specific location. Additionally, it is possible that the calculation of the wind variances on constant pressure surfaces, rather than constant height levels may influence the results (see chapter VII).

Comparing the summer and winter data sets, it can be seen that the zonal wind variances are generally smaller in summer than winter, in agreement with the results presented in section (IV.1.3). During summer, the wind variances increase with altitude, whereas during winter there is a marked decrease in wind variance above 82 km, indicating that there is strong wave dissipation in this region during winter. This may explain the larger standard deviations in the winter data than the summer.

The quadratic trends during the two summer data sets are in agreement, but the winter data sets show little agreement, possibly due to the variability of the data. This

shows that although the solar-cycle signatures during the summer are quite weak, they are recurrent in the model simulation. However, comparing the summer and winter trends at each height, there is little agreement between the two. This may be indicative of the seasonal dependence observed by Bošková and Laštovička (2001), although the high standard deviations of the winter data sets make it difficult to draw conclusions. It is therefore not worth making a detailed comparison between the observed seasonal dependence and that modelled here, but still worth noting that the positive correlation between solar activity and total gravity-wave activity during summer and the negative or zero correlation modelled at 95 km during winter is in agreement with the observations at Průhonice. The cause of such effects in the model results will be addressed in the following section.

In order to show the latitudinal dependence in the STEVE model results, figure (IV.19) shows a latitude height plot of the difference in zonal mean zonal wind variances between solar maximum and minimum, again averaged over 30 days around the June solstice. From this figure, the height dependence shown above can be seen to exist at all latitudes. Further, the seasonal differences can be seen. In the extra-tropical summer hemisphere, changes in the zonal wind variance in excess of $10 \text{ m}^2\text{s}^{-2}$ are only seen above ~ 100 km, whereas in the extra-tropical winter hemisphere, changes of this size are seen above ~ 75 km. This corresponds to the high-altitude wind variance maximum at mid to high latitudes seen during summer in the STEVE model (figure IV.8b). Close to the equator, a strong change in the zonal wind variance is seen at ~ 50 km, which is associated with the zonal wind instability at this location (see section III.4.2).

Regions of correlations and anti-correlations between the gravity wave wind variance and solar activity are seen throughout the middle atmosphere, in agreement with the macro-regional interpretation of Laštovička (1999). These regions have vertical scales of 10 – 25 km (5 – 12 grid points), which broadly agrees with the observations of Gavrilov *et al.* (2002) and horizontal scales of $10 - 80^\circ$ ($\sim 1000 - 8000$ km or 2 – 16 grid points). Gravity-wave activity in the middle atmosphere has been observed to vary considerably on similar horizontal scales (e.g. Manson and Meek, 1997) and similar horizontal scales are seen in the response of the stratosphere to variable solar forcing (e.g. Labitzke and van Loon, 1988, 1995). It seems likely that longitudinal variations also exist in the solar-cycle signature of middle atmosphere gravity-wave activity, although their typical scales cannot be estimated using the current version of the STEVE model.

Table (IV.1) shows a summary of the agreements and disagreements between the STEVE model results and the various observations detailed in this chapter. Having demonstrated that the STEVE model is able to replicate some of the observed features of the solar cycle in gravity-wave activity in the middle atmosphere, the subsequent effect which this has on other model parameters can be considered. Both the gravity-wave drag and eddy diffusion coefficients used by the STEVE model are dependent upon the gravity-wave activity in the middle atmosphere and should therefore be affected by changes in the zonal wind variance. Figure (IV.20a) shows the percentage change in the zonal gravity-wave drag in the northern hemisphere around the June solstice. In order to show the change in the bulk forcing of the atmosphere, this has been averaged over the extra-tropical MLT region between 76 and 90 km altitude. This region avoids the equatorial instability in the model and lies within the height range considered above. The least-squares quadratic fit shows a positive correlation between solar activity and gravity-wave drag in this region, with an increase of $\sim 3\%$ at solar maximum compared with solar minimum, which represents a 3% increase in the acceleration of the background flow by dissipating gravity waves.

Figure (IV.20b) shows the percentage change in the zonal gravity-wave drag in the northern hemisphere around the December solstice, averaged in the same way as figure (IV.20a). Again, a positive correlation between solar activity and gravity-wave activity is seen. As with the zonal wind variances, the standard deviation of the gravity-wave drags is higher during winter than during summer. However, a more clear correlation between solar activity and gravity-wave drag is seen than with zonal wind variance data presented earlier. This may be a result of the volume averaging, or variations in other model parameters used to calculate the gravity-wave drag.

As a positive correlation between gravity-wave drag and solar activity is seen in the STEVE model, it might be expected that a positive correlation would exist between the calculated eddy diffusion coefficient and solar activity in the region considered above. Figures (IV.20c and IV.20d) show that this is the case during summer, while during winter a clear negative correlation is seen. Considering equations (II.30 and III.5), and given that an increase in the net gravity-wave drag a_h is seen in the model, this shows that a smaller number of wave frequencies are propagating through this volume during solar maximum, compared with solar minimum. This may be related to a change in the gravity wave filtering of the atmosphere below 76 km, or an increase in the number of waves dissipating completely in the lower altitude range of the volume considered here. The increase in the calculated volume averaged gravity-wave drag shown in figures (IV.20a and IV.20b) would support the latter possibility, although a clear distinction between these two possible explanations cannot be made as the fate

The Effect of the Solar Cycle on Gravity Waves in the Middle Atmosphere Chapter IV
of each individual gravity wave simulated was not recorded due to computational constraints.

From equations (I.25 and I.28) it can be seen that a change in the zonal gravity-wave drag produces a change in the residual circulation of the middle atmosphere. This can be seen by plotting the mean vertical wind, averaged in the same way as the gravity-wave drags shown in figures (IV.20a and IV.20b), shown in figures (IV.20e and IV.20f). Around both summer and winter solstices, increases in the mean vertical wind of $\sim 3\%$ are seen during solar maximum, compared with solar minimum. This shows one possible mechanism by which changes in the solar UV flux might affect lower regions of the atmosphere, such as the stratospheric winter polar region, which shall be discussed in more detail in chapter (V). Possible explanations for the modelled changes in gravity-wave activity shown here will be discussed in the following section.

IV.3 EXPLORING A POSSIBLE MECHANISM FOR A SOLAR CYCLE IN GRAVITY-WAVE ACTIVITY IN THE MLT REGION

This section will investigate the factors affecting the response of gravity waves in the MLT region to the solar cycle in the STEVE model. The impacts of a variable gravity-wave source in the lower atmosphere, changes in the filtering of gravity waves in the stratosphere and lower mesosphere and changes in the background mesosphere will be shown. Using these results, a quantitative assessment of the theories detailed in section (IV.2.2) will be made.

In this section, the impact of six sets of parameters on the response of gravity waves to the solar cycle in the STEVE model will be investigated. While the STEVE model has been designed in order to make multi-decadal simulations, the computational expense necessary to repeat the 21 year simulation a further six times is beyond the scope, or needs of this experiment. Instead, calculations will be made using only the MK2000 parameterisation and the daily averaged temperature and wind fields outputted from the 21 year simulation. This will also allow tests to be performed, which would not be possible in a fully-interacting atmospheric model.

In order to demonstrate that this simplified calculation may be used as a reasonable alternative to performing several additional solar cycle simulations, and to provide a control with which the others can be compared, the temperature and wind fields will be used to repeat the solar cycle effects shown in the previous section. The clearest solar cycle effects on the modelled gravity-wave activity were for around the June solstice in the Northern Hemisphere, and so only these results will be considered here. Figures

(IV.21a and IV.22a) show the results of this calculation for the zonal wind variance at various heights at 52.5° North and the volume averaged extra-tropical gravity-wave drag in the Northern Hemisphere MLT region, using thirty days of temperature and wind data from the second simulated solar cycle. Comparing these to figures (IV.18b and IV.20a), it can be seen that this calculation produces similar results to the full STEVE model. These differences arise from using daily averaged temperature and wind data, rather than data which changes every 900 s in the fully interactive STEVE model.

IV.3.1 THE IMPACT OF CHANGES IN THE MIDDLE ATMOSPHERE

Using the type of calculation outlined in the previous section, it is possible to show the effect of changes in the stratosphere and lower mesosphere associated with the solar cycle on gravity-wave activity in the MLT region. Using temperature and wind fields from around the solar cycle for altitudes up to 74 km and only those from solar minimum for above 76 km (the next model level), the effects of changes in gravity-wave propagation through the stratosphere and lower mesosphere can be found, independently from the effects of changes in the MLT region. This also means that there is a discontinuity in the temperature and wind fields between 74 and 76 km, although this is small as the effects of the solar cycle on temperatures and winds are small compared with the background values at this altitude (see figure IV.17b). Figures (IV.21b and IV.22b) show the zonal wind variance and volume averaged gravity-wave drag using this calculation. The gravity-wave activity above 76 km is seen to vary with the changes in the conditions below this altitude, although these changes are not in good agreement with the results of the control calculation. The change in wind variance at 82 km shows a similar pattern to that seen in figure (IV.21a), but the clear positive correlations seen at 87 and 95 km are not seen. The volume averaged gravity-wave drag in figure (IV.22b) shows a positive correlation with solar activity, although the response is less than that seen in figure (IV.22a). While the STEVE model may underestimate changes in the stratosphere due to variable solar forcing (see chapter V), it is clear that this cannot be the only factor responsible for the solar cycle effects on gravity-wave activity in the MLT region seen in the model.

Using solar minimum data for levels up to 74 km and variable data for levels above 76 km, it is possible to show the impact of the *in situ* temperature and wind changes on gravity-wave activity in the 76 – 95 km region. Figures (IV.21c and IV.22c) show the gravity-wave activity in this region using this calculation. Both the height profiles of the zonal wind variance and the volume averaged gravity-wave drag are very similar to the results of the control test (figures IV.21a and IV.22a). This shows that the

changes in the temperature and winds in the MLT region are an important factor in producing the response of gravity-wave activity to the solar cycle in the STEVE model. Comparing this agreement to the effect of changes below 74 km, it can be seen that the variations in wind and temperatures in the 76 – 95 km region are more important for producing the response of gravity-wave activity to the solar cycle seen in the STEVE model than the variations below 74 km. This is consistent with the FV87 theory of the semi-annual cycle in gravity-wave activity in this region, which states that the *in situ* winds are the controlling factor. This apparent dependence may also be a consequence of the general increase with height of the temperature and wind perturbations due to variable solar UV flux.

Comparing the effects of variable temperature and wind fields above 76 km and below 74 km to the control test results also highlights a further important point. Variations in the background atmosphere in both height ranges is seen to have an impact on gravity-wave activity in the MLT region, but the effect of changes in the whole atmosphere is not equal to the sum of the effects due to variations in the two height ranges. This is because, in the ray-tracing gravity wave parameterisation, both the history of the ray-path and the current conditions are important to the gravity wave calculation at each height.

The effects of variations in the *in situ* temperatures and winds can now be separated, to see if either one of these is the dominant factor. Figures (IV.21d and IV.22d) show the gravity-wave activity calculated when only the temperatures above 76 km vary with the solar cycle and figures (IV.21e and IV.22e) show the gravity-wave activity calculated when only the wind data above 76 km is allowed to vary with the solar cycle. Comparing both sets of results to figures (IV.21c and IV.22c), it can be seen that both variable winds and temperatures can produce responses the vertical structure in the gravity-wave activity in the MLT region which is similar to the combined *in situ* variable temperature and wind results. However, the volume averaged change in extra-tropical gravity-wave drag in figure (IV.22d) is inconsistent with other results. This can be explained by considering the effect of the change in temperature in terms of the linear theory of gravity waves.

It is worth noting that the increase in temperature with solar activity in the MLT region seen in the model may be expected to decrease the buoyancy frequency and hence the rate of growth of gravity waves in this region, as in the summer stratosphere. However, as this region is one of wave dissipation, this decrease in the growth rate leads to less gravity-wave dissipation close to 76 km. This can be shown by plotting the latitude-height structure of the change in gravity-wave drag in this region (figure

IV.23). Between $\sim 75 - 90$ km, there is a decrease in the gravity-wave drag in the Northern Hemisphere (as shown in figure IV.22d), corresponding to a decrease in the wave dissipation in this region. The slower growth rate of gravity waves leads to more waves propagating up to altitudes over 90 km, where there is a pronounced increase in the gravity-wave drag. In this case, figure (IV.21d) is unrepresentative of the changes occurring throughout the Northern Hemisphere as a whole.

In both the height structure and volume averaged plots, the variable wind results show the closest agreement to the combined variable wind and temperature results. This suggests that while the *in situ* temperature variations play a part, the *in situ* wind variations are the most important factor in creating the gravity wave response to the solar cycle in the STEVE model. While it is worth mentioning that there is no consensus on the relationship between the horizontal winds in the middle and upper mesosphere and solar activity (discussed below), the theory postulated here is complimentary to the theory of FV87 which states that the *in situ* winds are the dominant factor in controlling long-term changes in gravity-wave activity in this region.

Changes in the mean horizontal winds have been observed simultaneously with changes in gravity-wave activity in the mesosphere. Figure (IV.24) shows the interannual changes in the mean zonal and meridional winds measured at Collm, after Gavrilov *et al.* (2002). As with the zonal wind variances, different trends are seen at different heights, but there is some evidence of variations in the horizontal winds with the solar cycle (measurements from the radars at Shigaraki and Saskatoon show similar trends). It is possible that these observed wind variations are responsible, at least in part, for the variations in gravity-wave activity seen at these locations. However, it is worth noting that measurements of changes in the horizontal winds with solar activity at different locations do not agree (e.g. Gavrilov *et al.*, 2002). As with the gravity-wave activity observations, this may be interpreted as a macro-regional response of the atmosphere to variable solar forcing.

Plotting the latitude-height structure of the change in temperature and zonal wind between solar minimum and maximum calculated by the STEVE model for the June solstice (figures IV.25a and IV.25b), regions of increases and decreases are observed. These have similar horizontal and vertical scales to the macro-regions of change in the zonal wind variance with solar activity shown in figure (IV.19), although an exact match is not seen. In the extra-tropical MLT region, only the zonal winds show both increases and decreases, which reinforces the importance of this parameter in producing the changes in gravity-wave activity seen in this region in the STEVE model. The existence of such changes in the zonal winds explains the geographical

dependence in the changes in gravity-wave activity with solar activity seen in the STEVE model.

Using the full STEVE model, it is possible to clarify the origin of macro-regions seen in figures (IV.25a and IV.25b). By performing a second 21 year simulation of the STEVE model, using the Rayleigh friction approximation to gravity-wave drag, it is possible to determine whether these macro-regions are a natural consequence of the variable UV absorption in the model, or are created by the variable interaction between the MK2000 gravity wave parameterisation and the background atmosphere. Figures (IV.26a and IV.26b) show the changes in temperature and zonal winds between solar maximum and minimum around the June solstice calculated by this STEVE model simulation. These results are substantially different from those obtained with the STEVE model using the MK2000 parameterisation. In the equatorial, middle atmosphere, the instability seen in the wind and temperature data has been reduced to almost zero when Rayleigh friction is used. This can be explained by either the damping of the instability by Rayleigh friction, which acts to damp all winds to zero, or by a possible amplification of the instability by the MK2000 parameterisation, similar to the interaction between the parameterised gravity-wave drag and the diurnal tide in this region which will be discussed in chapter (VI), or some combination of the two. In the extra-tropical MLT region, the small-scale patches seen in changes in the winds when the MK2000 parameterisation is used are absent when Rayleigh friction is used. The hemispheric scale changes in the zonal wind are similar for the two model simulations, as these are the changes in the global-scale thermal winds due to the global-scale increase in the temperature of the MLT region. Again, these changes are either amplified when the MK2000 gravity wave parameterisation is used, or damped when Rayleigh friction is used, or a combination of the two.

Having shown that the macro-regional response of winds in the MLT region is linked to the nonlinear gravity wave response to variable background conditions, it is now possible to propose a mechanism to explain the results presented in this section. Changes in the solar UV flux lead to changes in the diabatic heating rate of the atmosphere which in turn leads to changes in the temperature of the MLT region. This would, in the absence of nonlinear gravity-wave drag, take the form of figure (IV.26a). These changes in the temperature alter gravity-wave activity and horizontal winds in this region (similar to figure IV.26b), which leads to further changes in the gravity-wave drag and horizontal winds. Even though the changes in the background atmosphere only vary on hemispheric scales, the nonlinear response of the gravity waves to these changes means that the gravity-wave drag shows a macro-regional

response. The change in this drag feeds back into the background temperatures and winds, producing the STEVE model results.

IV.3.2 THE IMPACT OF CHANGES IN THE LOWER ATMOSPHERE

In the previous section, the effects of variations in the stratosphere were shown to be a contributing factor to the modelled response of gravity-wave activity to changes in solar activity, but not as important as interactions between gravity waves and changes in the background atmosphere in the MLT region. In this section, the impact of a possible change in gravity-wave sources in the lower atmosphere on gravity-wave activity in the middle atmosphere, which was cited by Gavrilov *et al.* (1995, 2002) and Bošková and Laštovička (2001) as the most likely mechanism to explain their observations, will be explored.

Using the simplified model outlined in section (IV.3), the effect of a variable gravity-wave source, in the absence of variations in the background atmosphere, can be calculated. In the absence of any observations of the manner in which the global scale lower atmosphere gravity-wave source varies with the solar cycle, an 11 year sinusoidal variation in the gravity-wave source amplitude was chosen. A 10 % variation in the amplitude of the source was used, as this is comparable to the changes observed and modelled in the mesospheric gravity-wave activity. However, it is worth noting that this is far larger than the observed changes in the lower atmosphere temperatures and winds associated with variable solar activity (e.g. Hoyt and Schatten, 1997).

Figures (IV.27 and IV.28) show the changes in gravity-wave activity in the MLT region calculated using this variable source strength and the solar minimum background atmospheric conditions. A similar response is seen to the variable gravity-wave source as to the changes in the background atmosphere (figures IV.21a and IV.22a). This indicates that variations in the lower atmosphere gravity-wave source, if it were sufficiently strong, could account for the observations detailed in section (IV.2.1). This result provides the first quantitative evidence that the theories proposed by Gavrilov *et al.* (1995, 2002) and Bošková and Laštovička (2001) could explain their observations, although the *in situ* changes in the MLT region and changes in the filtering of gravity waves in the stratosphere and lower mesosphere have also been shown to be plausible contributing factors.

IV.4 SUMMARY

In this chapter, both changes in the stratospheric and mesospheric conditions have been shown to be important in producing the solar cycle in gravity-wave activity in the MLT region in the STEVE model. The effect of variable solar activity on gravity waves in the middle atmosphere in the STEVE model has been shown to have a complex latitude height structure, and it has been postulated (both in this chapter and by references herein) that a longitudinal structure should also exist, although no explicit tests of this can be carried out in the zonally averaged STEVE model. The STEVE-3D model, introduced in the last chapter, provides the ability to study the gravity waves in a three dimensional model with a good representation of the dynamics of the polar stratosphere region. The following chapter will examine the interaction between gravity waves and planetary waves and the effect which this has on the Northern Hemisphere winter polar stratosphere.

Haigh (1996) has shown that variations in atmospheric ozone concentrations, associated with variable solar UV flux, can have important consequences on the dynamics of the middle atmosphere and Gavrilov *et al.* (1994) have shown that there was some correlation between the geomagnetic Ap index and gravity-wave activity in the mesosphere. Chapter (VI) will introduce the CMAT model, which includes both interactive chemistry and the impact of ion-neutral collisions on the middle atmosphere. This model will then be used to study impact that gravity waves have on the dynamics and chemical composition of the MLT region.

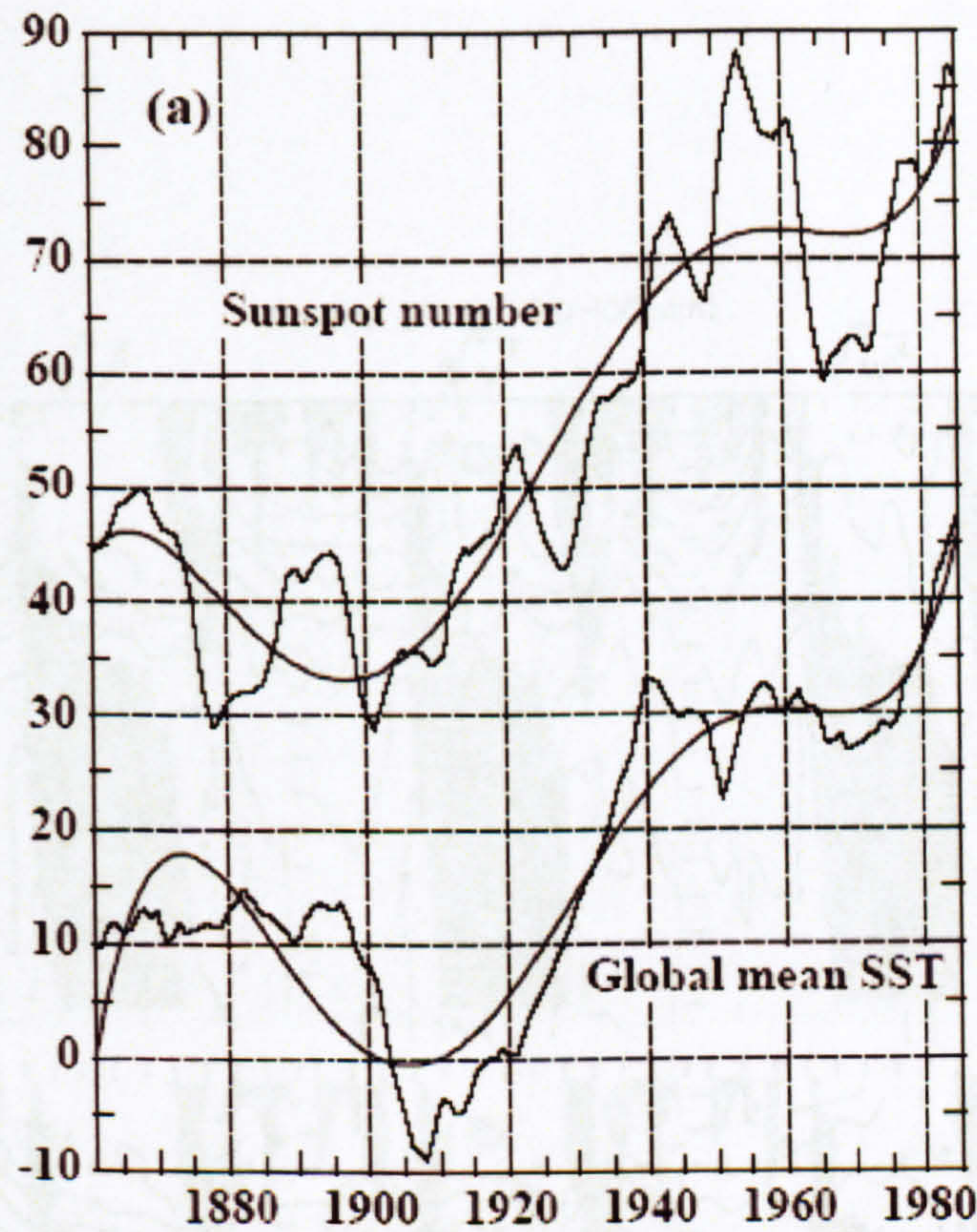


Figure IV.1, Eleven-year running mean of the annual sunspot numbers and the departure of the annual mean global sea-surface temperature from the 1951 – 1980 average, after Reid (2000). The temperature is in units of 0.01 K. The heavy line shows a least-squares 7th order polynomial fit to the data.

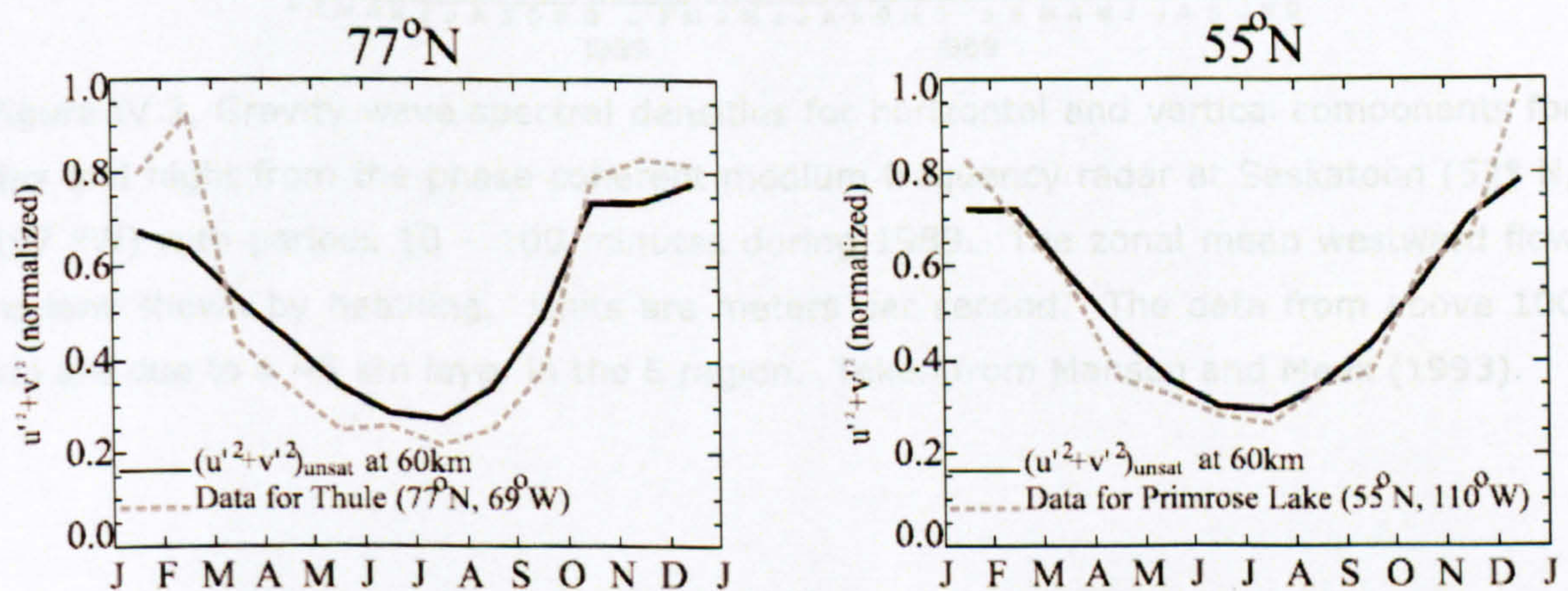


Figure IV.2, Seasonal variations observed in gravity-wave horizontal wind variance in rocket soundings (dashed lines) and the theoretical variation predicted by the amplitude saturation theory, after Eckermann (1995).

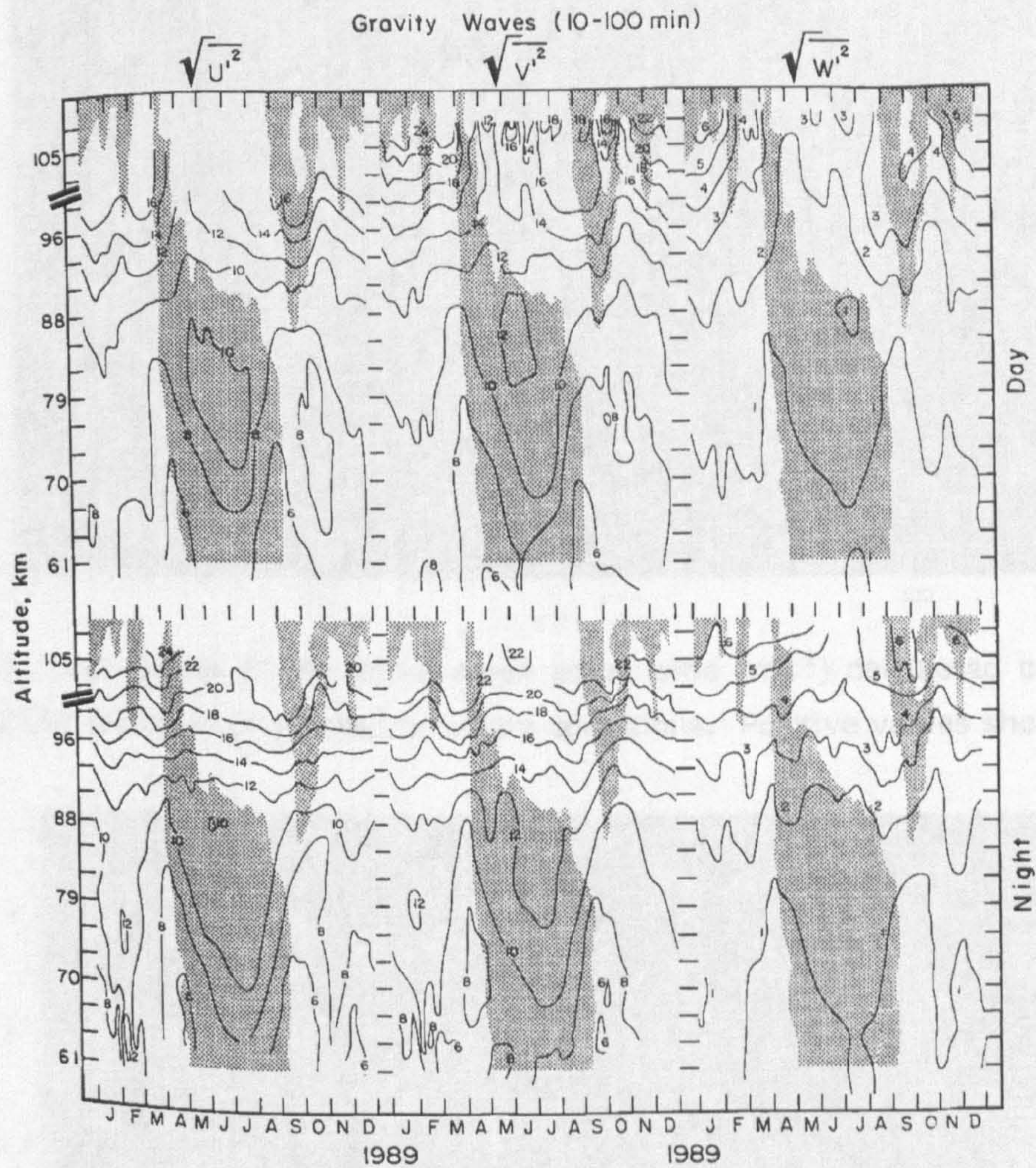


Figure IV.3, Gravity wave spectral densities for horizontal and vertical components for day and night from the phase coherent medium frequency radar at Saskatoon (52° N, 107 °W) with periods 10 – 100 minutes during 1989. The zonal mean westward flow regions shown by hatching. Units are meters per second. The data from above 100 km are due to a ~5 km layer in the E region. Taken from Manson and Meek (1993).

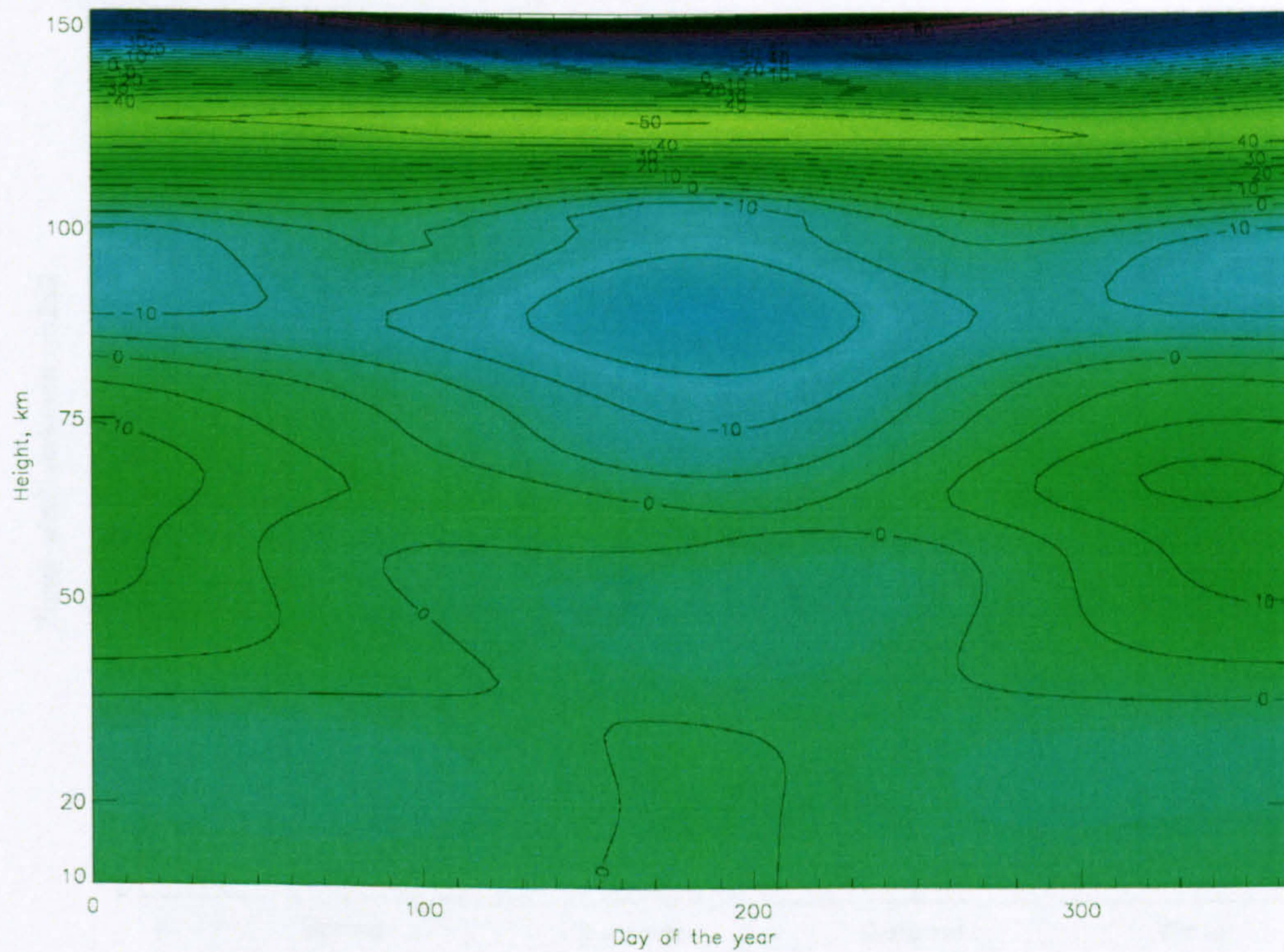


Figure IV.4, Time series of the zonal mean zonal wind (ms^{-1}) calculated by the HWM model at 67.5° North during solar minimum conditions. Positive values show eastward winds.

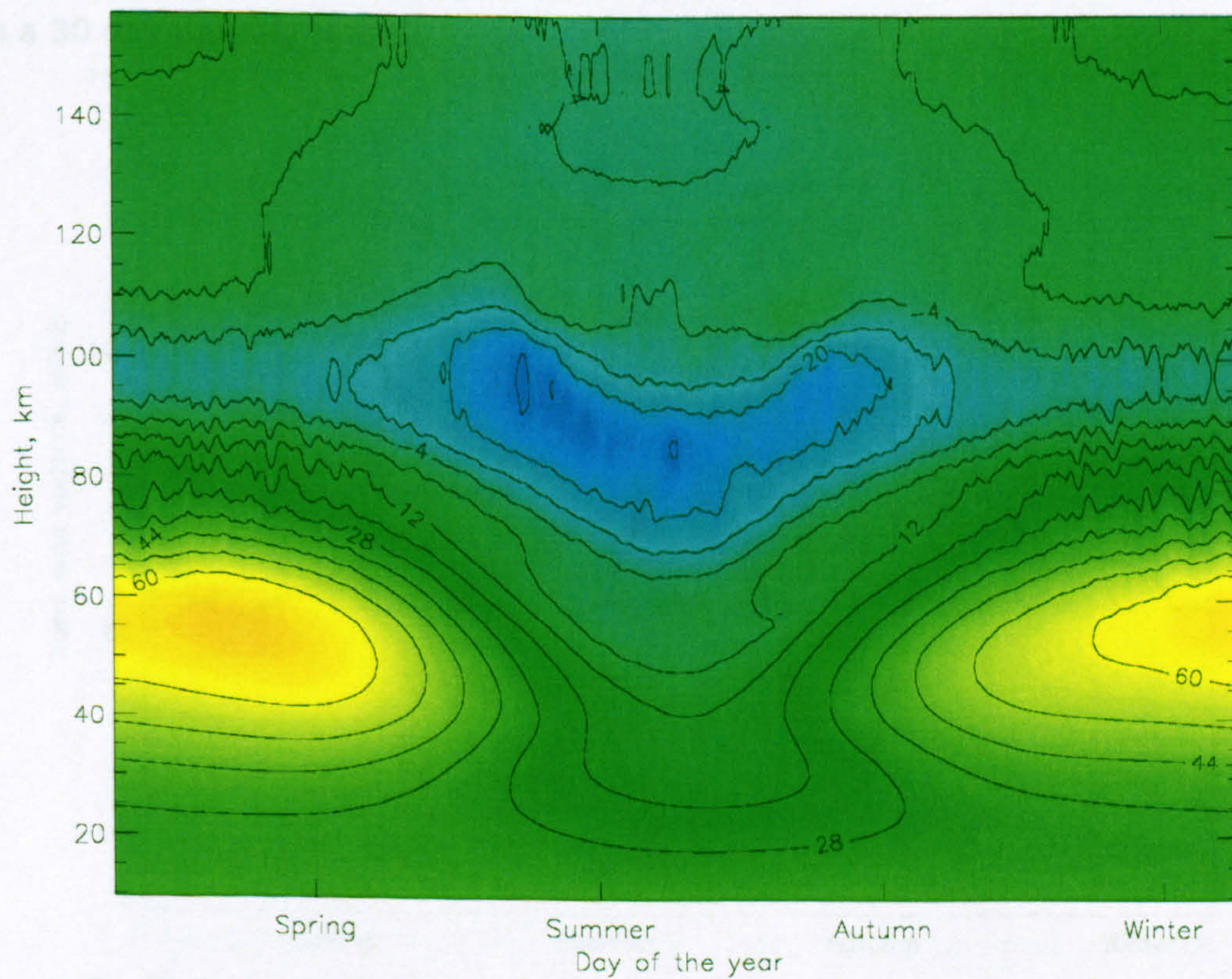


Figure IV.5, Time series of the zonal mean zonal wind (ms^{-1}) in the STEVE model at 67.5° N during solar minimum conditions.

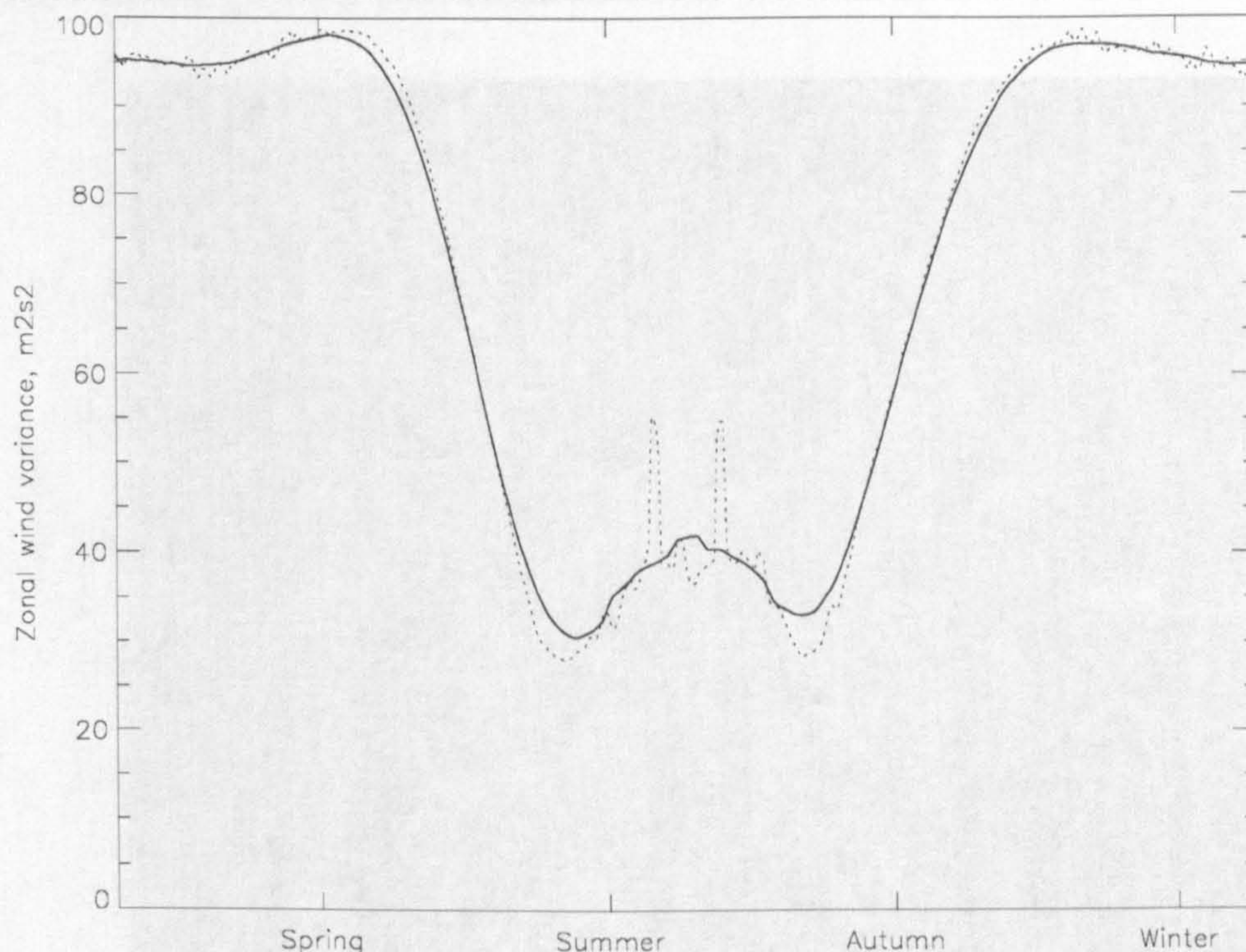


Figure IV.6, Time series of the zonal mean total horizontal wind variance produced by parameterised gravity waves in the STEVE model at 64 km altitude, 67.5° N during solar minimum conditions. The dotted line shows daily mean values and the solid line shows a 30 day running mean.

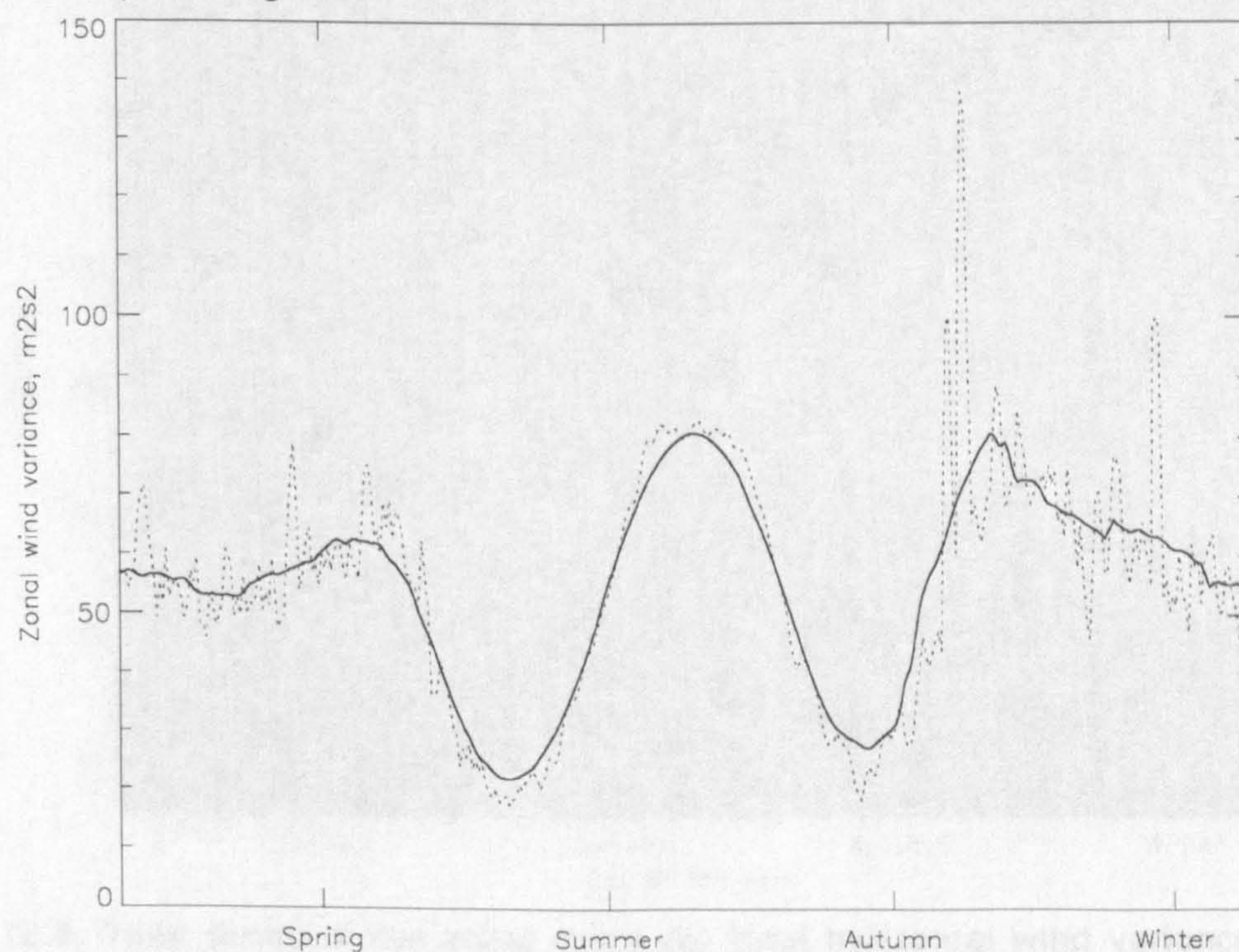
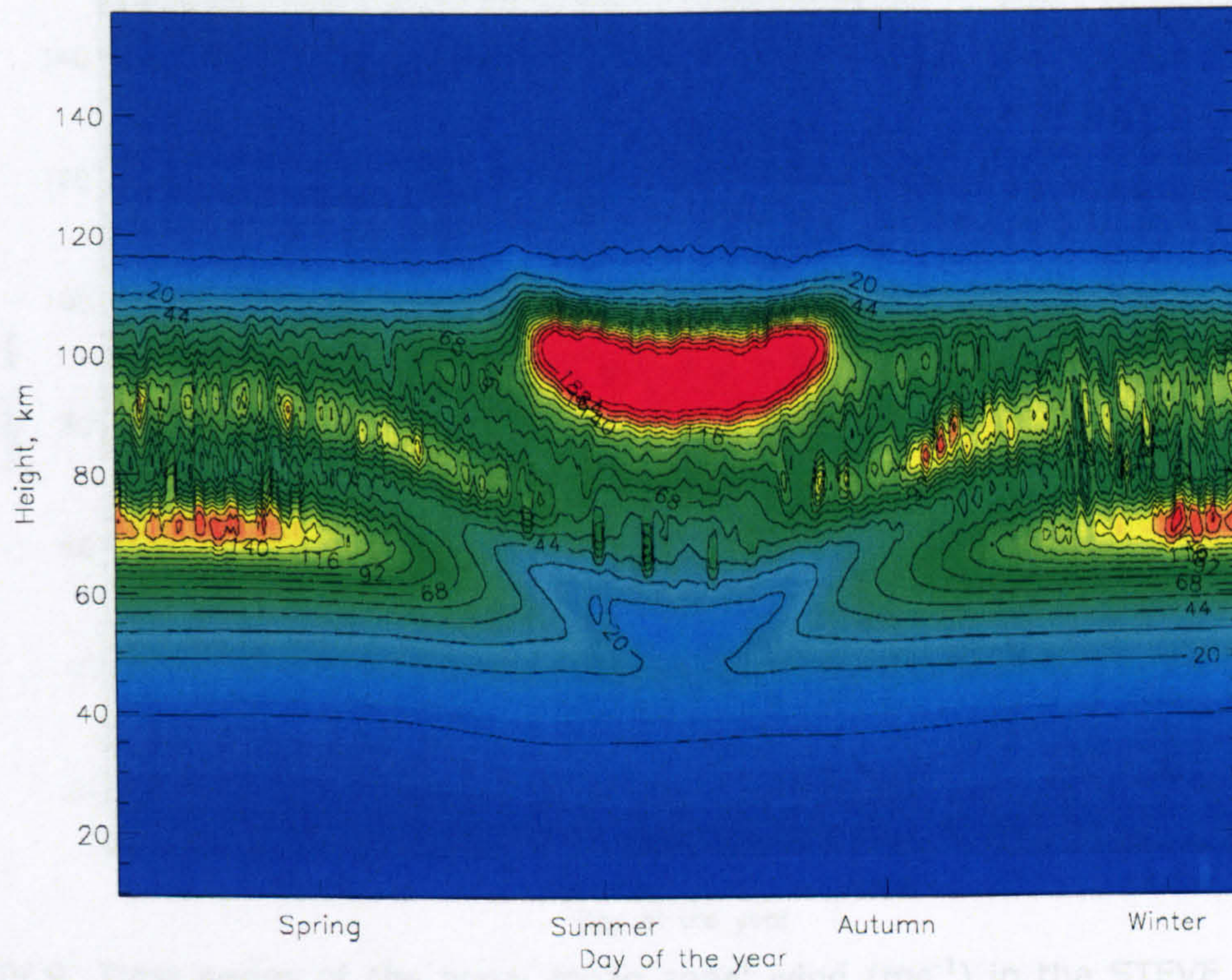


Figure IV.7, Time series of the zonal mean zonal wind variance produced by parameterised gravity waves in the STEVE model at 88 km altitude, 67.5° N during solar minimum conditions. The dotted line shows daily mean values and the solid line shows a 30 day running mean.

(a)



(b)



Figure IV.8, Time series of the zonal mean (a) total horizontal wind variance and (b) zonal wind variance (m^2s^{-2}) produced by parameterised gravity waves in the STEVE model at 67.5° N during solar minimum conditions.

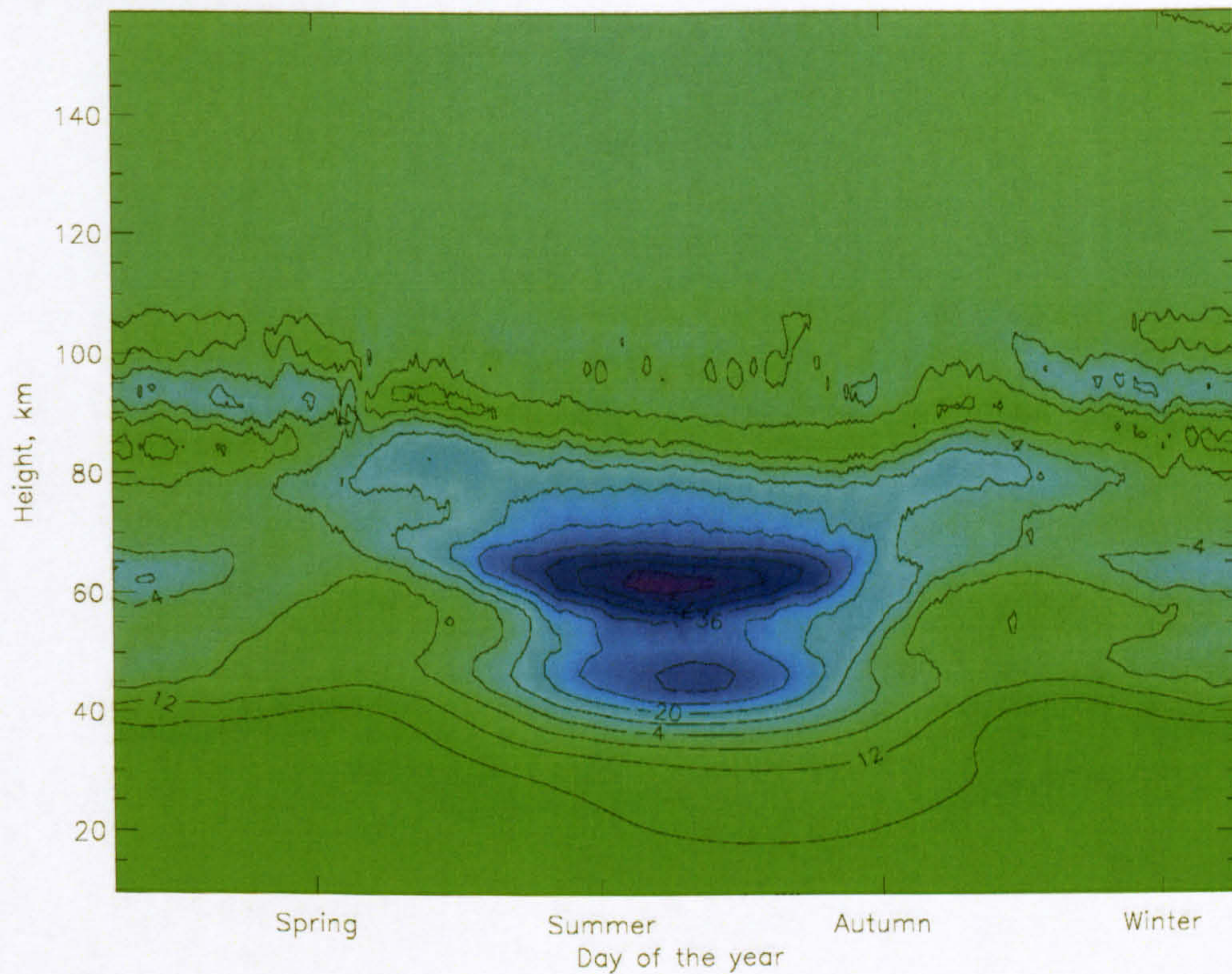


Figure IV.9, Time series of the zonal mean zonal wind (ms⁻¹) in the STEVE model at 2.5° N during solar minimum conditions.

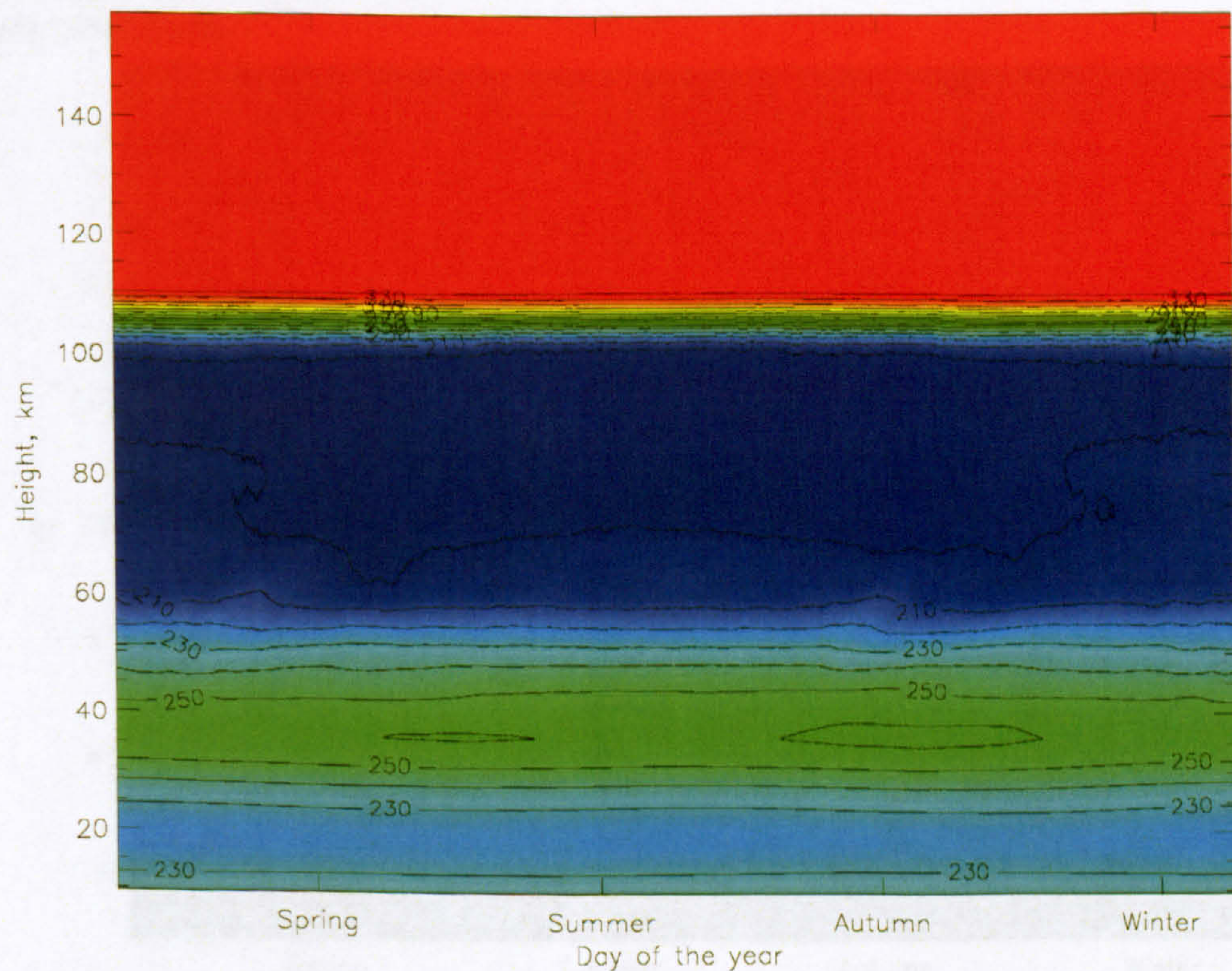


Figure IV.10, Time series of the zonal mean temperature (K) in the STEVE model at 2.5° N during solar minimum conditions.

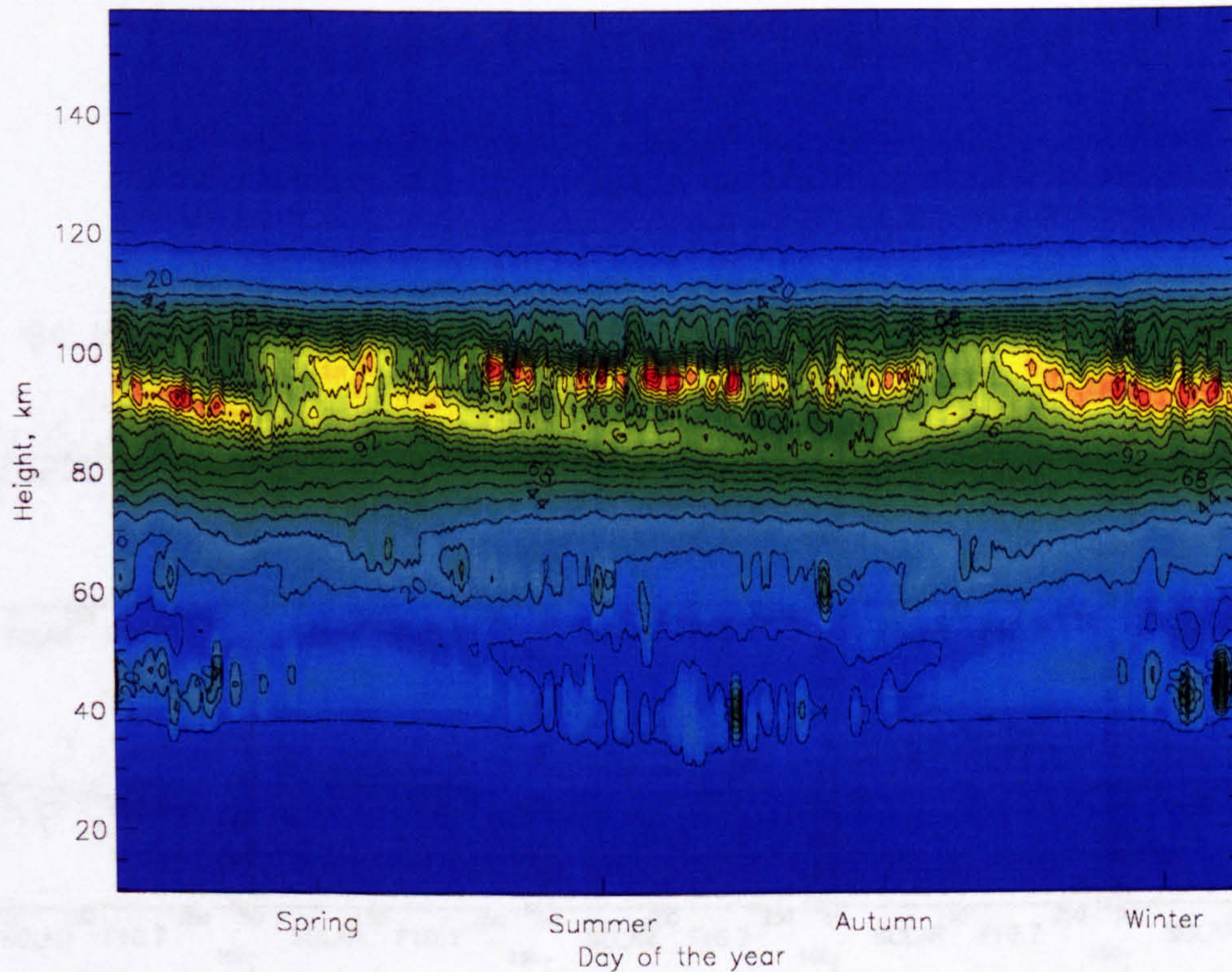


Figure IV.11, Time series of the zonal mean total horizontal wind variance (m^2s^{-2}) produced by parameterised gravity waves in the STEVE model at 2.5°N during solar minimum conditions.

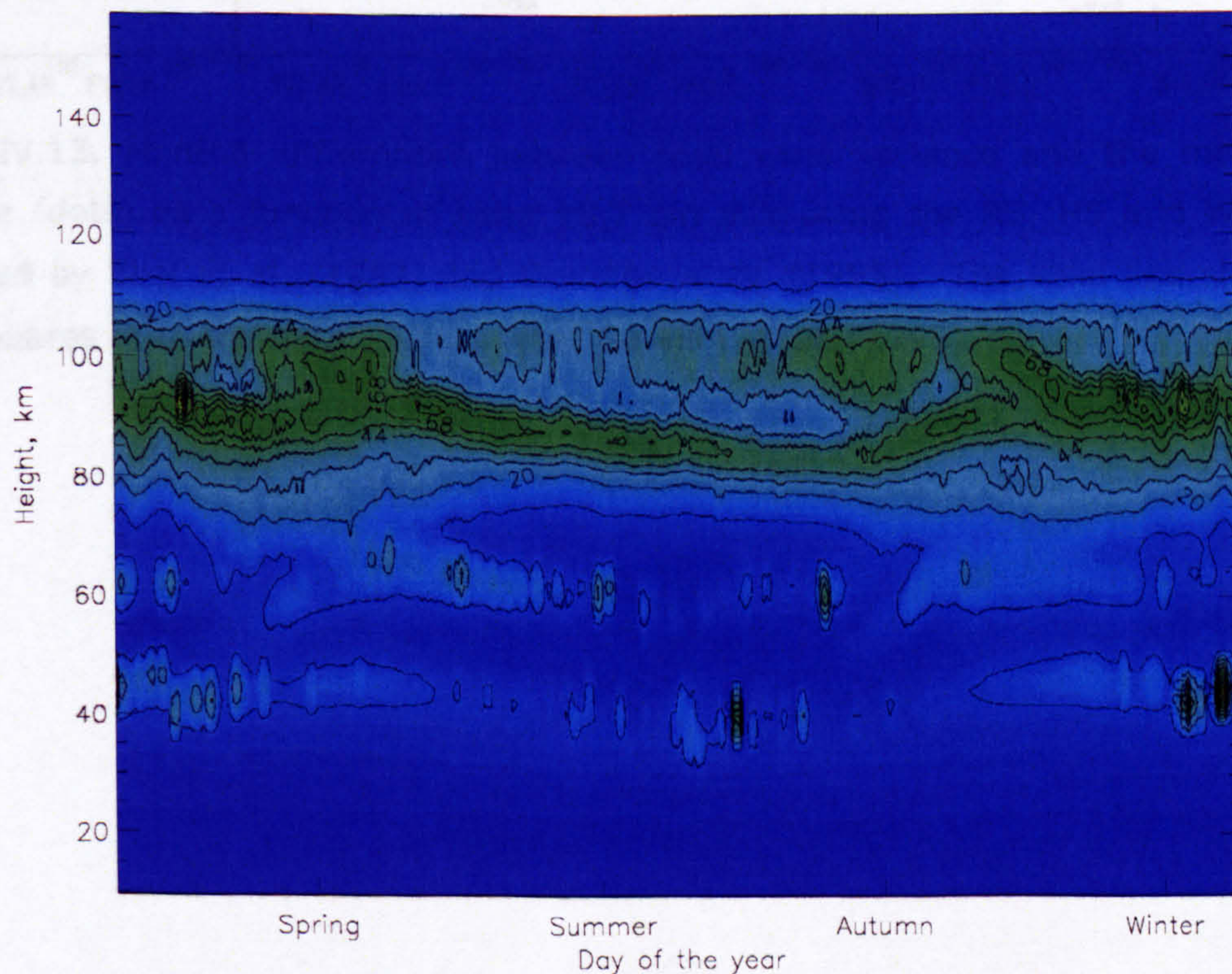


Figure IV.12, Time series of the zonal mean zonal horizontal wind variance (m^2s^{-2}) produced by parameterised gravity waves in the STEVE model at 2.5°N during solar minimum conditions.

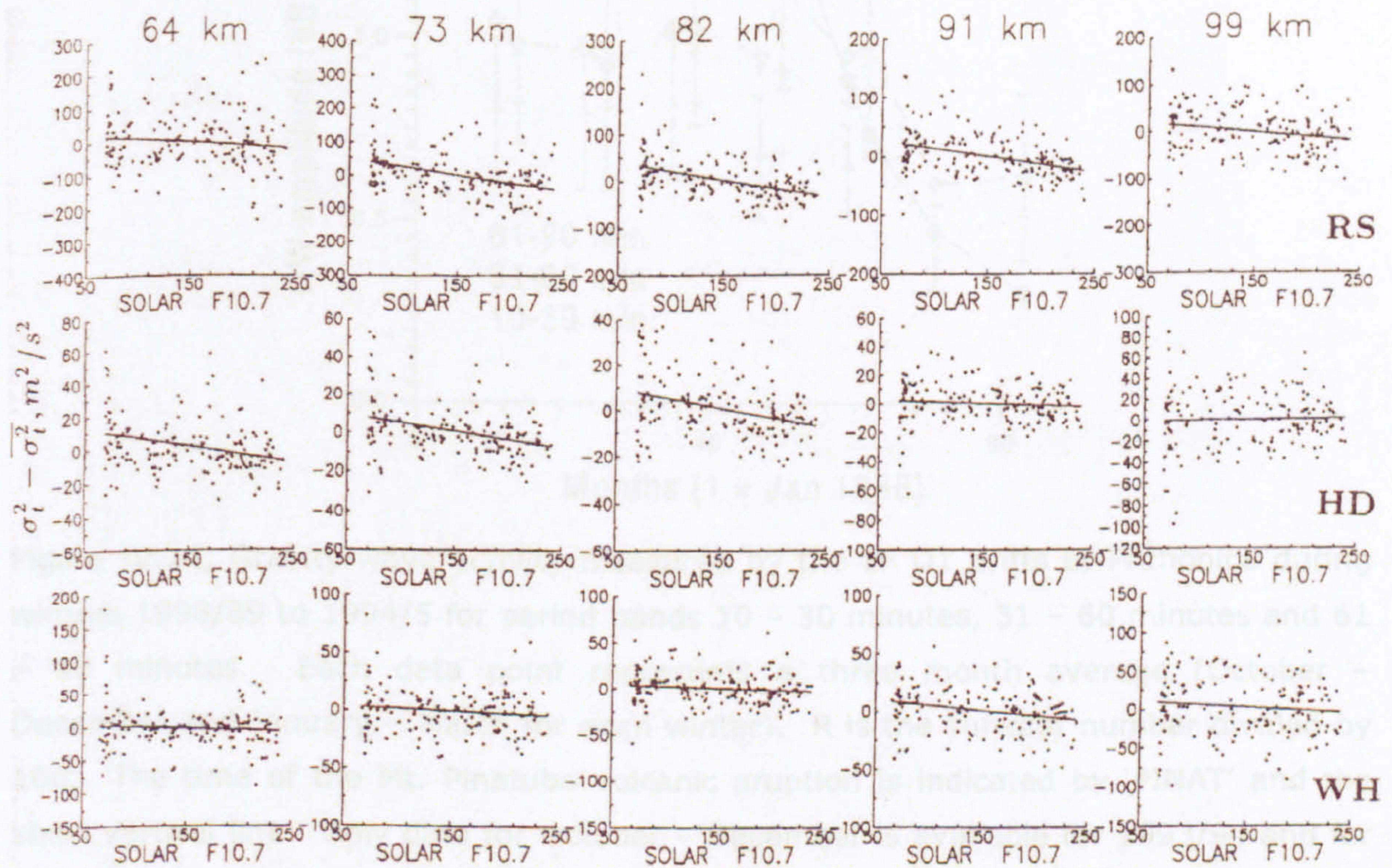


Figure IV.13, Monthly differences between total wind variance and the mean wind variance (dots) as a function of solar 10.7 cm flux using the RS, HS and WH filters described by Ebel *et al.* (1987) and Gavrilov *et al.* (1995). The solid line shows the least squares fit to the data. Figure after Gavrilov *et al.* (1995).

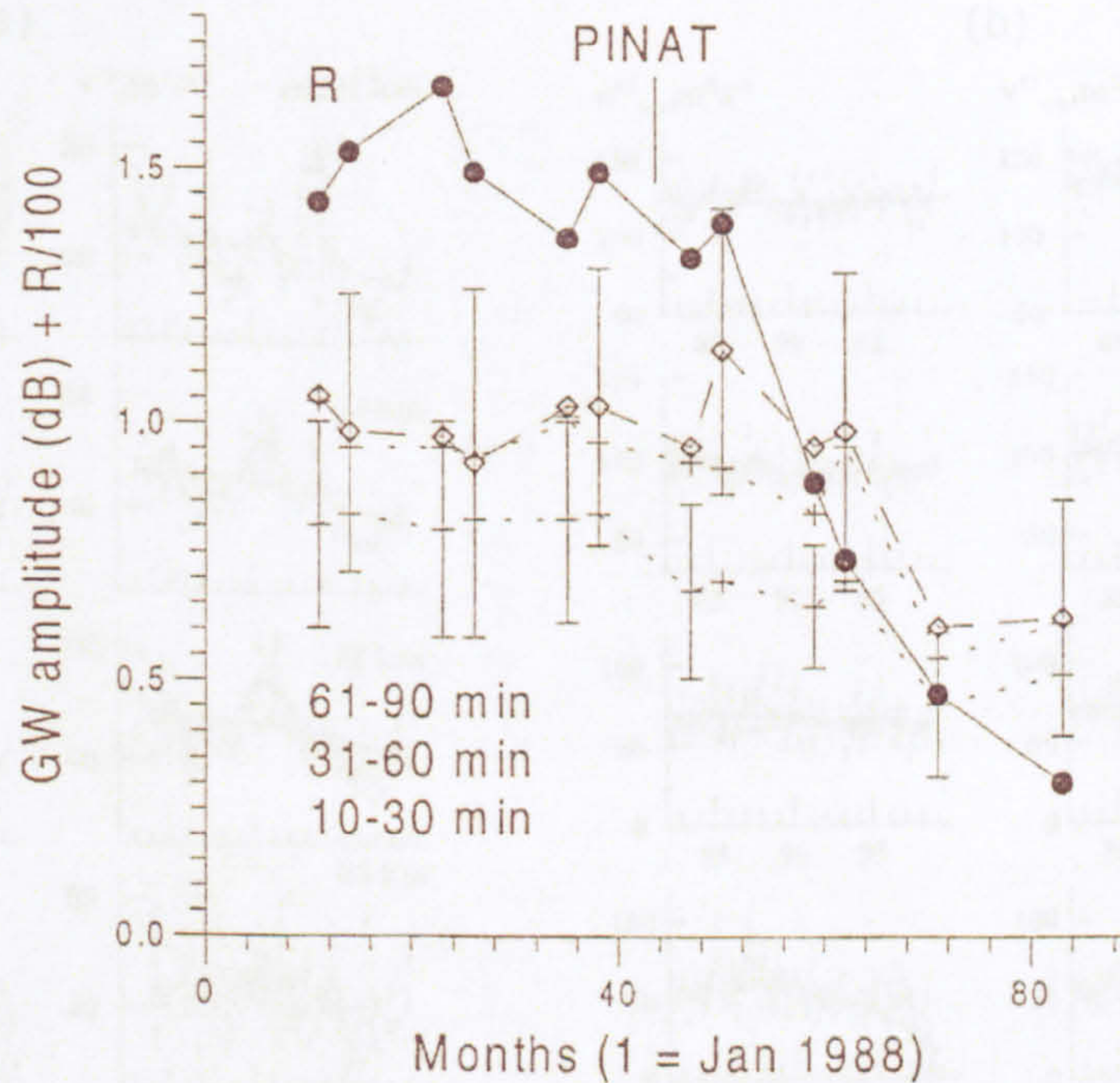


Figure IV.14, Gravity-wave activity measured by the LF D1 drifts at Průhonice during winters 1998/89 to 1994/5 for period bands 10 – 30 minutes, 31 – 60 minutes and 61 – 90 minutes. Each data point represents a three month average (October – December and January – March for each winter). R is the sunspot number divided by 100. The time of the Mt. Pinatubo volcanic eruption is indicated by 'PINAT' and the short vertical line. Only data for October – December is available for 1993/94 and for October – November for 1994/5. The standard deviations of the individual data included in each 3 month average is shown on certain points. After Laštovička (1999).

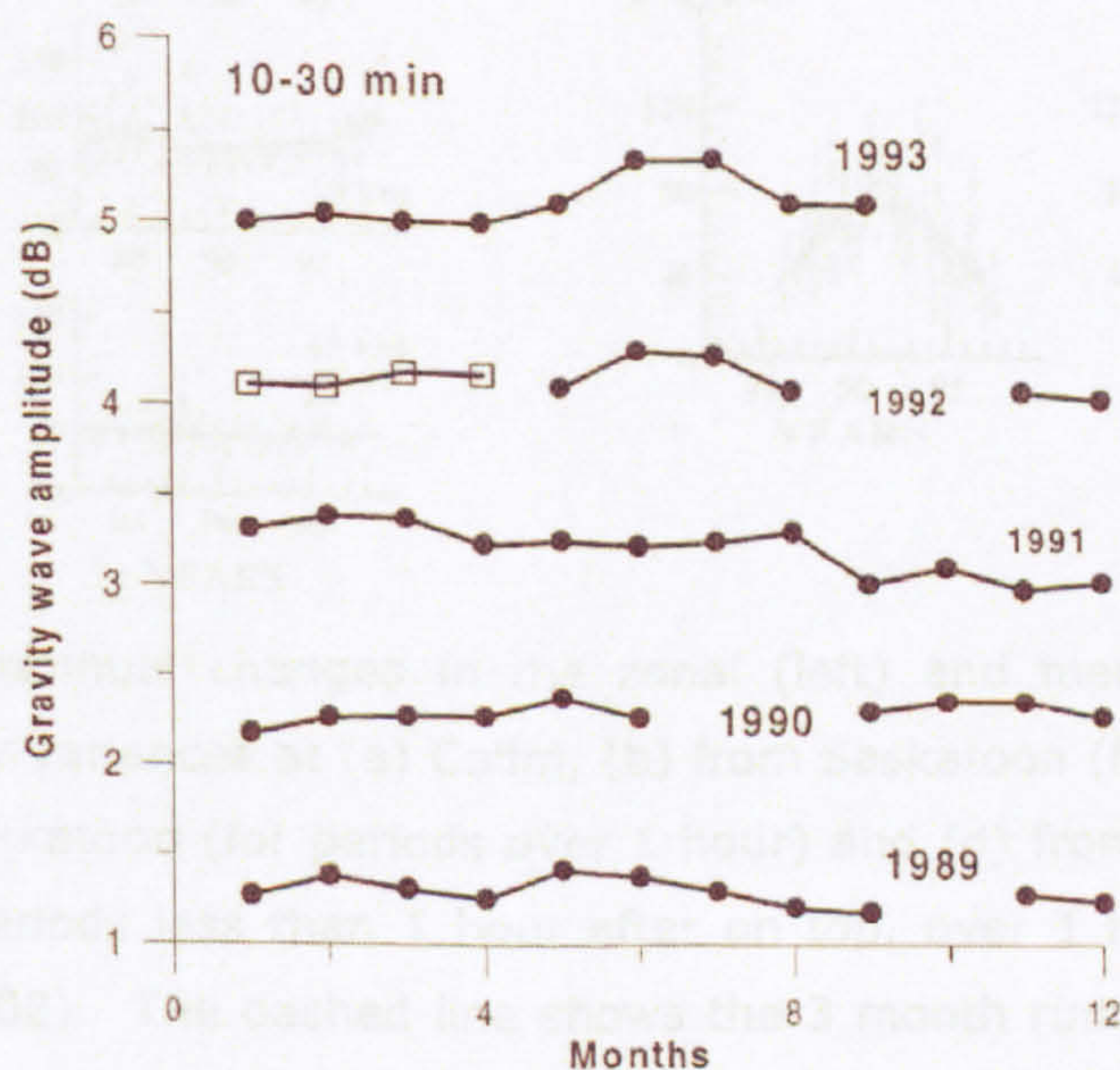


Figure IV.15, Gravity wave maximum amplitudes for 1989 – 1993 in the period band 10 – 30 minutes from the LF D1 drifts at Průhonice. The squares are computed from full – night data, while the dots are computed from 5 - hour data sets. Each year is offset from the preceding one to avoid overlap. From Bošková and Laštovička (2001).

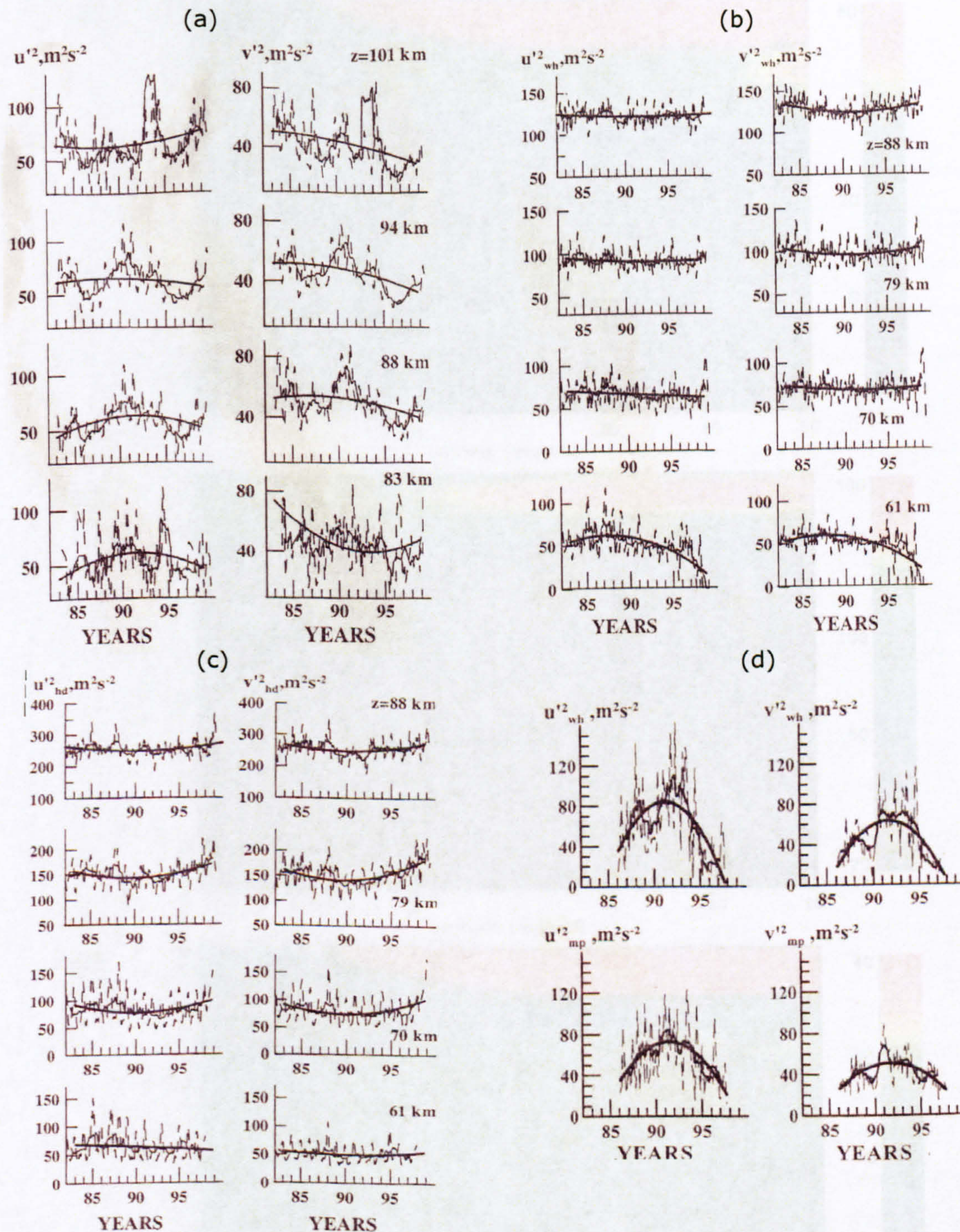


Figure IV.16, Interannual changes in the zonal (left) and meridional (right) short-period drift velocity variances at (a) Collm, (b) from Saskatoon (for periods less than 1 hour), (c) from Saskatoon (for periods over 1 hour) and (d) from Shigaraki (averaged over 65-80 km, periods less than 1 hour after on top, over 1 hour on bottom) after Gavrilov *et al.* (2002). The dashed line shows the 3 month running average, the thin solid line shows the 12 month running average and the thick solid line shows the quadratic fit to the data.

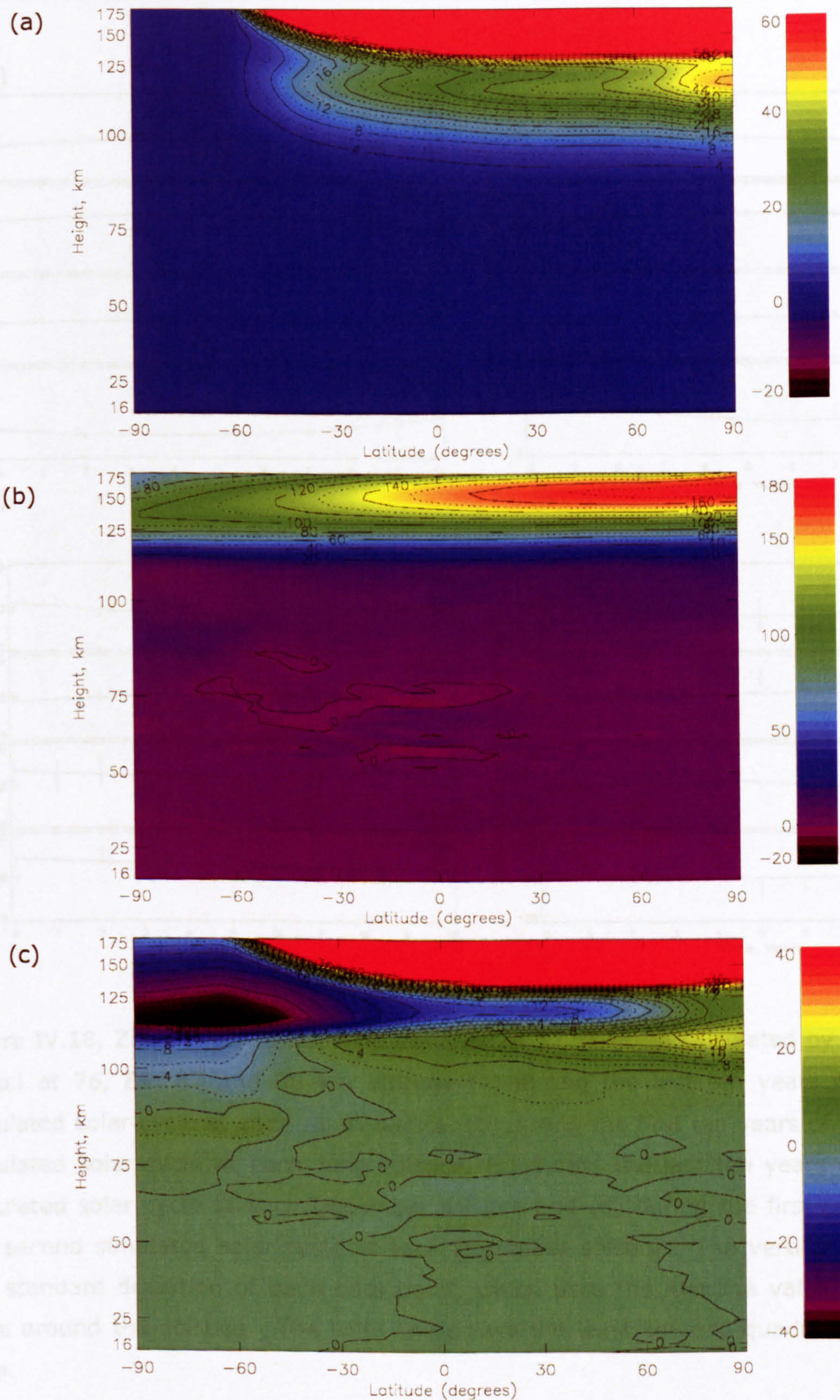


Figure IV.17, Zonal mean change in the (a) heating rate due to atmospheric absorption of solar UV (Kday^{-1}) at, (b) temperature (K) averaged over 30 days around the June solstice and (c) net diabatic heating rate (Kday^{-1}) at the June solstice between solar minimum and maximum, calculated by the STEVE model.

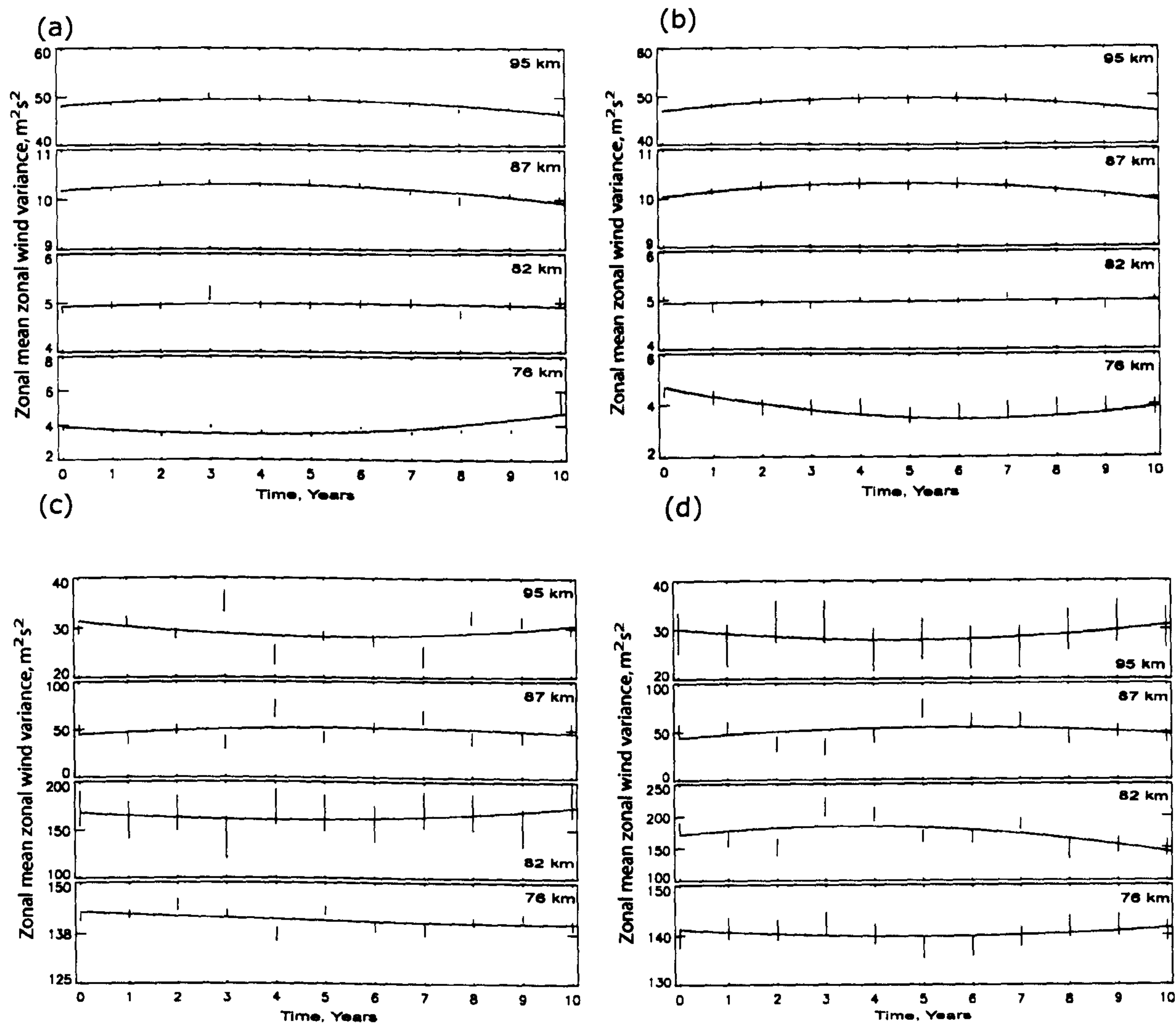


Figure IV.18, Zonal mean zonal wind variance at 52.5° North calculated by the STEVE model at 76, 82, 87 and 95 km altitude (a) during the last ten years of the first simulated solar cycle at each June solstice, (b) during the first ten years of the second simulated solar cycle at each June solstice, (c) during the last ten years of the first simulated solar cycle at each December solstice and (d) during the first ten years of the second simulated solar cycle at each December solstice. The vertical lines show the standard deviation of each data point, which uses the average value over thirty days around the solstice. The solid line shows the least squares quadratic fit to the data.

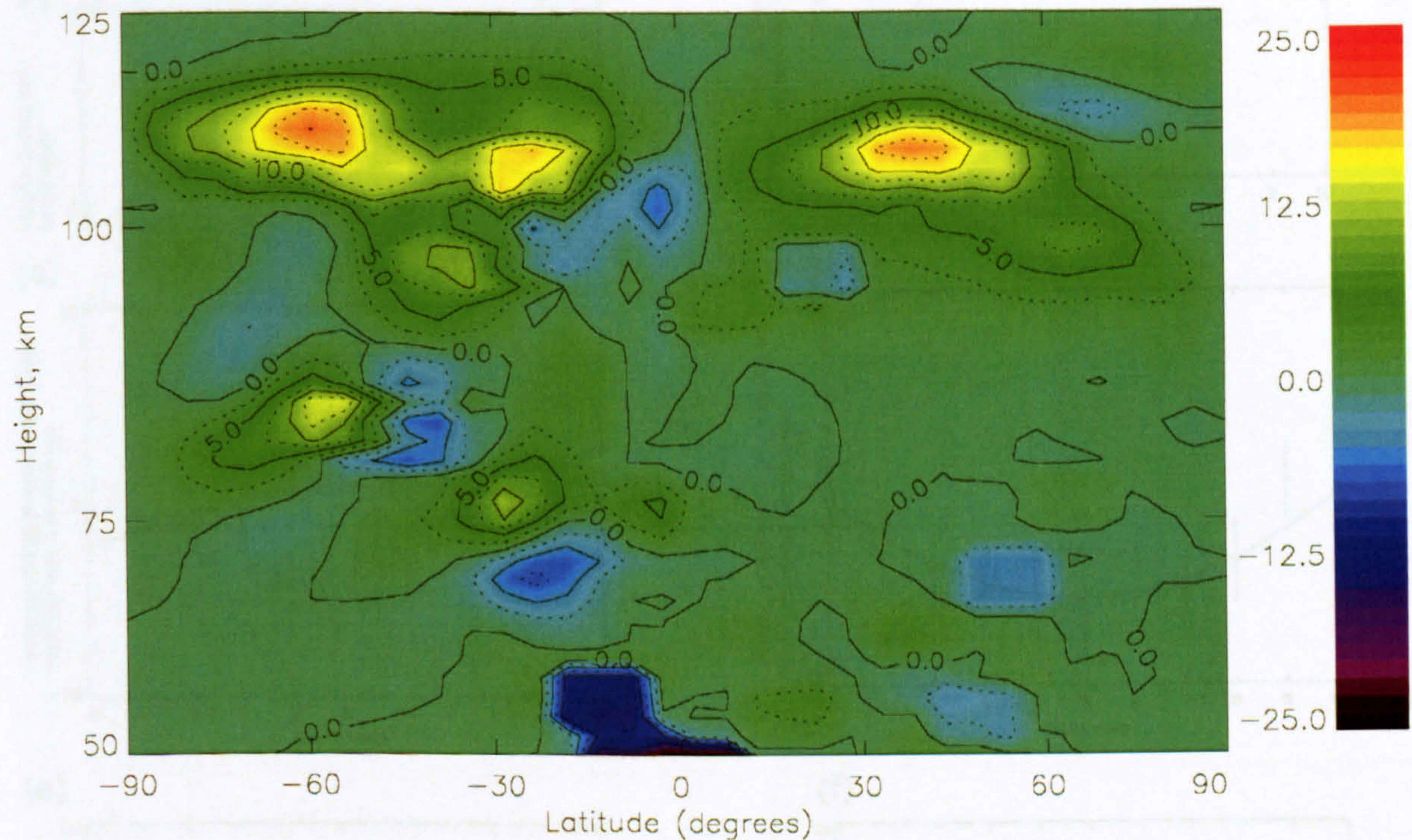


Figure IV.19, Change in the zonal mean zonal wind variance (m^2s^{-2}) from solar minimum to solar maximum calculated by the STEVE model for the June solstice. Data taken from the second simulated solar cycle and is averaged over thirty days around the solstice.

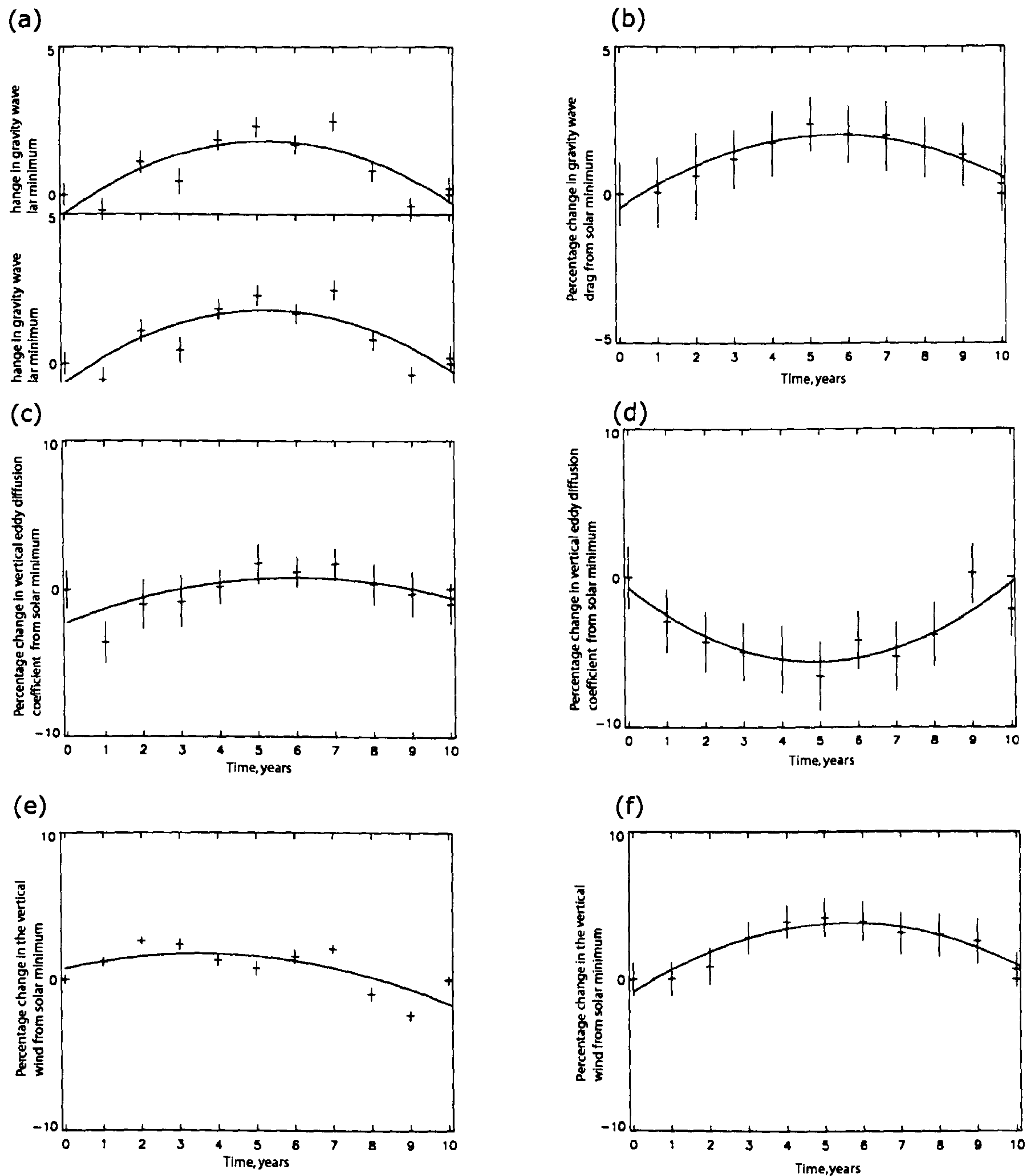


Figure IV.20, Percentage change from solar minimum in the zonal mean (a) zonal gravity-wave drag at the June solstice, (b) zonal gravity-wave drag at the December solstice, (c) thermal eddy diffusion coefficient at the June solstice, (d) thermal eddy diffusion coefficient at the December solstice, (e) vertical wind at the June solstice and (f) vertical wind at the December solstice calculated by the STEVE model during the first ten years of the second simulation. The drags are averaged over the extra-tropical northern hemisphere between 76 – 90 km altitude and over thirty days around each solstice. The vertical lines show the standard deviation. The solid line shows the least squares quadratic fit to the data.

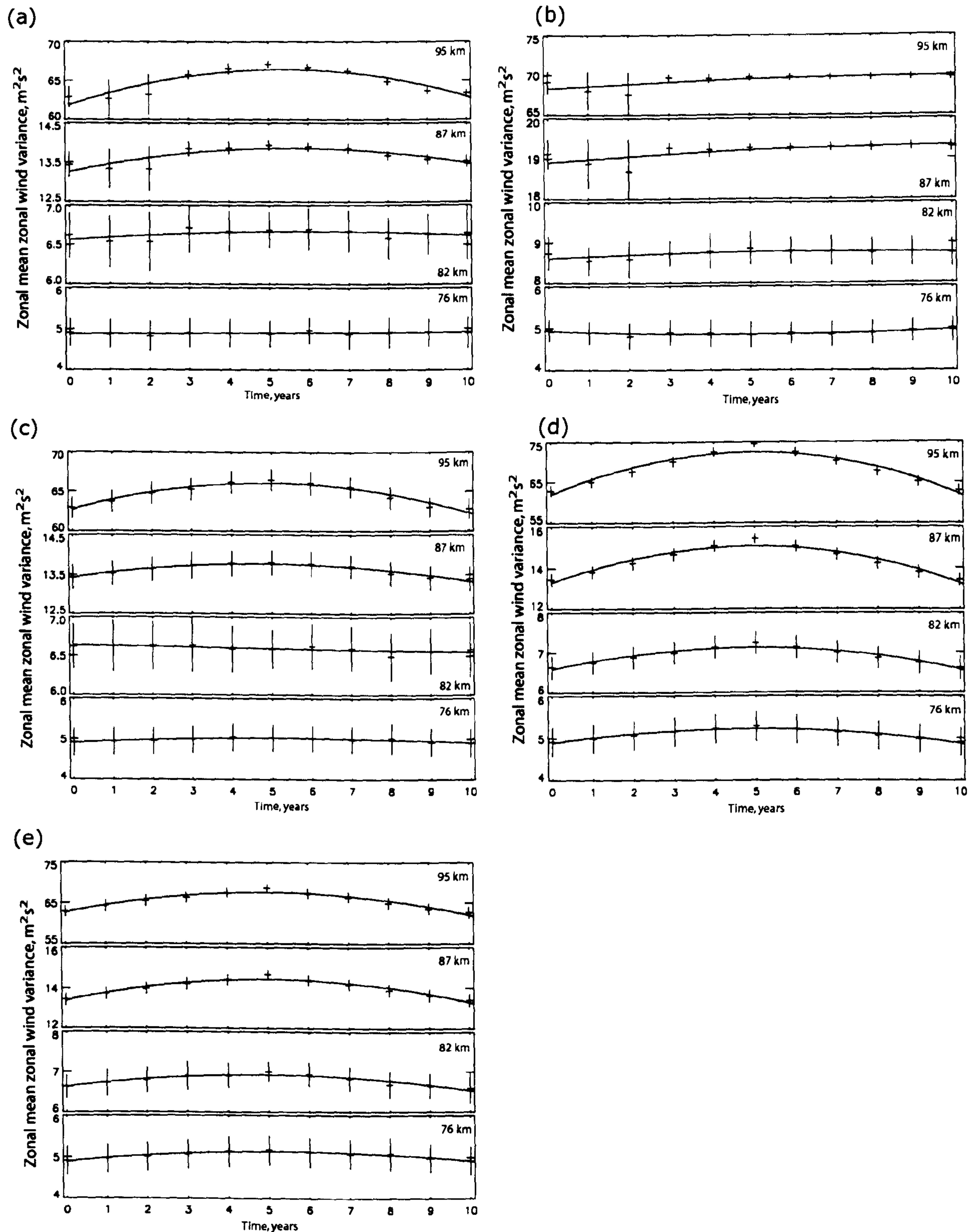


Figure IV.21, As figure (IV.18) but using the data calculated using daily average data from the STEVE model at each June solstice during the first ten years of the second simulated solar cycle for (a) variable conditions at all heights, (b) variable conditions below 74 km only, (c) variable conditions above 74 km only, (d) variable winds at all heights and variable temperature above 74 km only and (e) variable temperatures at all heights and variable winds above 74 km only.

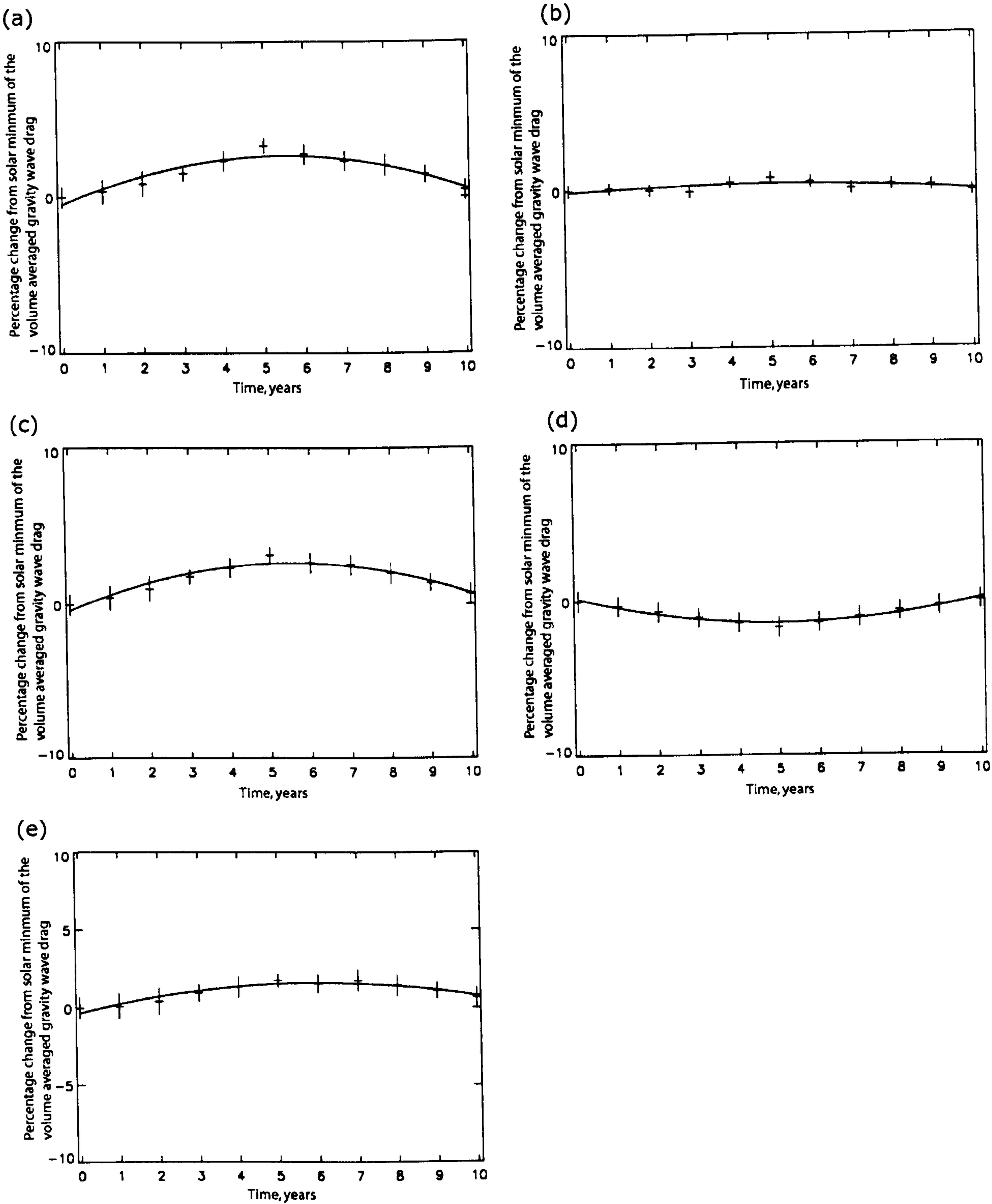


Figure IV.22, As figure (IV.21) but for zonal mean zonal gravity-wave drags averaged over the extra-tropical northern hemisphere between 76 – 90 km altitude.

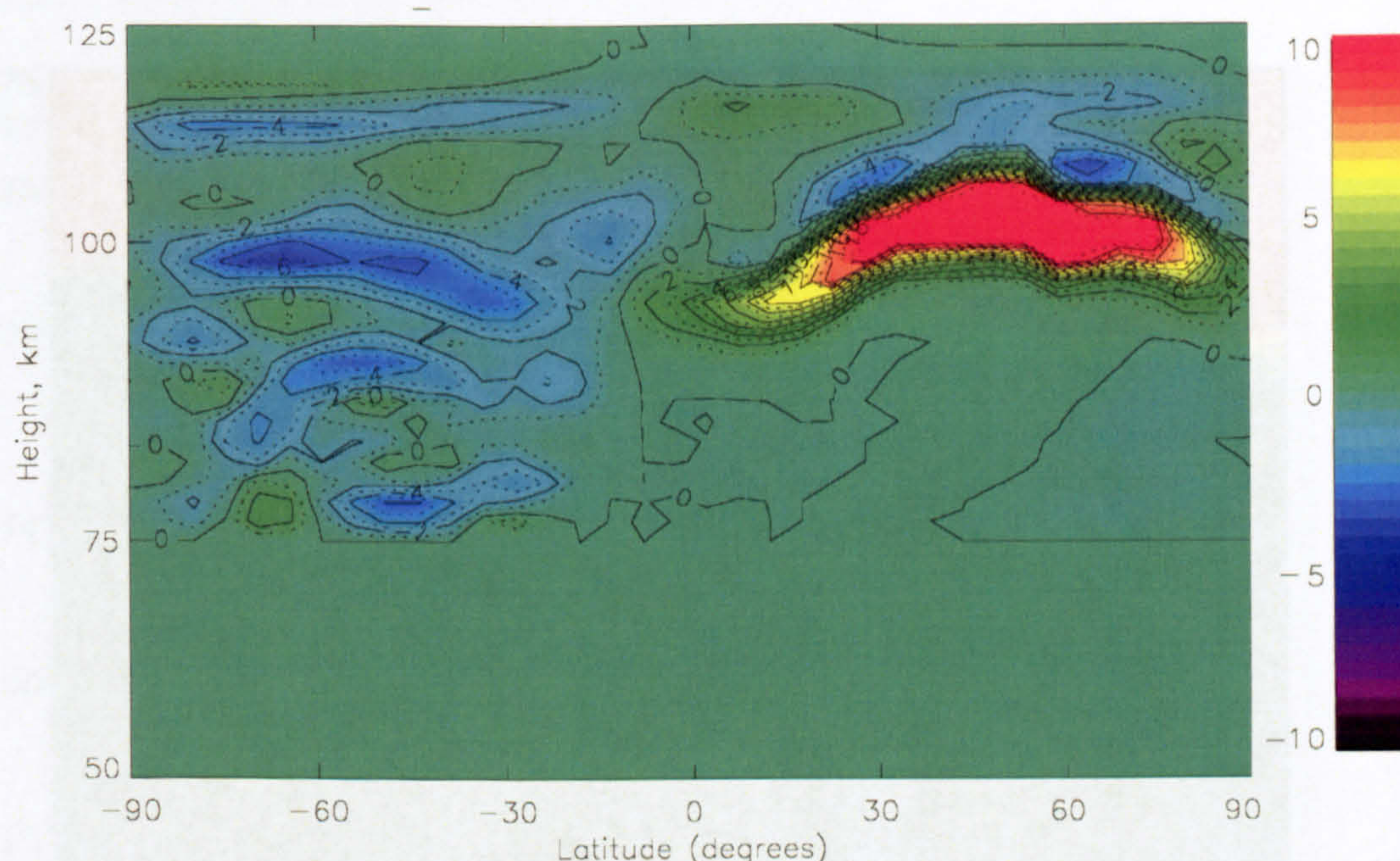


Figure IV.23, Zonal mean change in the zonal gravity-wave drag (m/s/day) between solar maximum and minimum at the June solstice calculated using daily averaged data from the STEVE model during the first ten years of the second simulation. The drags are averaged over thirty days around the June solstice. Temperature data at all heights and wind data above 76 km are for variable solar conditions while for 74 km and below, perpetual solar minimum wind values are used.

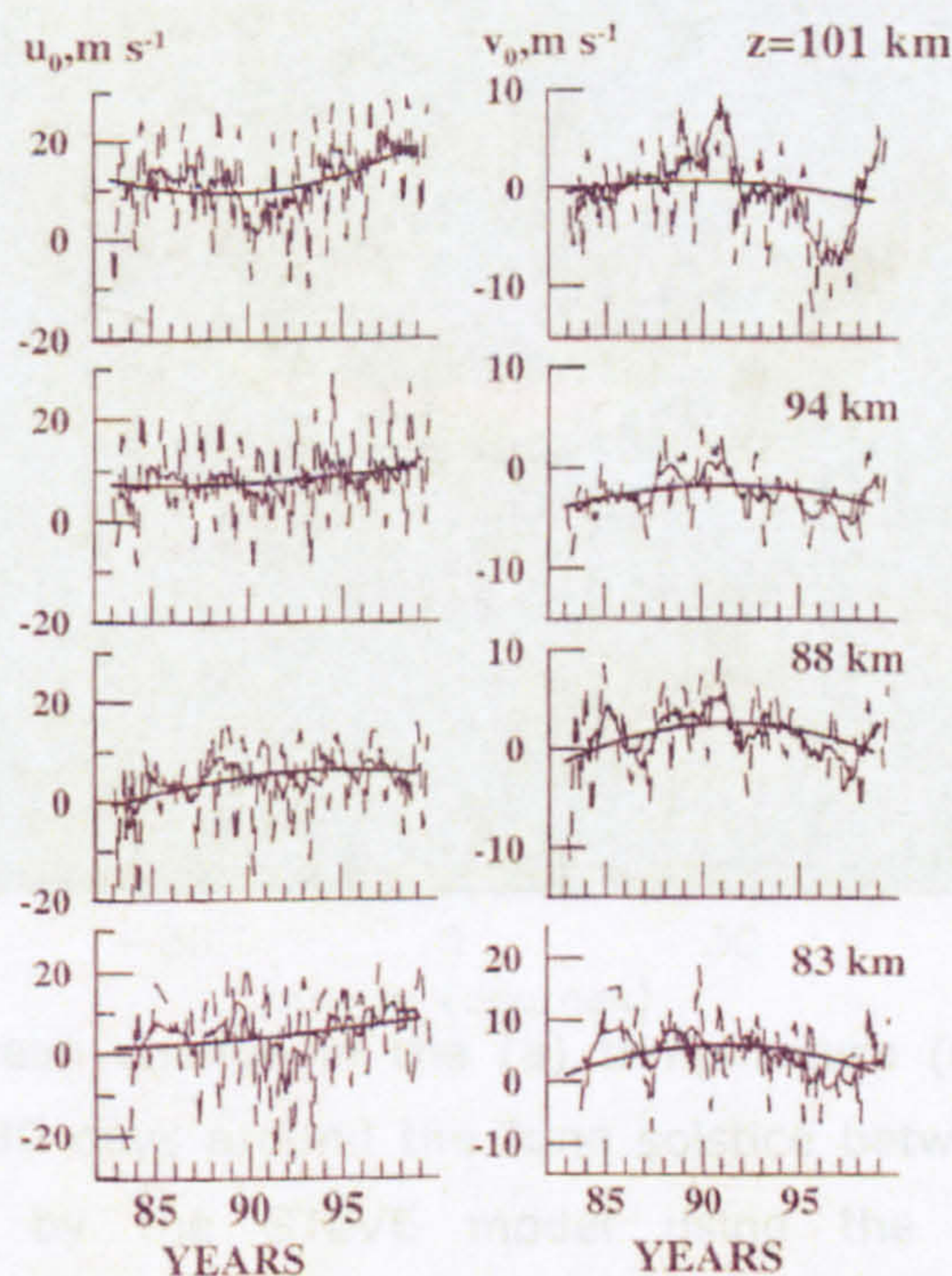


Figure IV.24, Interannual changes in the mean zonal and meridional winds measured at Collm, after Gavrilov *et al.* (2002). The dashed line shows the 3 month running average, the thin solid line shows the 12 month running average and the thick solid line shows the quadratic fit to the data.

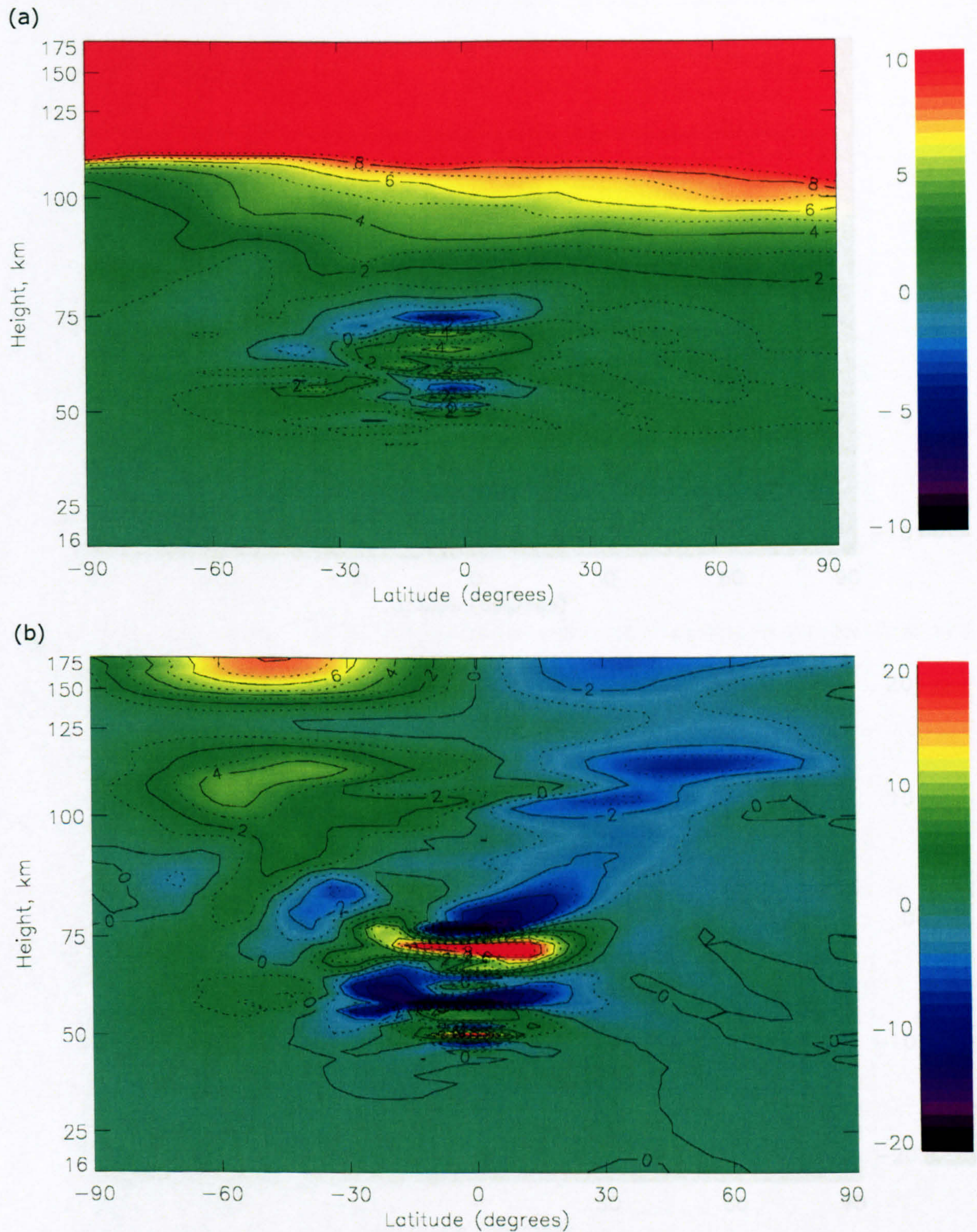


Figure IV.25, Zonal mean change in the (a) temperature (K) and (b) zonal winds (ms^{-1}) averaged over 30 days around the June solstice between solar minimum and maximum, calculated by the STEVE model using the MK2000 gravity wave parameterisation.

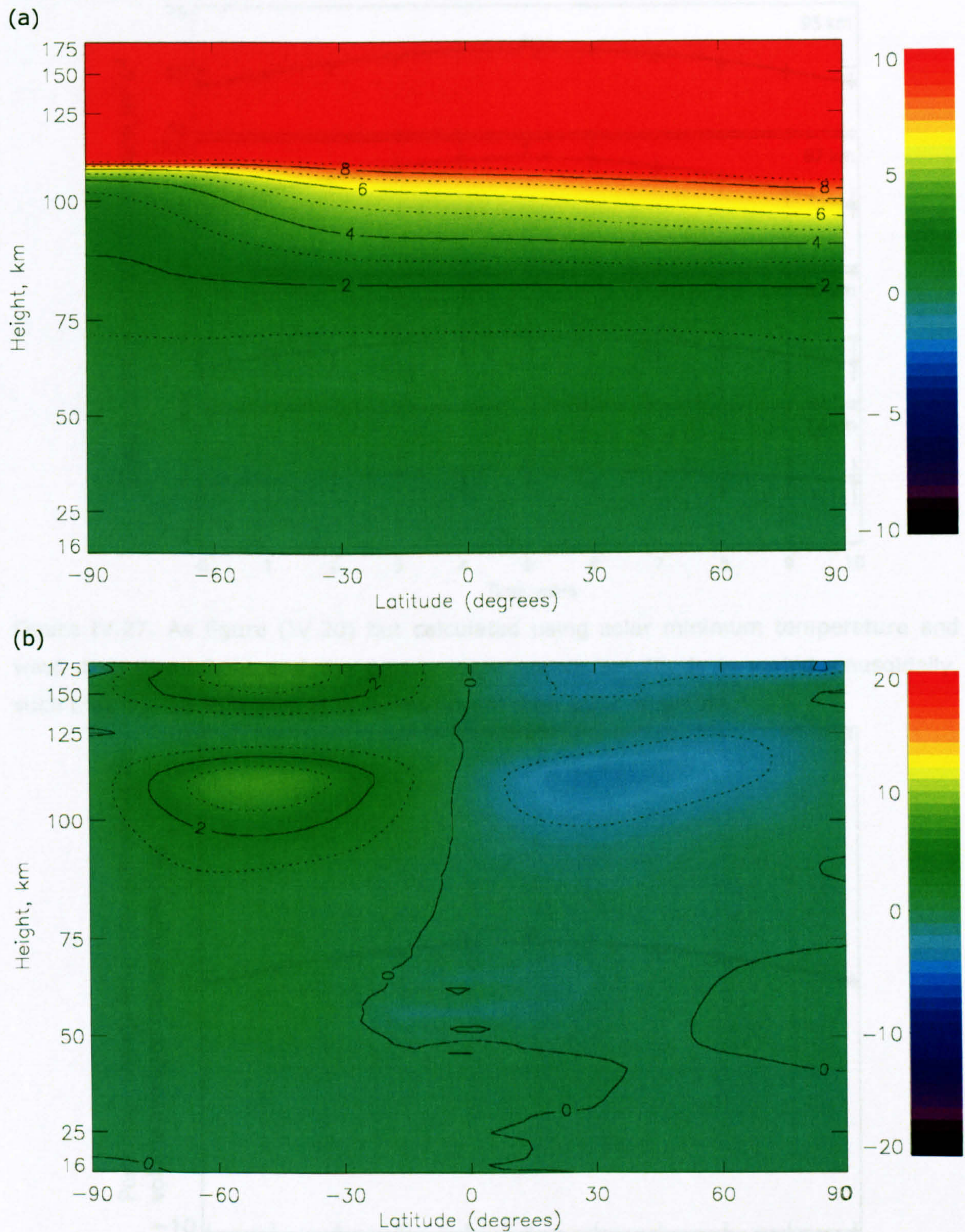


Figure IV.26, Zonal mean change in the (a) temperature (K) and (b) zonal winds (ms^{-1}) averaged over 30 days around the June solstice between solar minimum and maximum, calculated by the STEVE model using the Rayleigh friction gravity wave parameterisation.

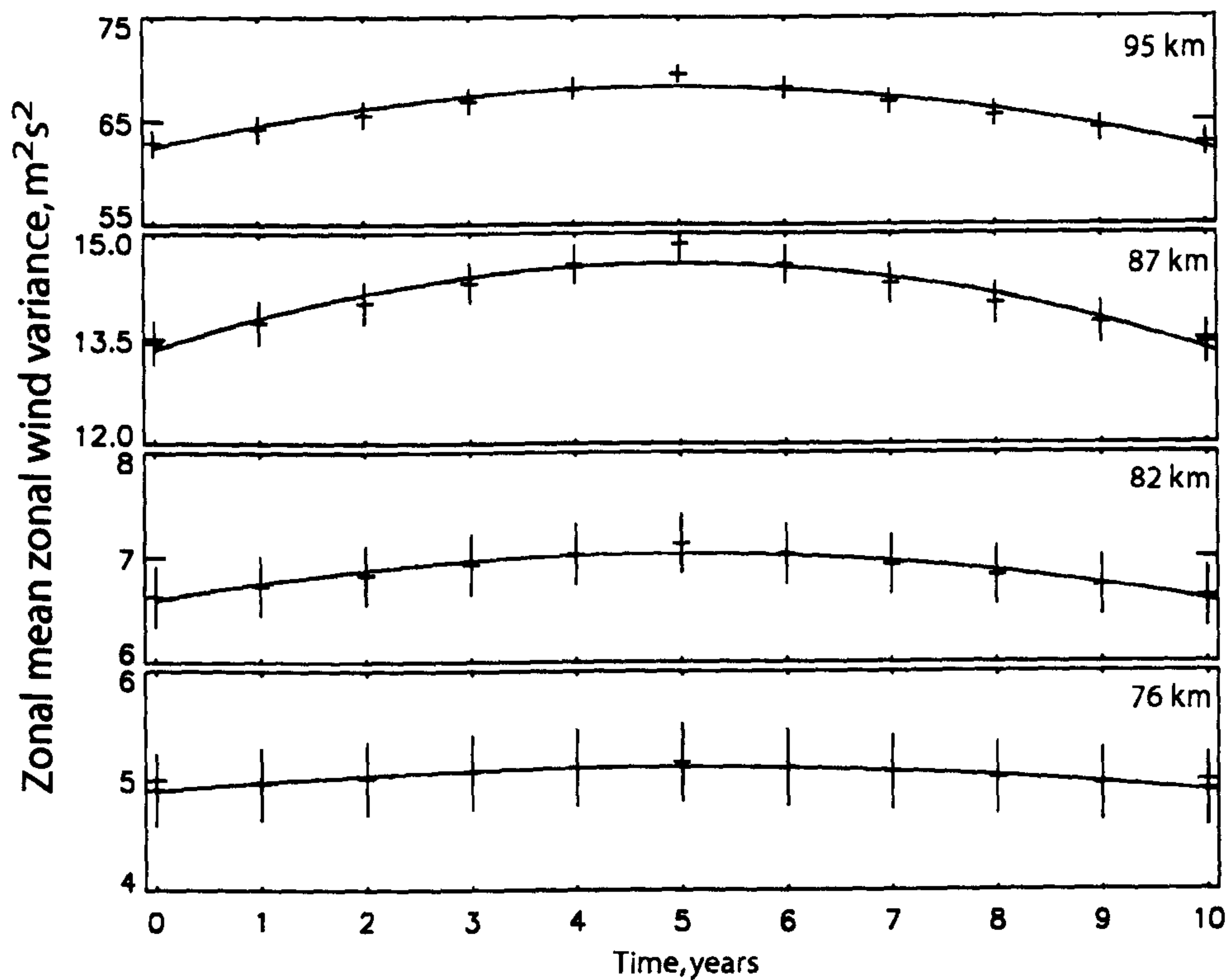


Figure IV.27, As figure (IV.20) but calculated using solar minimum temperature and wind data throughout and the gravity-wave source amplitude is varied sinusoidally, such that it is 10 % higher at solar maximum than solar minimum.

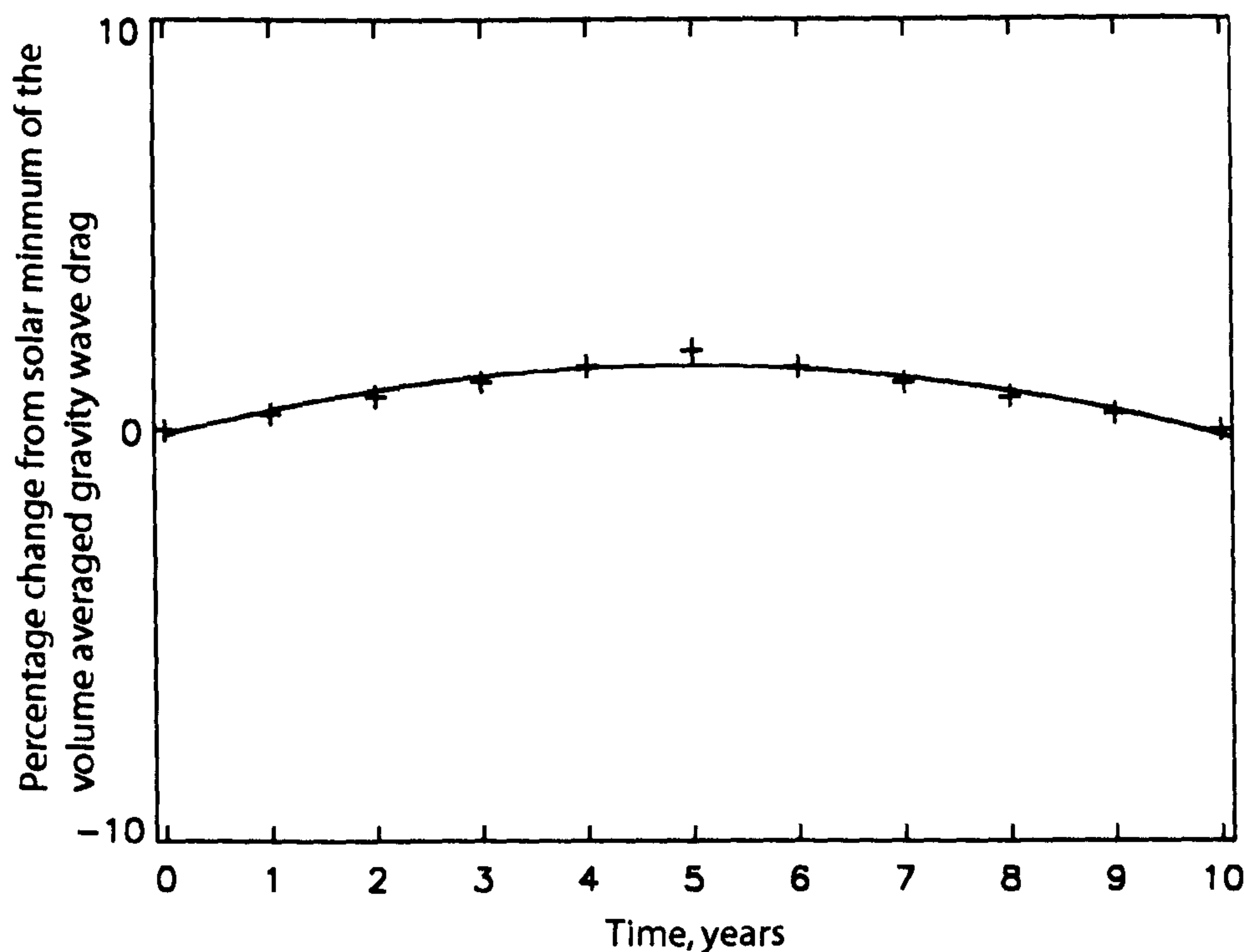


Figure IV.28, As figure 22 but calculated using solar minimum temperature and wind data throughout and the gravity-wave source amplitude is varied sinusoidally, such that it is 10 % higher at solar maximum than solar minimum.

Location	Height and / or Wave Period Range	Agreement with STEVE Model
Průhonice	95 km, 10 – 30 minutes	Summer only
	95 km, 31 – 60 minutes	Summer only
	95 km, 61 – 90 minutes	Summer only
Collm	83 km	Winter only
	88 km	Summer and winter
	94 km	Summer and winter
Shigaraki	65 – 80 km, less than 1 hour	None
	65 – 80 km, more than 1 hour	None
Saskatoon	70 km, less than 1 hour	Winter only
	79 km, less than 1 hour	Winter only
	88 km, less than 1 hour	None
	70 km, more than 1 hour	Winter only
	79 km, more than 1 hour	Winter only
	88 km, more than 1 hour	None

Table IV.1, Summary of the agreement and disagreement between the STEVE model results and the observations of zonal wind variances detailed in this chapter.

CHAPTER V, SUDDEN WARMINGS IN THE NORTHERN HEMISPHERE WINTER POLAR STRATOSPHERE: DEPENDENCE ON THE LOWER ATMOSPHERE GRAVITY WAVE SOURCE SPECTRAL AMPLITUDE

Chapter (IV) showed the effects of variable solar forcing on gravity-wave activity in the mesosphere in the STEVE model. However, no consideration was given to what impact changes in the mesospheric gravity waves might have on the rest of the atmosphere. Through downward control theory (Haynes *et al.*, 1991; see appendix I) it may be expected that this could have an impact on the stratosphere. As discussed in Chapter (III), the STEVE model does not simulate the stratosphere well, primarily due to the lack of planetary-wave activity in the model. The STEVE-3D model, introduced in Section (III.5), provides the means of studying the impact of variable gravity-wave activity on the stratosphere. The study presented in this Chapter will focus on the impact that gravity-wave drag in the mesosphere has on sudden stratospheric warmings (SSW). Section (V.1) will introduce SSW and their impact on the atmosphere. Section (V.2) will detail the results of previous studies which have considered the impact of gravity-wave drag on SSW. Section (V.3) will then detail the STEVE-3D model simulations, their results and insights gained into the effect of the mesosphere on the stratosphere.

V.1 SUDDEN STRATOSPHERIC WARMINGS

V.1.1 OBSERVED PROPERTIES OF SSW

SSW are major large-scale events, observed in the winter polar region of the middle atmosphere. They were discovered by Scherhag (1952) in radiosonde observations of temperatures over Berlin. They are seen as a rapid increase in localised temperatures in the polar stratosphere, which can rise by tens of degrees in a few days. Figure (V.1) shows a time series of temperatures derived from LIMS data during the winter of 1978 – 1979, which shows the rapid temperature rise produced by 3 such SSW. A major warming is defined as a reversal of the zonal mean zonal wind (from eastward to westward) at 60° North at 10 mb or below. From thermal wind balance, this is associated with a reversal in the meridional temperature gradient in this region and hence the localised increase in temperature. A minor warming occurs when a rapid rise in temperature is seen, but the zonal mean zonal wind does not reverse at 60° N at 10 mb. Figure (V.2) shows the effect of a SSW on the temperatures at 10 mb. As can be seen, there is a localised region of high temperature over Siberia, where the temperatures are ~30 K higher than at other regions at the same latitude.

Table (V.1) shows the occurrences of major warmings in the Northern Hemisphere between winter 1964 and spring 1986. As can be seen, major warmings occur somewhat sporadically, and do not occur every winter. Major warmings tend to occur during January and February and, with the exception of September 2002, all major warmings have been observed in the Northern Hemisphere (e.g. Hoppel *et al.*, 2003). By contrast, minor warmings occur more frequently and have been observed in both hemispheres across many years.

SSW can be split into two categories, wave 1 and wave 2, according to their observed characteristics. Figure (V.3) shows an example of a wave 1 warming, in which the winter polar vortex (cyclone) is shifted away from the pole by a warmer anticyclone (seen as a high in the geopotential height data). Figure (V.2) shows an example of a wave 2 warming, in which the polar vortex is distorted by two anticyclones. In extreme cases, these two anticyclones can split the polar vortex into two.

SSW affect not only the winter polar stratosphere, but much of the middle atmosphere. Using the model of Hsu (1980), Dunkerton *et al.* (1981) calculated the motion of air parcels during a model SSW. Figure (V.4) shows the mass stream function associated with the residual circulation during the modelled SSW. During the SSW, an enhanced descent of air parcels is seen near to the pole in the stratosphere and a corresponding ascent is seen in the equatorial stratosphere. This explains the cooling in the equatorial stratosphere which is often seen to accompany sudden warmings. In the mesosphere, a weak ascent is seen, associated with the observed cooling and corresponding changes in the winds and temperatures in this region (e.g. Lauter and Schminder, 1976). Further, recent studies such as Christiansen (2000, 2001) have indicated that zonal wind anomalies in the winter polar stratosphere, such as those produced by SSW, can propagate downwards into the troposphere and occasionally even reach the surface.

V.1.2 THE MATSUNO THEORY OF SSW

SSW are characterised by rapid increases in temperature in the winter polar stratosphere. However, as this region is in almost complete darkness, direct radiative heating cannot explain this temperature rise. Matsuno (1971; Matsuno hereafter) proposed a simple dynamical explanation of SSW which explains the observed features in the polar stratosphere, as described in the previous Section.

Quasi-stationary planetary waves are generated in the troposphere, typically by perturbations on the jet stream as it flows across mountain ranges. Those with low

zonal wavenumbers ($m = 1 - 3$) are able to propagate through the tropopause into the stratosphere. These waves can propagate vertically, provided the Charney-Drazin criterion is fulfilled (Charney and Drazin, 1961), which requires the background flow to be eastward and not too fast. As discussed in Section (I.5.2), the Eliassen-Palm flux \mathbf{F} gives the direction of propagation of the planetary wave in the meridional plane and the divergence of this flux gives a measure of the acceleration of the background flow which induces a residual circulation in the atmosphere (equations I.25 - I.30). Under normal winter conditions, \mathbf{F} can be found from observations. Figure (V.5a) shows the ray paths for planetary waves propagating into the stratosphere under normal winter conditions, along with the calculated acceleration due to their divergence. As can be seen, planetary waves propagating up from the troposphere at mid to high latitudes enter the stratosphere and then turn equatorwards. The associated acceleration drives the stratospheric Brewer-Dobson circulation (figure I.12).

Prior to a SSW, the background wind conditions must be such that planetary waves are anomalously refracted into the polar vortex, as in figure (V.5b). In this case, the planetary waves will grow in amplitude with height such that their wave-activity density (a measure of the wave amplitude) is given by,

$$A = \frac{1}{2} \rho_0 \overline{\eta'^2} \bar{q}_y, \quad \text{V.1}$$

where η' is the northward parcel displacement and \bar{q}_y is the northward mean potential vorticity gradient. As \bar{q}_y is positive then $\partial A / \partial t > 0$ when a planetary wave first enters the polar vortex. Andrews and McIntyre (1978b) showed that,

$$\nabla \cdot \mathbf{F} = -\frac{\partial A}{\partial t}, \quad \text{V.2}$$

(generalized Eliassen-Palm theorem), which means that a negative region of $\nabla \cdot \mathbf{F}$ is produced in this region. From equation (I.25), it can be seen that this represents a slowing of the zonal mean zonal wind in this region (figure V.5b). This weakens the polar vortex and, through the thermal wind equations, the temperature in this region rises.

This weakening of the polar vortex allows for more planetary waves to enter, such that $\partial A / \partial t > 0$ and further deceleration and heating occur (figure V.5c). Eventually the zonal mean zonal wind reverses (eastward to westward) and through the Charney-Drazin criterion, planetary waves are then prevented from propagating into this region (figure V.5d). However, their dissipation produces a rapid westward acceleration below this boundary.

Eventually, as planetary waves are excluded from this region, the wave-driven residual circulation decreases. Radiative cooling brings the stratosphere back towards its radiative equilibrium condition over a timescale of several days.

The Matsuno theory also explains the observed occurrences of major warmings. A SSW can only occur when planetary waves can propagate into the polar vortex. This requires the polar vortex to be sufficiently weak for the Charney-Drazin criterion to be satisfied. This necessary pre-conditioning of the polar stratosphere (e.g. Labitzke, 1981), along with the requirement for the planetary waves to have sufficient size to create the warming, provides a reason why major warmings do not occur regularly. Further, the strength of the polar vortex in the Southern Hemisphere and the associated lower amplitude of planetary waves in the Southern Hemisphere compared with the Northern Hemisphere (as there are fewer large mountain ranges in the southern hemisphere) explains why major warmings are not a common feature of the Southern Hemisphere winter.

Matsuno's theory of SSW is very simplistic and for example, does not account for wave-wave interactions (e.g. Butchart *et al.*, 1982; Hsu, 1981) or the influence of the equatorial QBO (e.g. Gray, 2003). However, the theory outlines the basic mechanism which is still used to explain the main features of SSW (e.g. Limpasuvan *et al.*, 2003). The discussion in the following Sections will use this theory as a means of explaining the model results in a concise manner.

V.2 PREVIOUS STUDIES INTO THE EFFECT OF GRAVITY-WAVE DRAG ON SSW

It has been known for many years that the inclusion of parameterised gravity-wave drag from orographic gravity-wave dissipation in the lower stratosphere can significantly improve numerical weather forecasting and climate models (Palmer *et al.*, 1986; McFarlane, 1987). This is because high resolution general circulation models (GCMs) are able to resolve some orographic gravity waves, but the momentum deposited by these does not balance the torque on the atmosphere created by friction with mountains, which leads to anomalous zonal mean features in these models. This torque can be balance by including the gravity-wave drag from unresolved waves.

The effects of non-orographic gravity waves and mesospheric gravity-wave drag on the stratosphere are less well understood. From a simple consideration of the mass of the mesosphere, it is not expected that mesospheric gravity-wave drag will have a large impact on the stratosphere at all times, so studies into its impact have concentrated on

the effects of mesospheric gravity-wave drag on SSW, which are highly sensitive to background conditions.

To date, relatively few studies have examined the effects of gravity-wave drag on SSW, a few key results of which will be summarised here. Pawson *et al.* (1998) investigated the effects of variable gravity-wave drag in the mesosphere on planetary waves in the stratosphere in a long, perpetual January model simulation. The authors used the Berlin troposphere-stratosphere-mesosphere general circulation model (TSM GCM), which is a spectral model with a horizontal resolution of $5.6^\circ \times 5.6^\circ$, which extends from the ground to ~ 85 km and which uses a Rayleigh friction scheme to simulate the effects of gravity waves in the mesosphere. They showed that when a Rayleigh scheme was used in the upper mesosphere only, varying the drag coefficient had a direct impact on the zonal winds in the mesosphere and upper stratosphere. However, when the height profile of the Rayleigh drag coefficient was extended down to act over the whole mesosphere, the zonal mean winds were seen to respond in a more complex way, which did not directly match changes in the eddy forcing or residual circulation. A decrease in the zonal mean winds in the stratospheric polar winter jet of up to 16 m s^{-1} were seen, which extended down into the troposphere where changes of up to 4 m s^{-1} were seen. The authors explained this as a fundamental change in the circulation related to changes in planetary-wave activity.

Pawson (1997) considered the effects of gravity-wave drag on SSW in a modified version of the TSM GCM. Using a combination of Rayleigh friction and the Palmer *et al.* (1986) orographic gravity wave parameterisation, Pawson investigated the effects of both stratospheric and mesospheric gravity-wave drag on SSW in relatively short model simulations (25 days), similar to those used for extended range weather forecasts. Varying the Rayleigh friction coefficient in the mesosphere (above the 5mb level) was seen to have an impact on the mesospheric jets, but to have no significant effect on the stratosphere or SSW. This was in contrast to the results of Pawson *et al.* (1998), which the author attributed to the relatively short simulation used, which was not long enough for changes in the mesosphere to propagate into the stratosphere (the term $\overline{w^*}$ in equation I.26 is of order 1 cm s^{-1}). However, the inclusion of stratospheric Rayleigh friction (below the 5 mb level) in addition to the Palmer *et al.* (1996) orographic scheme was seen to alter the timing and nature of a major warming and the timing of a minor warming simulated in the model. In the case of the major warming, the increased stratospheric gravity-wave drag was seen to decrease the planetary wave forcing in the winter polar stratosphere, which the author attributed to changes in the background wind.

Arnold and Robinson (1998, 2000, 2001) used modified versions of the UK Meteorological Office Stratosphere Mesosphere Model (see Chapter III for more details), to study the effects of solar variability on the middle atmosphere. They showed that, with the inclusion of the mesosphere and thermosphere in their model, the planetary-wave propagation / background atmosphere amplification mechanism proposed by Ramanathan *et al.* (1983) and Geller (1998) could explain how variable solar forcing could create changes in the Northern Hemisphere winter stratospheric temperature around the time of a SSW which were of similar order to those reported by Labitzke and Van Loon (1988). Arnold and Robinson (2003) went on to show that by including a 10 % decrease in the Rayleigh drag coefficient during solar maximum (based on the observations of Gavrilov (1995)), a larger change in temperature with solar activity could be produced in the model. They also showed that, even in the absence of changes in the solar UV flux, decreasing the gravity-wave drag in the model produced changes in the temperature of the winter polar stratosphere of order 1 K.

Both of the models discussed above have used Rayleigh friction to calculate the effects of non-orographic gravity waves. As this is not a ray-tracing gravity wave parameterisation, it does not take into account the propagation of gravity waves and as such does not produce a realistic gravity-wave drag. While McIntyre (1992) has argued that the detailed structure of the gravity-wave drag is not important to the long term residual circulation of the middle atmosphere, it has been shown that Rayleigh friction does not respond to changes in the background atmosphere in the same manner as more sophisticated gravity wave parameterisations (see Chapter IV). Other simulations such as Lawrence (1997) have performed similar simulations to those discussed above using ray-tracing parameterisations, but have only used a small number of gravity wave parameters (2 in the case of Lawrence, 1997). In the following Section, the STEVE-3D model will be used to investigate the impact of a variable gravity-wave source on SSW, using the MK2000 gravity wave parameterisation and a wide range of the lower atmosphere gravity-wave source strengths.

V.3 EFFECT OF VARIABLE GRAVITY WAVE SOURCES ON SSW IN THE STEVE-3D MODEL

In Chapters (III and IV) the suitability of the STEVE model to simulate the effects of gravity waves in the middle atmosphere has been demonstrated and the STEVE-3D model has been shown to simulate the dynamics of the Northern Hemisphere winter stratosphere reasonably well. In the following Section the simulation of a SSW in the STEVE-3D model will be demonstrated, before any perturbations due to variable gravity-wave sources will be considered. Section (V.3.2) will detail the theoretical

consideration of the impact of gravity-wave drag on SSW and Section (V.3.3) will present the STEVE-3D model results.

V.3.1 A SSW DURING JANUARY IN THE STEVE-3D MODEL

The STEVE-3D model was run for 90 days from the 1st December using the lower boundary geopotential height data described in Section (III.5). A SSW is seen to occur in the model during early to mid January. Figure (V.6) shows the temperatures in the stratosphere at the 10 mb pressure level every 2 days from the 10 – 20th January. A SSW of wave 1 type is clearly visible in the Northern Hemisphere temperature data. A temperature difference of ~80 K is seen between the warm and cold regions around the North Pole. This is comparable with that seen in the NCEP data (figures V.2 and V.3) which show differences of up to ~70 K at this pressure level. As expected, the temperature distribution around the South Pole (not shown here) is more zonally symmetric than in the north as planetary waves cannot propagate into the summer hemisphere. The SSW is seen to last for around 10 days, which is comparable to their observed duration.

In order to classify this as a major or minor warming, the zonal mean zonal wind must be found. Figure (V.7) shows the zonal mean zonal wind throughout the middle atmosphere in the STEVE-3D model, averaged over the 10 – 20th January. As can be seen, while the zonal mean winds are low in much of the Northern Hemisphere polar stratosphere region, no reversal in the zonal mean zonal wind occurs (specifically not at 60° North, 10 mb). This means that the SSW seen in the STEVE-3D model at this time is classed as a minor warming. Data from around this period will be used in Section (V.3.3) when the impact of a variable gravity-wave source on SSW in the STEVE-3D model is studied.

V.3.2 THEORY OF THE EFFECT OF GRAVITY-WAVE DRAG ON SSW

Pawson (1997) briefly describes a theory by which gravity-wave drag in the middle atmosphere can affect SSW. In this Section, this description will be expanded on, in order to provide a framework in which to interpret the STEVE-3D results in the following Section.

Figure (V.8) shows the sense of the residual circulation in the middle atmosphere during January (see figure I.12 for a more detailed depiction). SSW are clearly visible in changes in the temperature seen in the winter polar stratosphere. The Matsuno theory shows that they are sensitive to the preconditioned zonal winds in this region

(roughly indicated by the black dashed region). The zonal winds in this region can be affected by gravity-wave drag in the winter polar stratosphere. The results of Pawson (1997) indicate, within 25 days, changes in this drag can affect SSW.

The results of Pawson *et al.* (1998) show that, over longer time-scales, changes in the mesospheric gravity-wave drag can also affect SSW. By altering the zonal mean gravity-wave drag, \bar{X} term in equation (I.25), this can alter the residual circulation. Specifically, through equations (I.25 and I.27), this can alter the residual vertical wind, \bar{w}^* , in the mesosphere. As this vertical wind causes adiabatic heating in the winter polar atmosphere (responsible for creating the mesopause anomaly, discussed in Section I.3.4), variations in this can alter the temperature structure, and therefore the zonal wind in the winter polar stratosphere. Through this residual circulation, changes in the mesospheric gravity-wave drag can affect SSW. Pawson (1997) argued that this was only important in longer-term model simulations, as \bar{w}^* is slow and therefore changes in the mesosphere require longer than 25 days to propagate into the winter polar stratosphere.

In the following Section, the size of the impact of gravity-wave drag on a minor warming in January will be explored in the STEVE-3D model over a series of 90-day simulations. The effect of the contribution due to the mesospheric and stratospheric portions of the gravity-wave drag will also be investigated, using a wide range of tropospheric gravity-wave source amplitudes.

V.3.3 THE EFFECT OF GRAVITY-WAVE DRAG ON SSW IN THE STEVE-3D MODEL

V.3.3.1 STEVE-3D MODEL CONFIGURATION

In order to quantify the impact of gravity-wave drag in the middle atmosphere on SSW, five sets of simulations were performed using the STEVE-3D model. In all five sets, the spectral amplitude of tropospheric (and lower stratospheric) gravity-wave source (S_0 in equation III.4) was varied over a wide range (0 – 275 or 500 m^3s^{-2}). This range is perhaps larger than the observed variations in gravity waves averaged over three months, but is interesting to show the behaviour of the system in extreme cases. Varying the gravity-wave source spectrum, rather than directly altering the gravity-wave drag as in Pawson (1997) and Arnold and Robinson (2003), is also a more physically reasonable test to perform as the gravity waves sources in the troposphere and lower stratosphere are observed to vary with time, but the physics of gravity-wave dissipation does not.

The five sets of simulations will be referred to as experiments (I – V). In experiment (I), the effect of varying the gravity-wave source on the winter polar stratosphere, in the absence of a SSW (or, at least, with only a very weak warming event) is investigated. This is performed by using a version of the STEVE-3D model with only 10° zonal resolution. This version of the model is able to resolve some planetary-wave activity, but not enough planetary wave dissipation is present for a SSW to occur in the simulation. This experiment uses a total of 11 simulations with gravity-wave source spectral amplitude in the range $0 - 500 \text{ m}^3\text{s}^{-2}$.

Experiment (II) uses the ordinary $5^\circ \times 5^\circ$ resolution STEVE-3D model, as described in Section (III.5). Experiment (III) uses the same model, but under solar maximum conditions (all other experiments are performed using solar minimum conditions). Due to the nonlinearity of planetary waves, it is not possible to make a thorough investigation of the effects of variable solar conditions on SSW in the STEVE-3D model as in England *et al.* (2004) using only two different solar conditions, but it is useful to see the effects of gravity waves on SSW under two different solar conditions. These experiments use 16 simulations each with gravity-wave source spectral amplitude in the range $0 - 275 \text{ m}^3\text{s}^{-2}$ (computational expense prevented further simulations).

Experiments (IV) and (V) are used to quantify the effects of stratospheric and mesospheric gravity-wave drag on SSW. In experiment (V), only gravity-wave drag up to the 4.2 mb (~ 40 km, depending on latitude) pressure level is included in the model. In experiment (IV), only gravity-wave drag above the 4.2 mb pressure level is in the model, although the filtering due to winds below this level is still included in the calculation. This experiment complements the test of Pawson (1997) who investigated the effect of including only gravity-wave drag above the 5 mb pressure level in the TSM GCM. As noted by Pawson (1997), experiments such as these are not strictly self-consistent as it is possible that a net zonal or meridional acceleration is applied to the model, even though an isotropic gravity-wave source is used, as not all the momentum deposition is included. However, this is also true of simulations using the Rayleigh friction approximation to gravity-wave drag, which has been used in the previous studies outlined in Section (V.2) without creating notable problems. It therefore seems reasonable to compare the results of these experiments with experiments (I – III) in order to separate the effects of stratospheric and mesospheric gravity-wave drag on SSW. These experiments use 11 simulations each with gravity-wave source spectral amplitude in the range $0 - 500 \text{ m}^3\text{s}^{-2}$.

V.3.3.2 STEVE-3D MODEL RESULTS AND DISCUSSION

As noted by Medvedev and Klaassen (1998), the relationship between the gravity-wave drag calculated by the MK2000 parameterisation and the gravity-wave source amplitude used is complex. This is because the drag is a function of the background flow and the background flow is modified by the gravity-wave drag (see appendix II). In order to see the change in gravity-wave drag in the mesosphere with changes in the gravity-wave source, the zonally averaged gravity-wave drag averaged over 40° – 90° North, between 2.4×10^{-1} – 2.4×10^{-3} mb (~ 55 – 88 km) and between 10^{th} – 20^{th} January is found (this region is roughly indicated by the green enclosed region in figure V.8). Figure (V.9) shows the change in this drag, compared with zero gravity-wave source (zero drag) for experiments (I-V).

In experiment (V), there is no mesospheric gravity-wave drag, so this value remains zero throughout. In experiments (I – IV), there is a general increase in the drag as the source amplitude increases. This increase is rapid at low values of the source and appears to saturate at higher values of the source. This is to be expected as, considering the theory of Lindzen (1981), as the source amplitude of the gravity waves increase, they become convectively unstable at progressive lower altitudes. This means that, at increasing source amplitudes, more gravity waves dissipated before reaching the mesosphere. However, the general increase seen in the mesospheric drag with increasing gravity-wave source is sufficient to see the effect of variable mesospheric gravity-wave drag on SSW.

Following experiment (I), the effect of this variable gravity-wave drag on the residual circulation in the mesosphere (the region indicated by the purple box in figure V.8) can be seen in figure (V.10). As may be expected from considering the transformed Eulerian mean equations of motion, as the gravity-wave drag in this region increases, so too does the meridional circulation. This picture is complicated by changes in the background atmosphere and planetary wave dissipation, which means that a perfect correlation is not seen.

Figure (V.11) shows the change in the EP flux divergence in the winter polar stratosphere (indicated by the blue boxed region in figure V.8). In experiment (I), planetary-wave propagation in the winter polar stratosphere is not altered significantly by the varying gravity-wave source. This is because, as a clear SSW is not seen in this experiment, there is very little planetary-wave activity in this region. This means that, as might be expected, the variable gravity-wave source has very little impact on the temperature of the winter polar stratosphere in this experiment (see figure V.12). This shows that the changes in the meridional circulation seen in figure (V.10) are not able

to alter the temperature of the winter polar stratosphere by more than 0.5 K under normal conditions (in the absence of a SSW). In the mesosphere, there is a general increase in the temperature as the gravity-wave source increases (figure V.13), which is associated with the increased meridional circulation leading to an increase in the diabatic heating in this region.

In experiments (II) and (III), a clear SSW was visible during January 10th – 20th (see, for example, figure V.6). In both of these experiments, as the gravity-wave source spectral amplitude is increased from 0 – 225 m³s⁻², a general decrease in the vertical component of the mesospheric circulation in the winter hemisphere is seen. Clearly, as the acceleration due to gravity-wave dissipation increases over this range, a decrease in the meridional circulation must be related to a change in other parameters. The most likely cause of this seems to be a change in the EP flux divergence term, which is the other main driver of the meridional circulation. What is produced is a pattern of residual circulation similar to that shown in figure (V.4). A complex pattern of change in the EP flux divergence in the winter polar stratosphere can be seen in figure (V.11). Following a rise up to 50 m³s⁻², the EP flux divergence decreases to a minimum around 100 m³s⁻² in both cases. This corresponds to an increase in the deceleration of the winter polar vortex and a stronger SSW with temperature increases reaching 3 K, shown in figure (V.12).

This decrease in the EP flux is caused by a change in the zonal mean zonal wind in the middle atmosphere. Figure (V.14) shows the change in the zonal mean zonal wind between the zero gravity-wave source simulation and the 100 m³s⁻² mean spectral amplitude simulation, averaged over the 10th – 20th January for experiment (II). Comparing this to the zonal mean zonal winds during this period (figure V.7), it can be seen that increasing the gravity-wave drag has led to a decrease in the wind in the winter polar stratosphere and mesosphere, which will allow planetary waves to propagate into the polar vortex more easily. This leads to the increase in EP flux divergence and the strengthening of the SSW seen in the model. The increase in the zonal winds in the low latitude winter hemisphere is associated with the decrease in planetary waves propagating into this region associated with a change in their mean ray paths (see figure V.5a-c). In the upper mesosphere and lower thermosphere, the strong decrease in the zonal mean zonal winds shows the *in situ* effect of the increase in the gravity-wave dissipation in this region.

Above 100 m³s⁻², the EP flux divergence increases up to a peak at a source strength of 200 m³s⁻². This peak in EP flux divergence corresponds to a decrease in the strength of the SSW and a decrease in the temperatures in the winter polar stratosphere. This

is because the zonal mean zonal winds in the stratosphere have changed to a less favourable state and fewer planetary waves can propagate into the winter polar vortex. Above $225 \text{ m}^3\text{s}^{-2}$ the residual circulation appears to become stronger, the EP flux divergence decreases and the temperatures begin to increase again. This is the effect of the increased gravity wave driven meridional circulation in the mesosphere.

The change in temperatures in the mesosphere associated with the change in the strength of the sudden warming can be seen in figure (V.13). As expected, in experiments (II) and (III), these are approximately anti-correlated with the changes in temperature in the stratosphere. However, it appears that the temperature changes seen in the mesosphere may be made up of this anti-correlated signature and a small net increase with the gravity-wave source, related to the increase in the gravity wave driven circulation in this region. Further simulations at higher spectral amplitudes would be needed to confirm this.

Experiments (IV) and (V) can now be used to investigate which of the effects seen in experiments (II) and (III) may be related to either the stratospheric or mesospheric portion of the gravity-wave drag on its own. In experiment (IV), a clear SSW is not seen. This is because, in the absence of gravity-wave drag in the stratosphere, the winter polar vortex remains too strong for enough planetary waves to enter. This is seen in the meridional circulation in the mesosphere, which shows no sign of the decrease seen in experiments (II) and (III). The EP flux divergence in the stratosphere show an increase with the increasing mesospheric gravity-wave drag. This is similar to the increase seen in experiments (II) and (III) at high values of gravity-wave drag, but no increase in the size of the SSW is seen in experiment (IV). In the absence of a clear SSW, the increase in the residual circulation leads to a general increase in the temperature in the winter polar stratosphere, in agreement with the results of experiment (I). The difference in the temperature change seen in experiments (I) and (IV) may be an effect of the difference in resolution of the two models used, which affects planetary-wave activity throughout the middle atmosphere.

In experiment (V), a clear SSW is not seen as, in the absence of the gravity wave driven meridional circulation, the polar vortex in early January is too strong for enough planetary waves to enter. The residual circulation in the mesosphere in this experiment behaves in a similar manner to that seen in experiments (II) and (III), although even larger changes are seen in this case. As there is no mesospheric gravity-wave drag in this simulation, figure (V.10) shows the effect of varying planetary-wave activity on the meridional circulation in the mesosphere (figure V.4). This shows that the decrease seen in the vertical component of this circulation in

experiments (II) and (III) for gravity-wave source spectral amplitudes up to $200 \text{ m}^3\text{s}^{-2}$ is related to changes in planetary-wave activity. At values of the gravity-wave source spectral amplitude above $200 \text{ m}^3\text{s}^{-2}$, the vertical wind increases in experiment (V), which is related to the exclusion of more planetary waves from the winter polar region which was suggested in experiments (II) and (III). The change in EP flux divergence with gravity-wave source strength does not appear to show a clear trend in experiment (V), although the results from the change in the meridional circulation suggest that planetary-wave activity is changing in a manner similar to that in experiments (II) and (III).

In experiments (I), (II), (III) and (V), a pattern in the change in temperature with gravity-wave source strength is visible. In all four experiments an increase in temperature, with gravity-wave source spectral amplitude is seen up to around $150 \text{ m}^3\text{s}^{-2}$. Around $200 \text{ m}^3\text{s}^{-2}$, all four show a local minima in temperature, and above this value the temperature begin to increase again, albeit only by a small amount. However, the dramatic increase in temperature up to a gravity-wave source spectral amplitude of $100 \text{ m}^3\text{s}^{-2}$ is only seen when a strong SSW is present in the simulations. While varying the stratospheric gravity-wave drag can alter the temperature in the stratosphere, only when the polar vortex has the correct preconditioning (through mesospheric and stratospheric gravity-wave drag) and the planetary waves are of sufficient amplitude can a SSW occur. Further, when a clear SSW is seen in the model, gravity-wave drag can have a large impact on the strength of this SSW.

Throughout this discussion, only data from 10th – 20th January has been considered. Figure (V.15) shows the change in temperature with gravity-wave source spectral amplitude for the five experiments, averaged over the whole of January. While the absolute values are different as the data is averaged over a longer period of time, this figure shows the same overall pattern as figure (V.12), showing that the 10 day interval considered in this analysis was sufficient to capture the behaviour of the system. It also suggests that the variable gravity-wave drag is altering the size of the SSW, rather than significantly altering its timing.

V.4 SUMMARY

The strength of the gravity-wave source used in the STEVE-3D model simulations has been shown to have an impact on SSW. Variable gravity-wave drag can alter the background zonal wind conditions in the stratosphere, making them more or less favourable for a SSW to occur. Monthly averaged temperatures in the winter polar

stratosphere in the model can be altered by up to 2 K through altering the strength of the simulated SSW.

The experiments presented in the previous Section have only considered a minor warming event during January. The values of the gravity-wave source spectral amplitude which produce favourable or less favourable background conditions for the SSW will no doubt vary from warming to warming, but it seems likely that the overall pattern should be similar to that presented here. This highlights a need for further study and observational constraint of gravity-wave sources around the time of SSW, which may vary greatly from event to event.

This Chapter has highlighted an important aspect of the interaction between gravity waves and other waves in the middle atmosphere, over a timescale of several days. The STEVE-3D model was suitable for this type of study, but this model does not take into account the coupling between the dynamics and composition of the atmosphere. In the following Chapter the coupling between gravity waves and the atmospheric diurnal tide, on timescales of up to one day will be simulated in the CMAT model. The effects this has on the dynamics and composition of the MLT region will be addressed.

Year	Number	Month
1964-1965	0	
1965-1966	1	F
1966-1967	0	
1967-1968	1	J
1968-1969	0	
1969-1970	1	J
1970-1971	1	J
1971-1972	0	
1972-1973	1	J-F
1973-1974	0	
1974-1975	0	
1975-1976	0	
1976-1977	1	J
1977-1978	0	
1978-1979	1	F
1979-1980	0	
1980-1981	1	F
1981-1982	0	
1982-1983	0	
1983-1984	0	
1984-1985	1	J
1985-1986	0	

Table V.1, Occurrences of major stratospheric warmings in the Northern Hemisphere winter between 1964 – 1965 and 1985 – 1986, after Andrews *et al.* (1987).

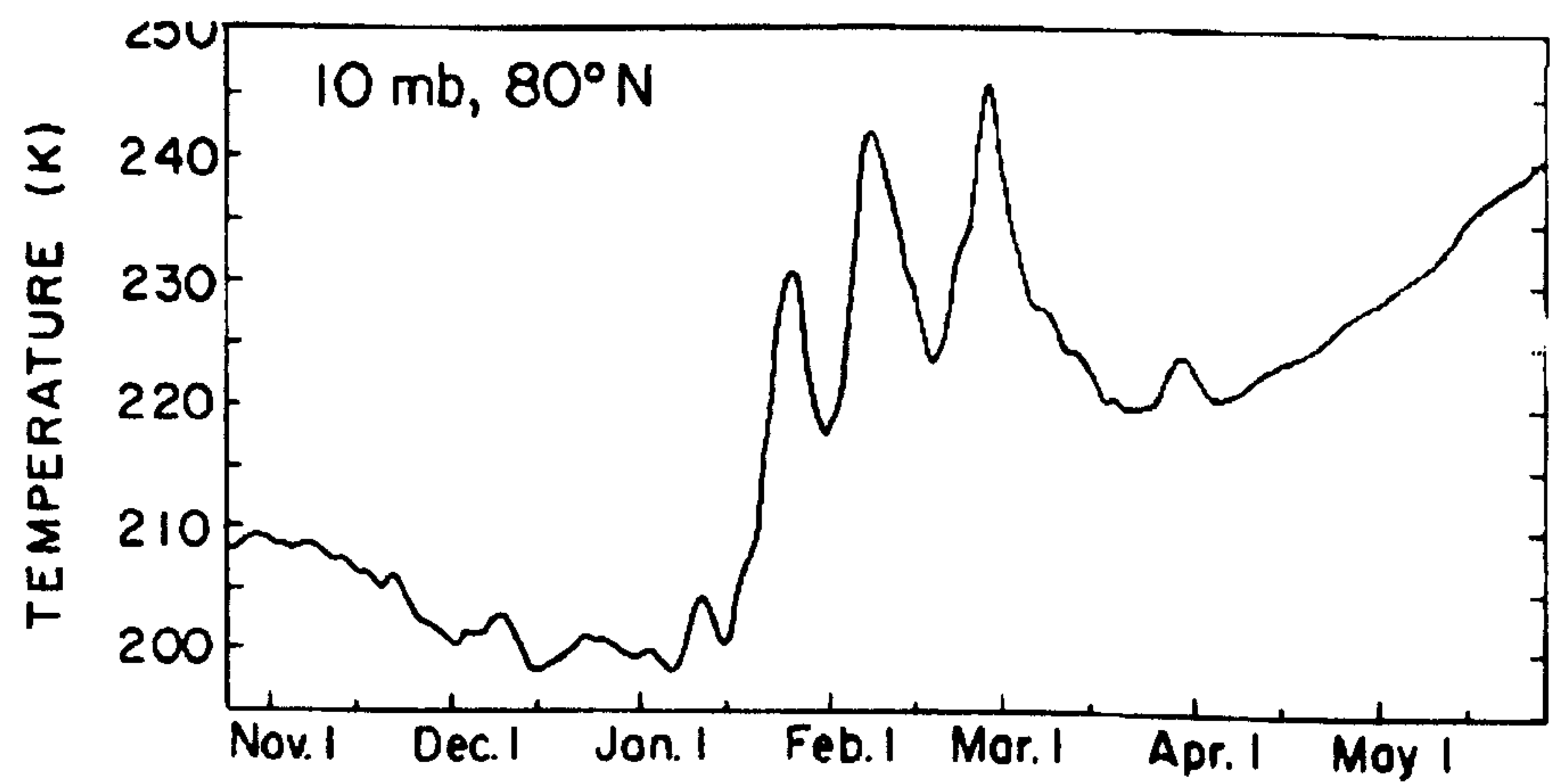


Figure V.1, Variation of the zonal mean temperature derived from LIMS data at 10 mb, 80° N from October 1978 to May 1979. After Gille and Lyjak (1984).

lon: plotted from 0.00 to 360

lat: plotted from 0. to 90

lev: 10.00

t: Jan 17 1992

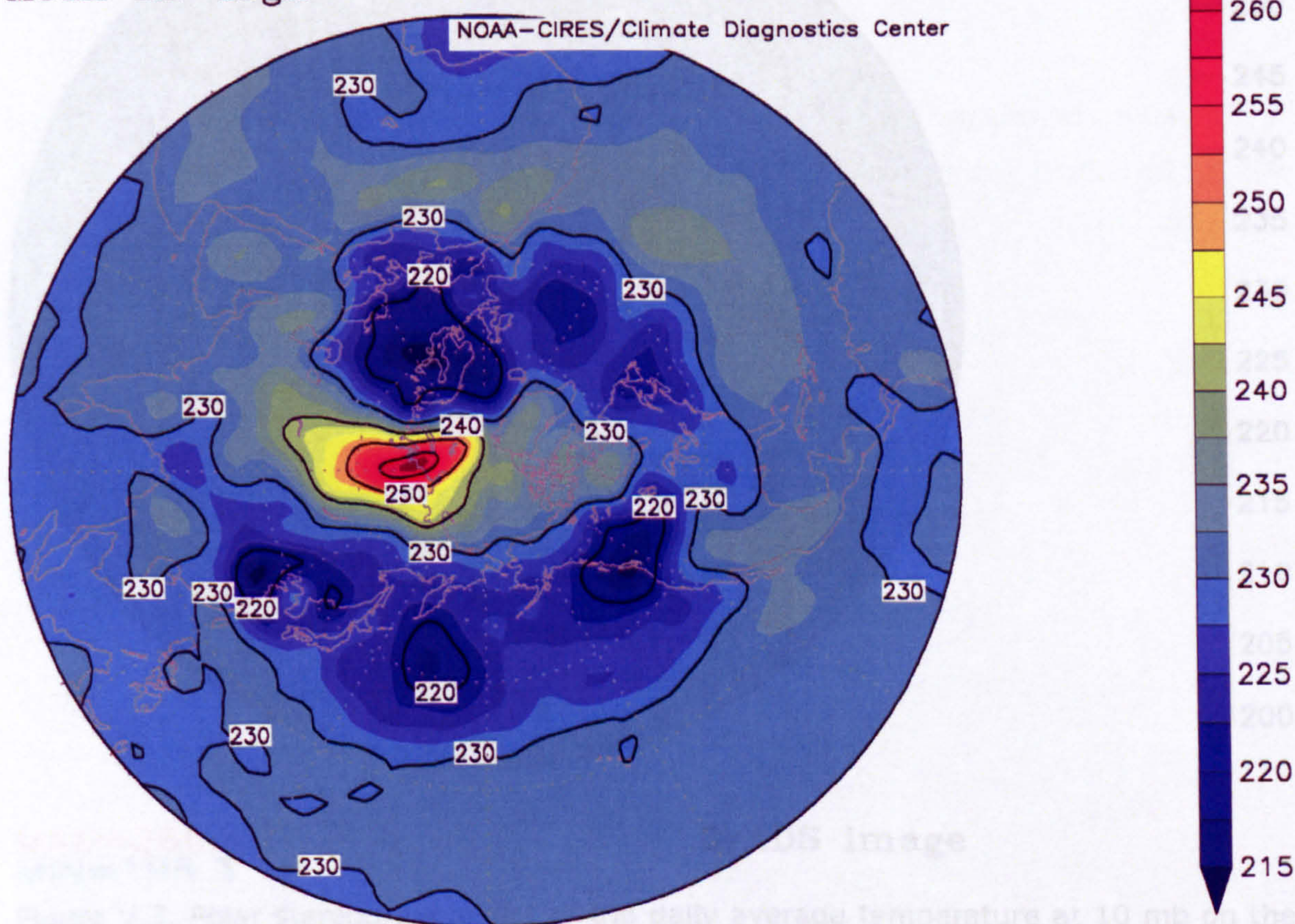
lon: plotted from 0.00 to 360

lat: plotted from 0. to 90 NOAA-CIRES/Climate Diagnostics Center

lev: 10.00

t: Feb 26 1979

Mean air degK

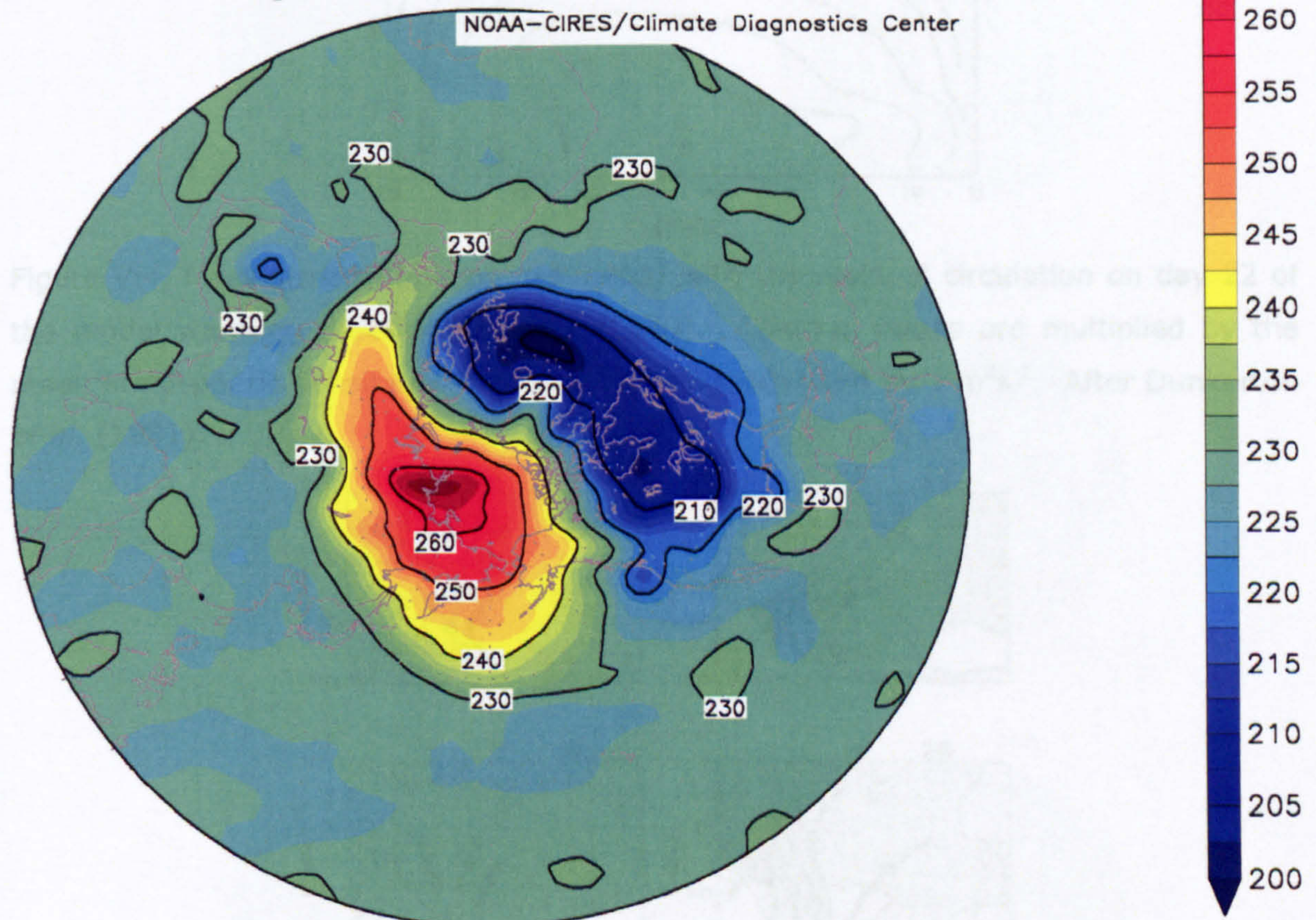


MAX=263.48
MIN=213.48

GrADS image

Figure V.2, Polar stereographic plot of the daily average temperature at 10 mb on the 26th February 1979. Data is taken from the NCEP reanalysis database. Image courtesy of the NOAA-CIRES Climate Diagnostics Centre.

lon: plotted from 0.00 to 360
 lat: plotted from 0. to 90
 lev: 10.00
 t: Jan 12 1992
 Mean air degK



MAX=266.4
 MIN=198.3

GrADS image

Figure V.3, Polar stereographic plot of the daily average temperature at 10 mb on the 12th February 1992. Data is taken from the NCEP reanalysis database. Image courtesy of the NOAA-CIRES Climate Diagnostics Centre.

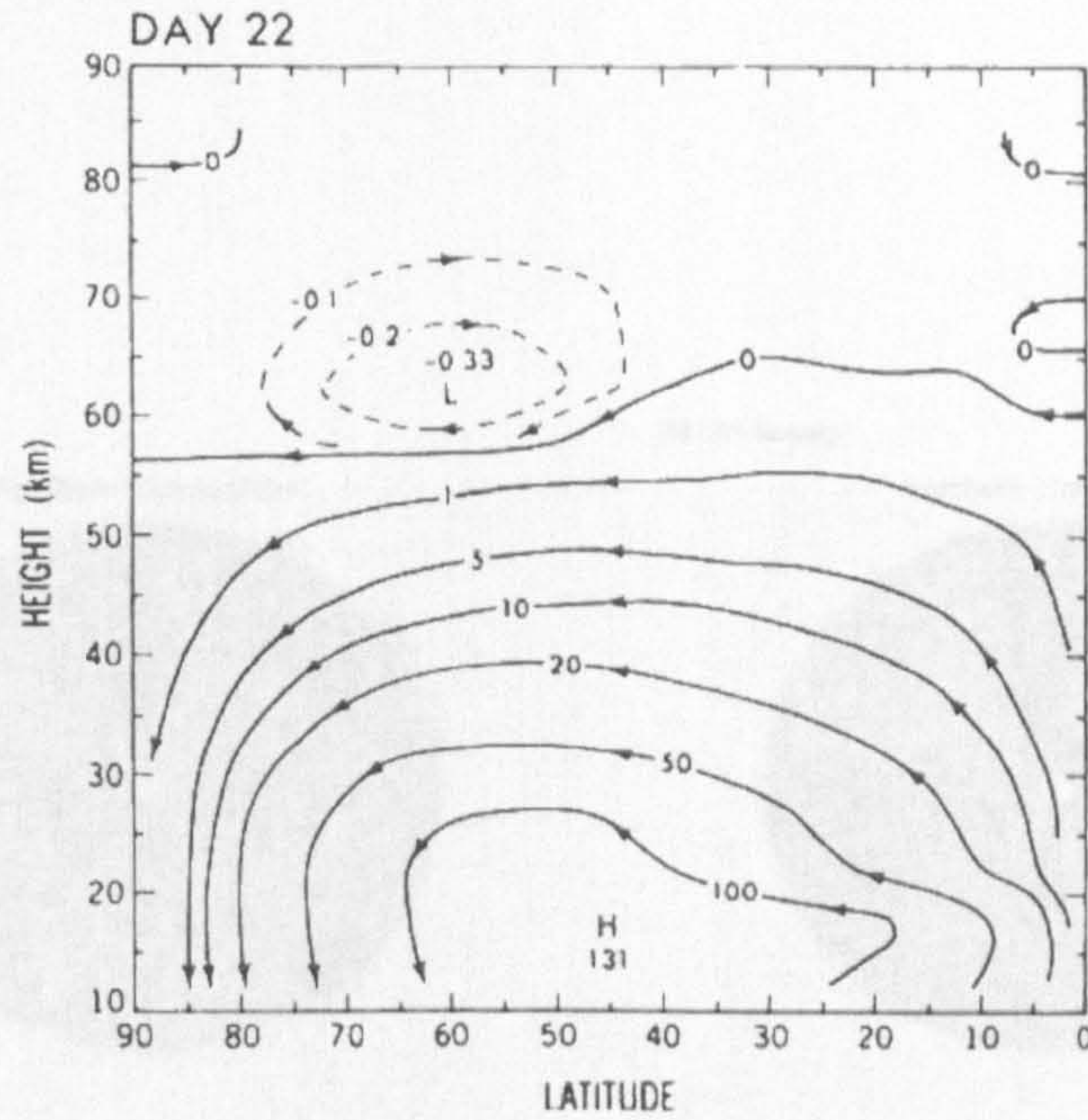


Figure V.4, Mass stream function associated with the residual circulation on day 22 of the model warming described by Hsu (1980). Contour values are multiplied by the mean atmospheric density at the bottom of the model and by $1 \text{ m}^2 \text{ s}^{-1}$. After Dunkerton *et al.* (1981).

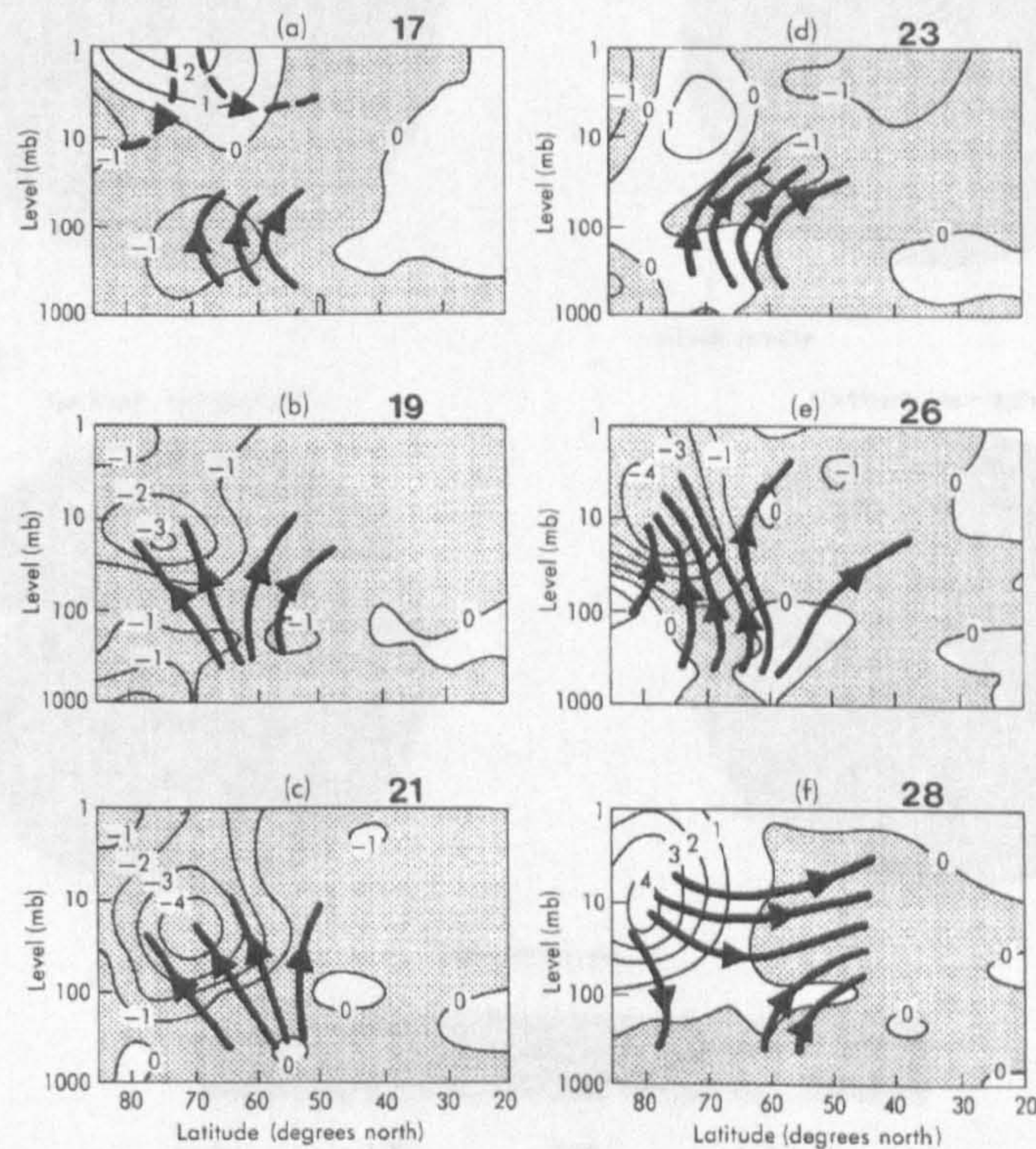


Figure V.5, Contours of acceleration due to EP flux convergence (units are 10^{-4} ms^{-2} , negative values are shaded) and calculated ray paths for (a) 17th February, (b) 19th February, (c) 21st February, (d) 23rd February, (e) 26th February and (f) 28th February 1979, after Palmer (1981). The solid ray paths denote the wavenumber 2 contribution to \mathbf{F} and the dashed line denotes the wavenumber 1 contribution.

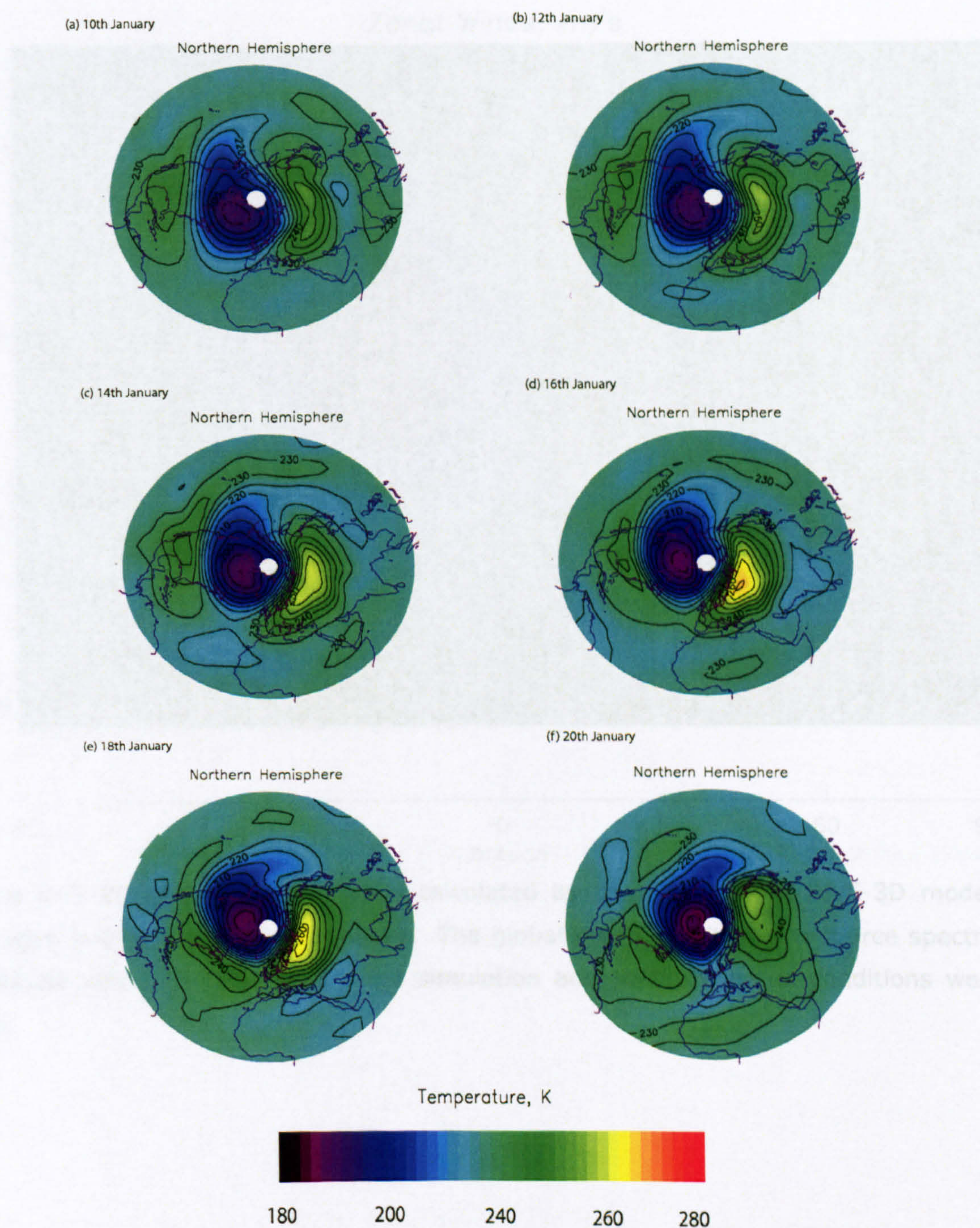


Figure V.6, Polar stereographic plot of the daily averaged temperatures at 2.4 mb in the $5^\circ \times 5^\circ$ STEVE-3D model for (a) 10th January (b) 12th January (c) 14th January (d) 16th January (e) 18th January (f) 20th January. The global mean gravity-wave source spectral amplitude was $100 \text{ m}^3\text{s}^{-2}$ during the simulation and solar minimum conditions were used.

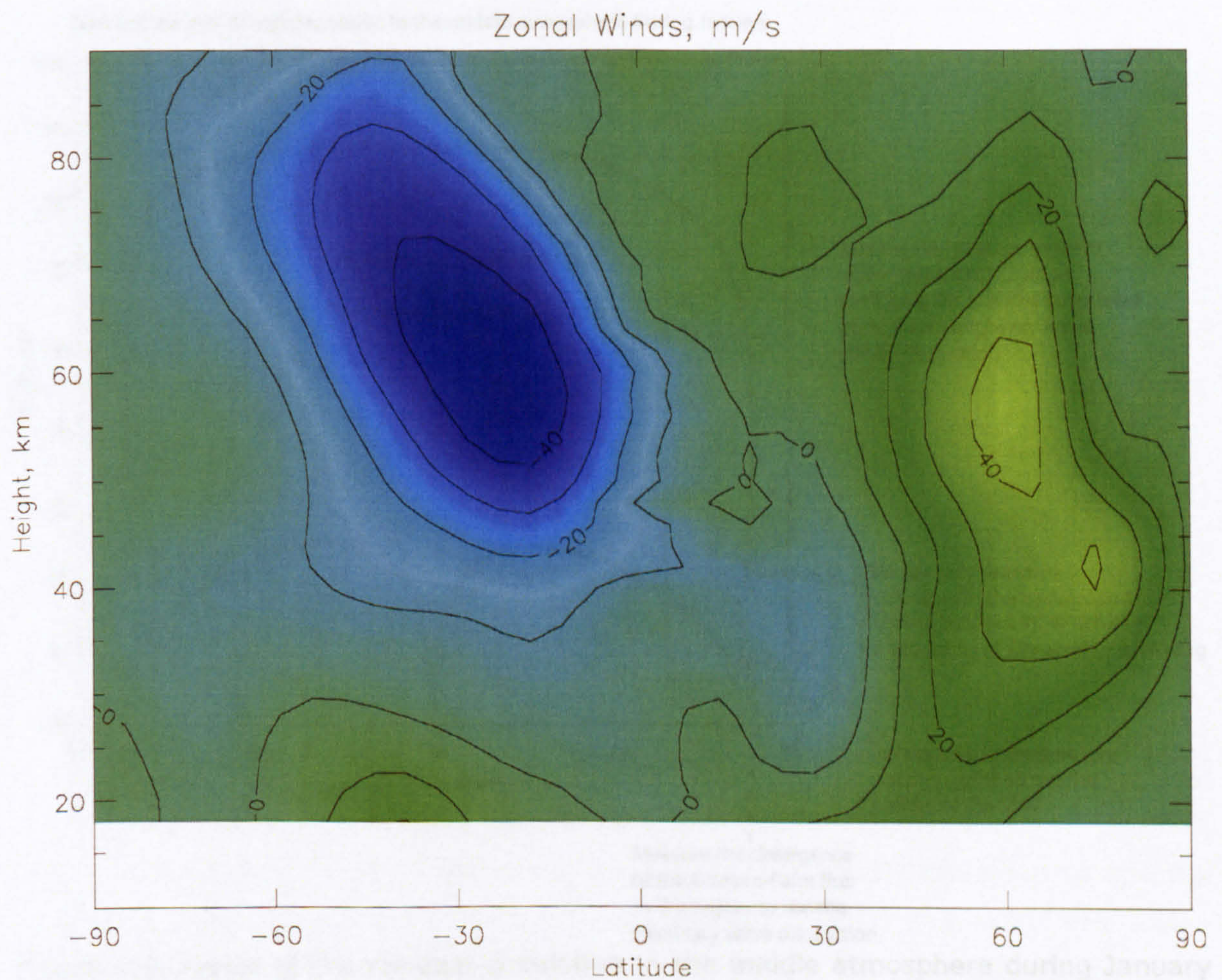


Figure V.7, Zonal mean zonal winds calculated by the $5^\circ \times 5^\circ$ STEVE-3D model, averaged over the 10th - 20th January. The global mean gravity-wave source spectral amplitude was $100 \text{ m}^3\text{s}^{-2}$ during the simulation and solar minimum conditions were used.

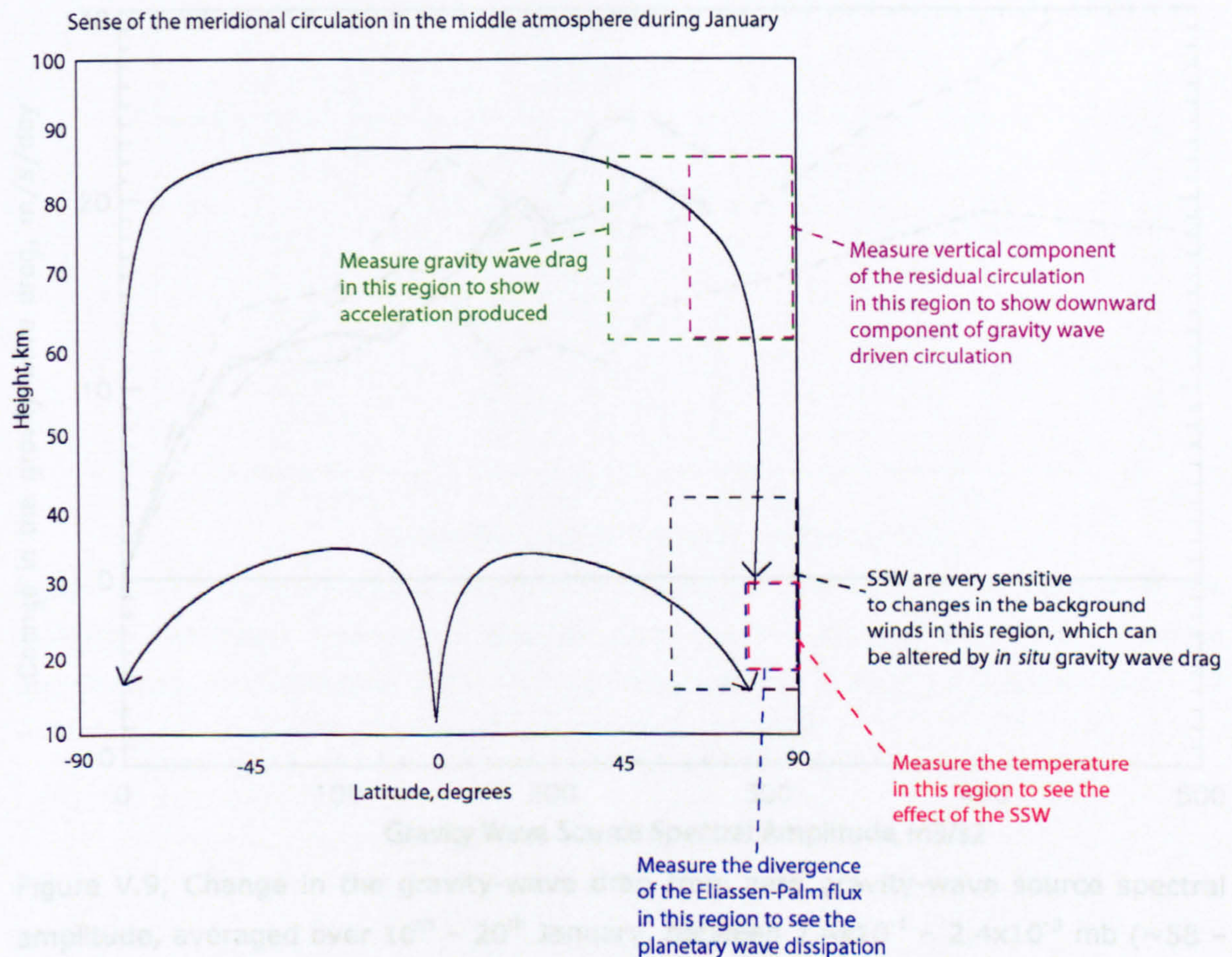


Figure V.8, Sense of the residual circulation in the middle atmosphere during January and the effects of gravity-wave drag on stratospheric sudden warmings. The boxed regions show the approximate areas over which STEVE-3D model parameters are averaged for analysis in Section (V.3.3).

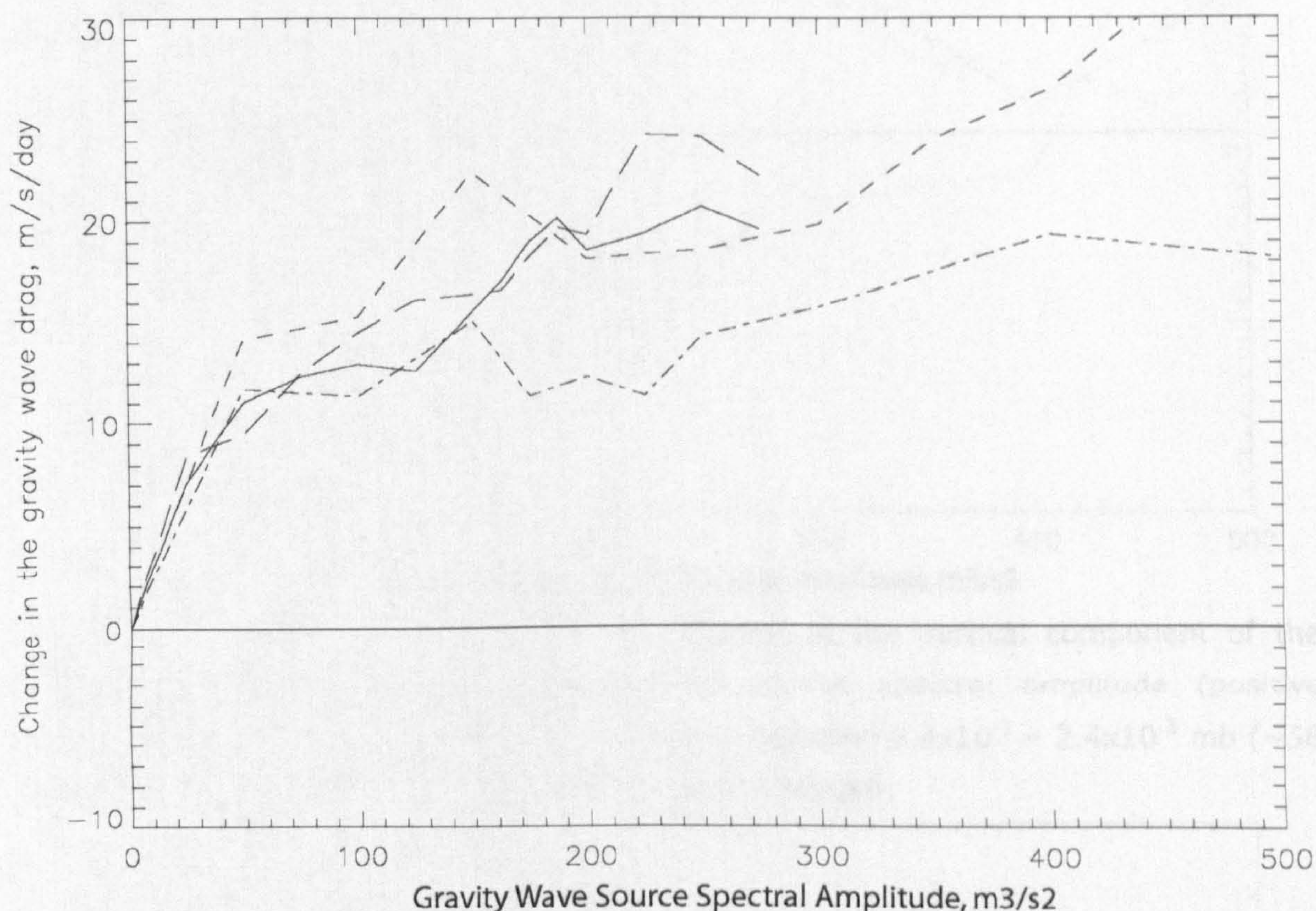


Figure V.9, Change in the gravity-wave drag from zero gravity-wave source spectral amplitude, averaged over 10th – 20th January, between 2.4×10^{-1} – 2.4×10^{-3} mb (~ 58 – 88 km), between 40° – 90° North and zonally averaged. The solid line depicts the 5° x 5° STEVE-3D model results for solar minimum, the long dashed line depicts the 5° x 5° STEVE-3D model results for solar maximum, the dot-dashed line depicts the 10° x 5° STEVE-3D model results for solar minimum, the short dashed line depicts the results from the 5° x 5° STEVE-3D model at solar minimum with gravity-wave drag applied above 4.2 mb only and the dotted line (zero for all values in this case) depicts the results from the 5° x 5° STEVE-3D model at solar minimum with gravity-wave drag applied below 4.2 mb only.

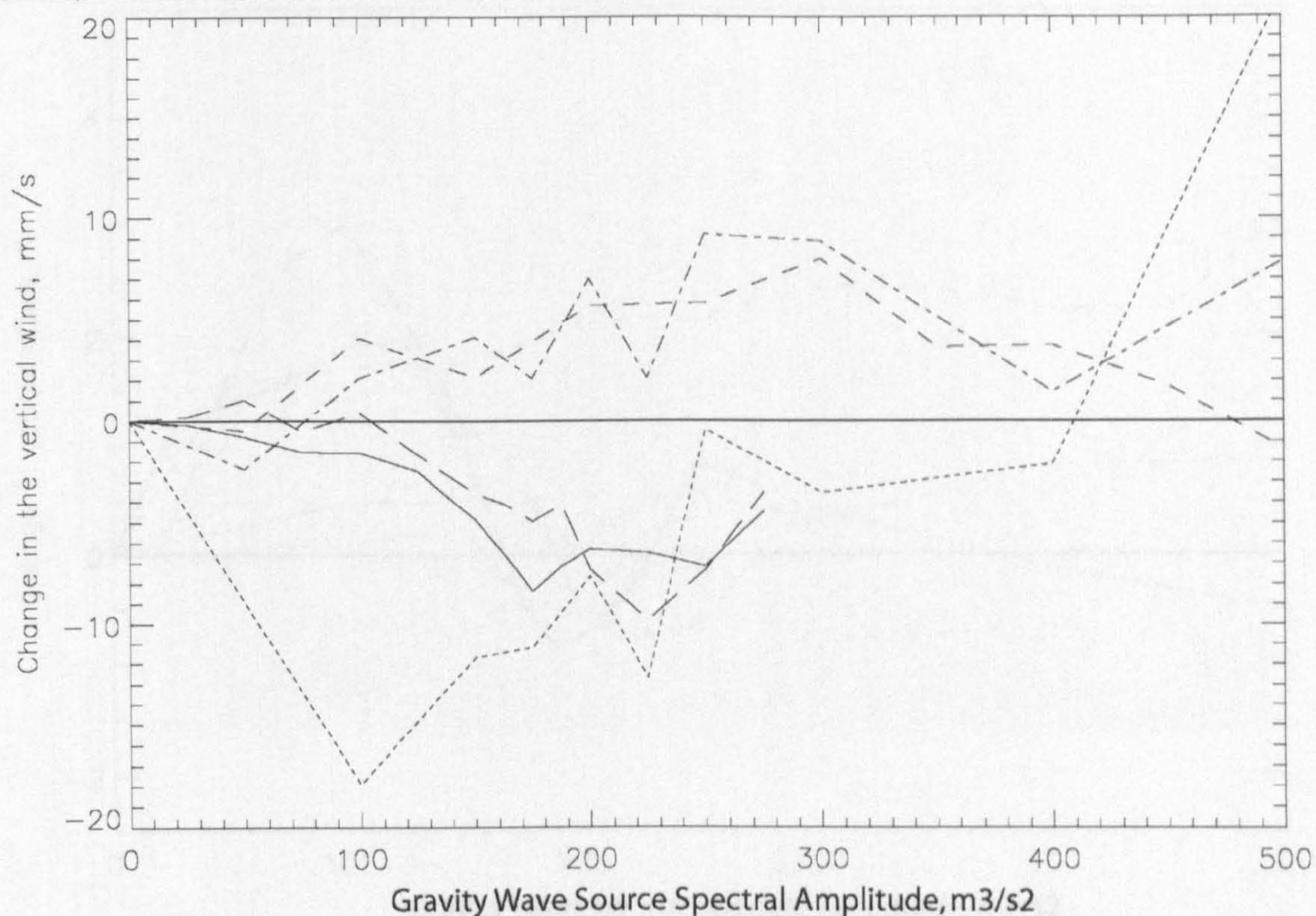


Figure V.10, As figure (V.9), but for the change in the vertical component of the residual circulation from zero gravity-wave source spectral amplitude (positive downwards), averaged over 10th – 20th January, between 2.4×10^{-1} – 2.4×10^{-3} mb (~ 58 – 88 km), between 60° – 90° North and zonally averaged.

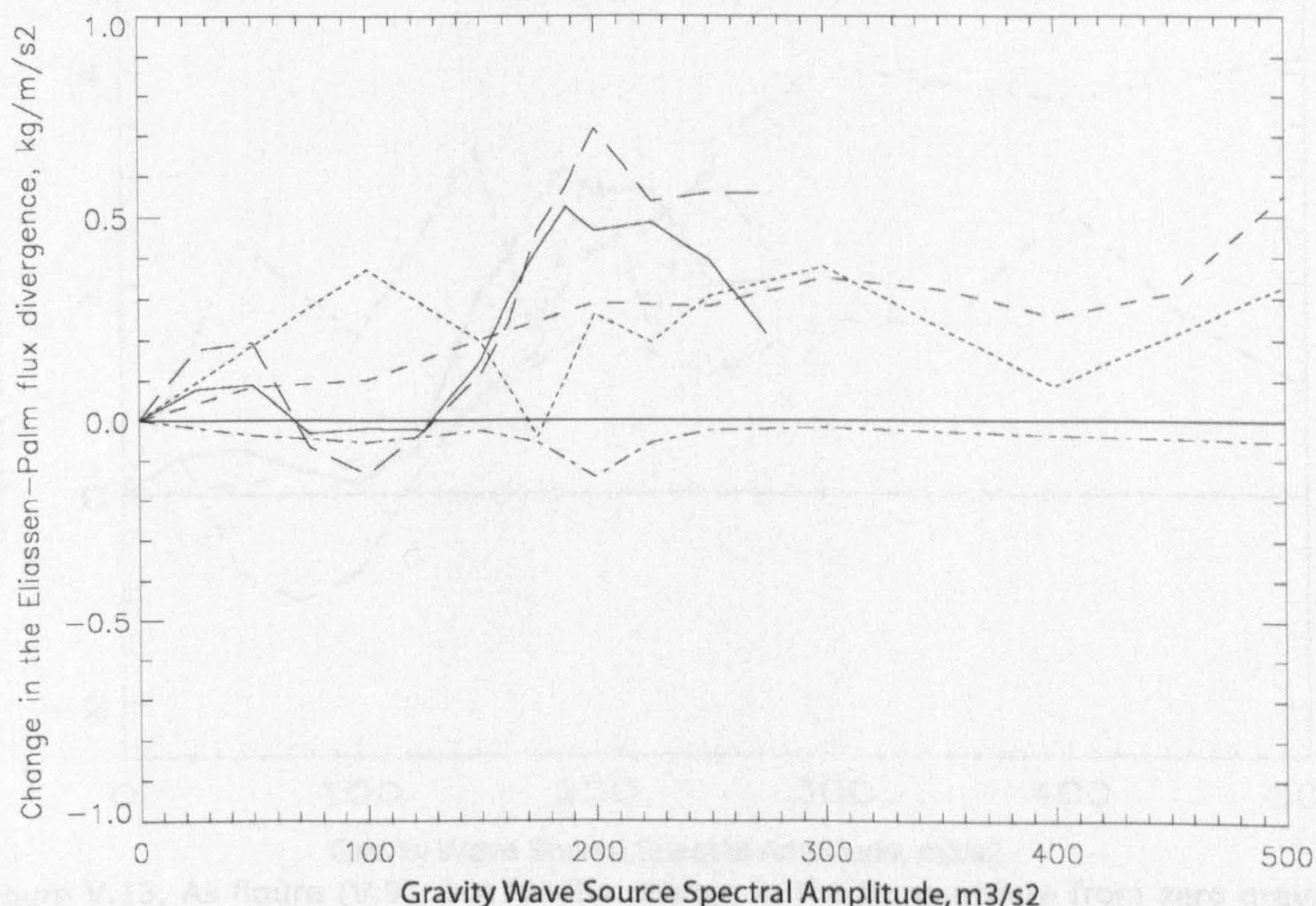


Figure V.11, As figure (V.9), but for the change in the Eliassen-Palm flux divergence from zero gravity-wave source spectral amplitude, averaged over 10th – 20th January, between 32 – 4.2 mb (~ 19 – 32 km), between 75° – 90° North and zonally averaged.

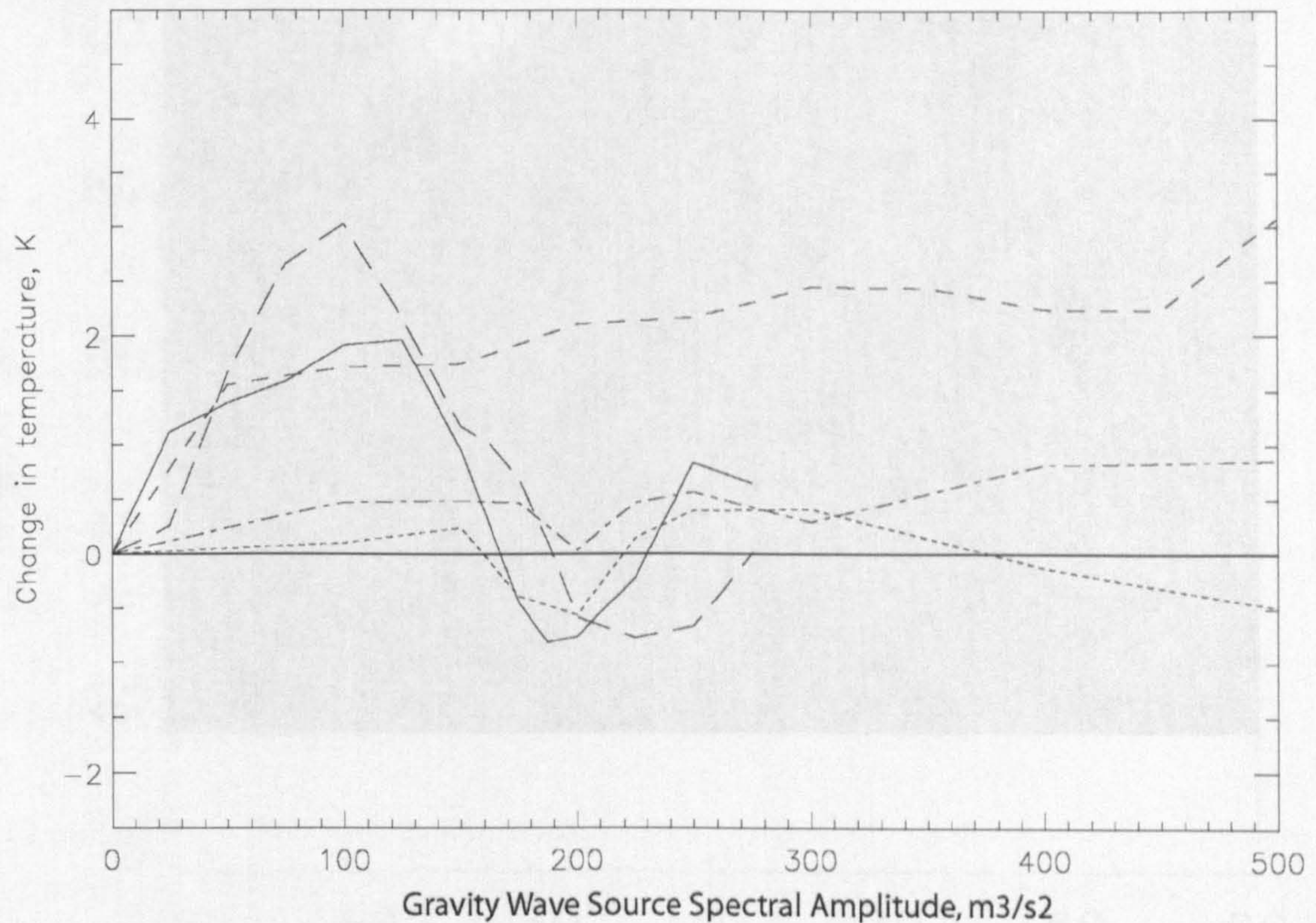


Figure V.12, As figure (V.9), but for the change in the temperature from zero gravity-wave source spectral amplitude, averaged over 10th – 20th January, between 32 – 4.2 mb ($\sim 19 - 32$ km), between 75° – 90° North and zonally averaged.

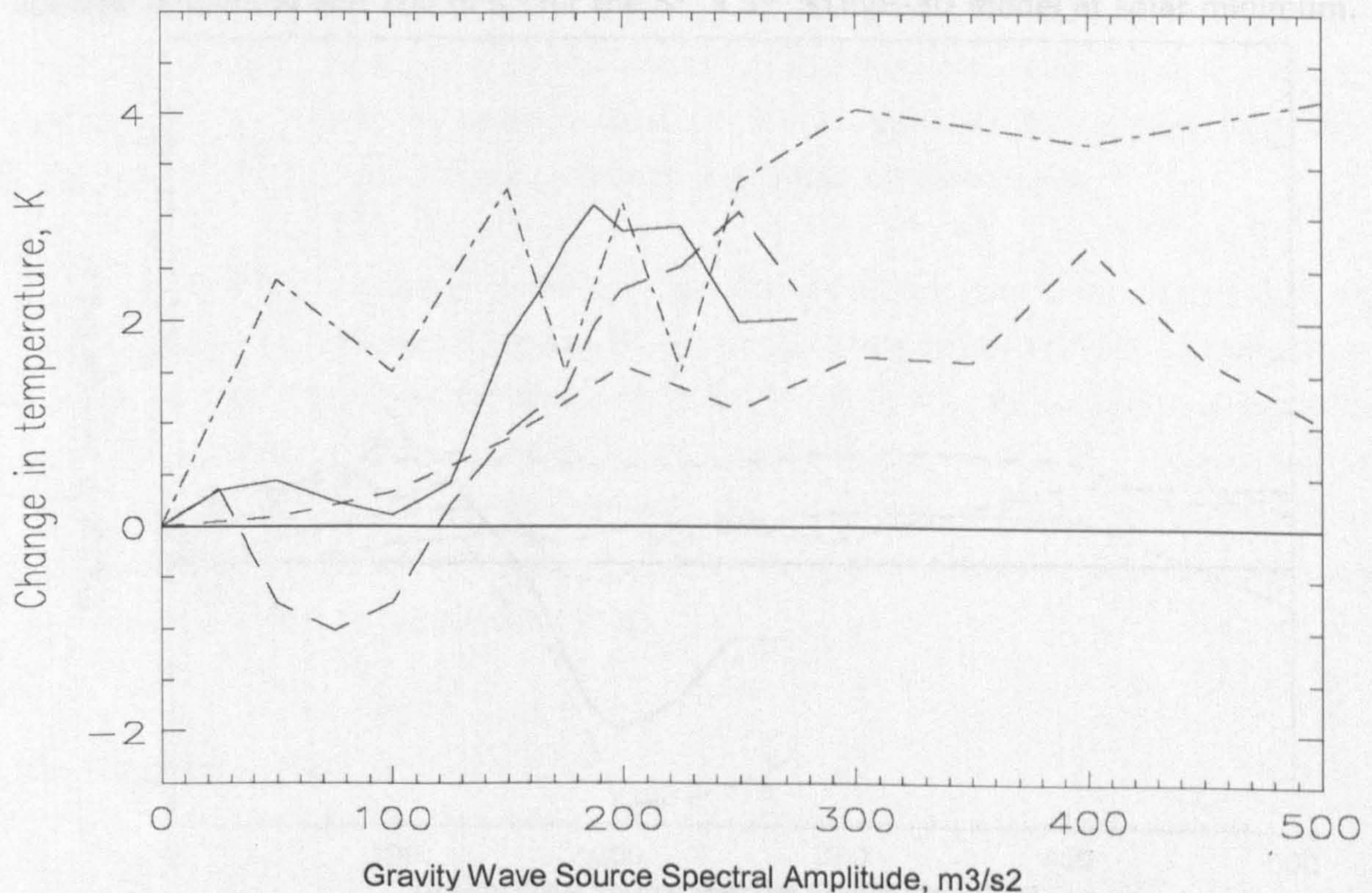


Figure V.13, As figure (V.9), but for the change in the temperature from zero gravity-wave source spectral amplitude, averaged over 10th – 20th January, between 2.4×10^{-1} – 2.4×10^{-3} mb ($\sim 58 - 88$ km), between 75° – 90° North and zonally averaged.

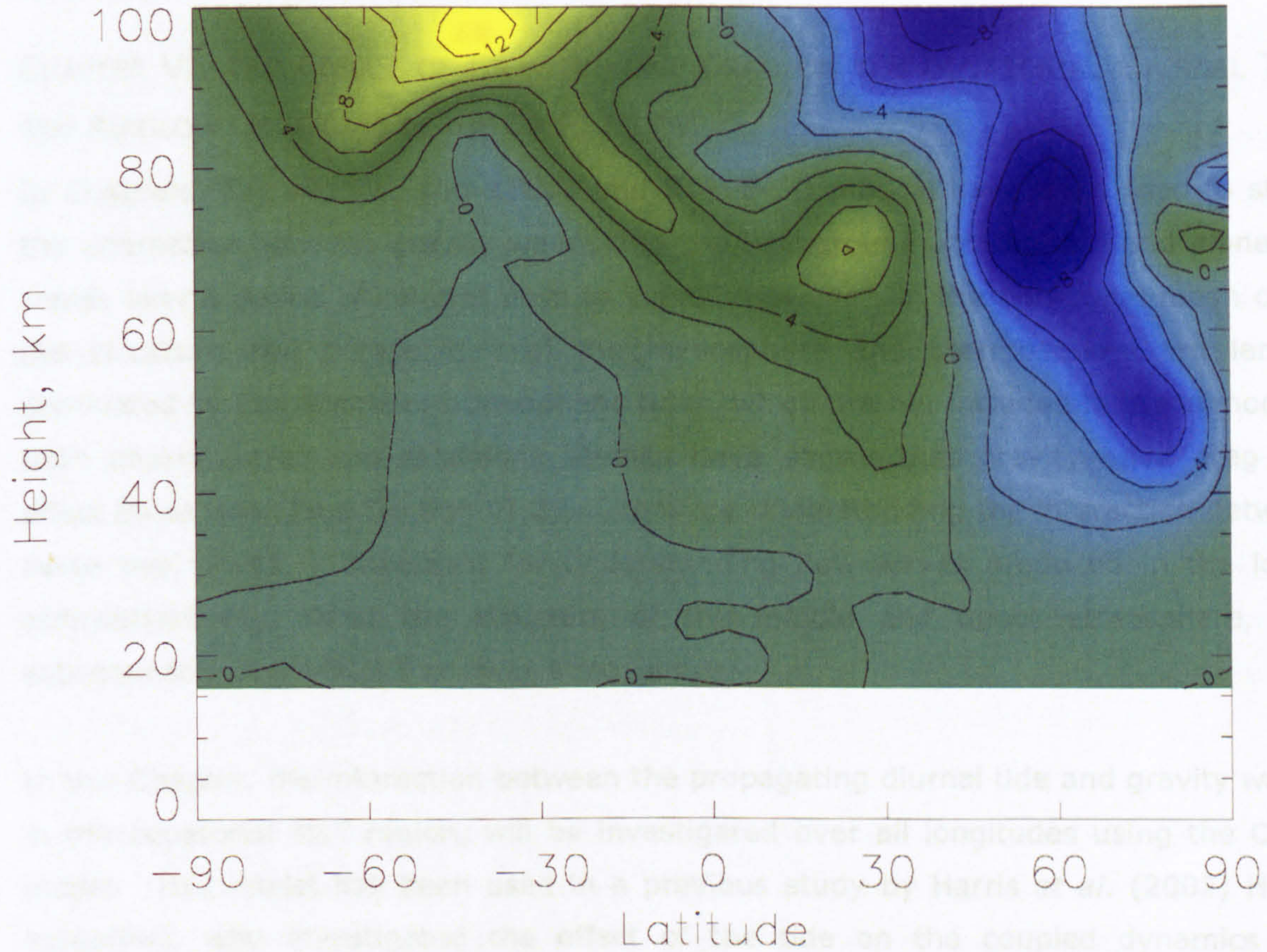


Figure V.14, Change in the zonal mean zonal wind between zero gravity-wave source spectral amplitude and $100 \text{ m}^3\text{s}^{-2}$ for the $5^\circ \times 5^\circ$ STEVE-3D model at solar minimum.

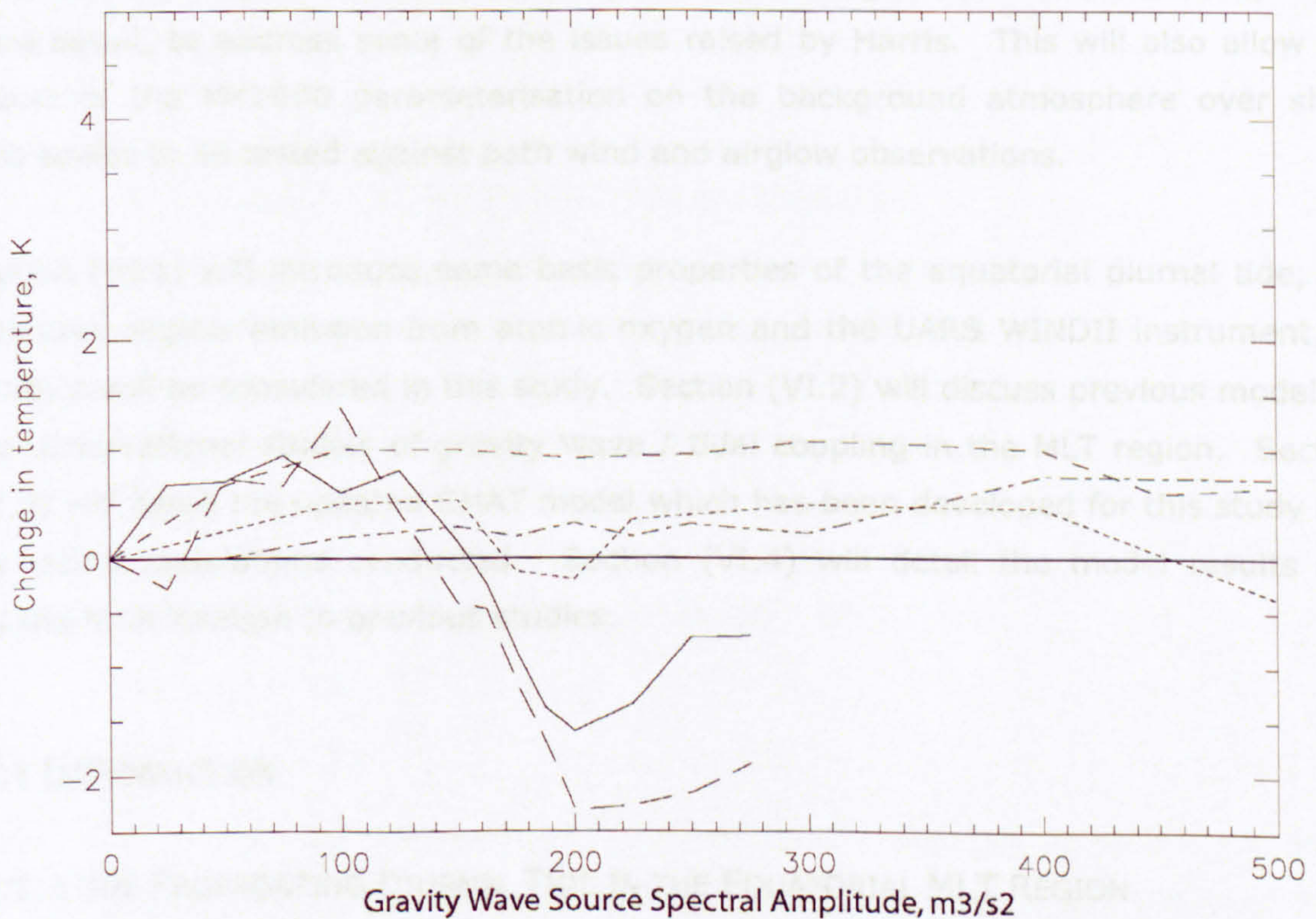


Figure V.15, As figure (V.9), but for the change in the temperature from zero gravity-wave source spectral amplitude, averaged over 1st – 31st January, between 32 – 4.2 mb (~ 19 – 32 km), between 75° – 90° North and zonally averaged.

CHAPTER VI, THE EFFECT OF GRAVITY-WAVE DRAG ON THE EQUATORIAL DIURNAL TIDE AND AIRGLOW EMISSIONS IN THE CMAT MODEL

In Chapters (IV) and (V), the STEVE and STEVE-3D models have been used to study the interaction between gravity waves and the background atmosphere and planetary waves over a period of several days to years. However, in the non-zonal mean case, the structure and composition of the mesosphere and thermosphere are largely dominated by the effects of atmospheric tides, which are not included in these models. Both observational and modelling studies have shown that gravity-wave drag can affect these tides (see Section VI.2). Therefore understanding the interaction between these two waves is important for understanding how waves produced in the lower atmosphere can affect the structure of the middle and upper atmosphere, and subsequently may affect the lower atmosphere.

In this Chapter, the interaction between the propagating diurnal tide and gravity waves in the equatorial MLT region, will be investigated over all longitudes using the CMAT model. This model has been used in a previous study by Harris *et al.* (2002; Harris hereafter), who investigated the effect of the tide on the coupled dynamics and composition of the MLT region. This Chapter will consider an improved version of the CMAT model, which will be used to study the effects of gravity wave / tidal coupling in more detail, to address some of the issues raised by Harris. This will also allow the effects of the MK2000 parameterisation on the background atmosphere over short time-scales to be tested against both wind and airglow observations.

Section (VI.1) will introduce some basic properties of the equatorial diurnal tide, the greenline airglow emission from atomic oxygen and the UARS WINDII instrument, all of which will be considered in this study. Section (VI.2) will discuss previous modelling and observational studies of gravity wave / tidal coupling in the MLT region. Section (VI.3) will detail the updated CMAT model which has been developed for this study and the model simulations conducted. Section (VI.4) will detail the model results and discuss their relation to previous studies.

VI.1 INTRODUCTION

VI.1.1 THE PROPAGATING DIURNAL TIDE IN THE EQUATORIAL MLT REGION

The propagating diurnal tide plays a major role in the dynamics of the equatorial mesosphere and lower thermosphere (MLT) region. Its general structure as described by classical tidal theory (Chapman and Lindzen, 1970) is in overall agreement with

observations, with a vertical wavelength in the region of 30 km, period of 24 hours and a zonal wave number of 1. The tide is primarily excited by the absorption of solar radiation by water vapour in the troposphere (Forbes, 1982), with secondary excitations including the release of latent heat in the tropics, the absorption of solar radiation by O₃ in the stratosphere and mesosphere and by O₂ in the lower thermosphere (e.g. Hagan, 1996). Figure (I.11) shows the meridional winds measured by the UARS HRDI instrument in the equatorial MLT region during March 1993 (Yudin *et al.*, 1997). The meridional wind pattern is dominated by the strong tidal signature of alternate compression and expansion regions, which peak at around $\pm 30^\circ$ latitude, ~ 100 km altitude, with a vertical wavelength of ~ 25 km. The temperature perturbations associated with the diurnal tide are observed to maximise at the equator. The diurnal tide is generally the largest amplitude wave in the equatorial MLT region and has a major impact on not only the dynamics of the region, but also its composition (e.g. Akmaev and Shved, 1980).

Gravity waves propagating through this region would therefore be affected by these tidal wind and temperature fluctuations. This will alter the gravity-wave drag in the middle atmosphere, which in turn will affect the winds and temperatures. This interaction, and its effects will be discussed in Sections (VI.2) and (VI.4). The effects of this tide on the greenline airglow from atomic oxygen will be discussed in the following Section.

VI.1.2 THE ATOMIC OXYGEN GREENLINE AIRGLOW EMISSION

557.7nm greenline airglow emission is seen in the MLT region, peaking in amplitude at ~ 95 km altitude. This frequency is known to be associated with the radiative de-excitation of O(¹S) to O(¹D), but no firm proof of the dominant excitation mechanism which explains the abundance of O(¹S) in the MLT region exists.

Barth (1964; Barth hereafter) proposed a mechanism which could explain the excitation of O(¹S) through the following reactions:



where M is a molecular species (most likely N₂) and * denotes an excited molecular state. Murtagh *et al.* (1990) showed that the Barth mechanism was the most likely excitation mechanism to explain the observed emission rates in the MLT region.

Tidal dissipation has been shown to affect the mean atomic oxygen density in the MLT region by altering its recombination rate (Akmaev and Shved, 1980) through an induced mean circulation (figure VI.2). The diurnal oscillation in density and motion of the air induced by the tide also creates a diurnal cycle in the $O(^1S)$ emissions (see Section VI.4). The impact of altering the gravity wave / diurnal tide interactions in the CMAT model on $O(^1S)$ emission rates in the equatorial MLT region will be discussed in more detail in Section (VI.4).

VI.1.3 THE UARS WIND IMAGING INTERFEROMETER (WINDII)

The UARS WINDII instrument (Shepherd, 1985) measured the Doppler shift and width of spectral emission lines from airglow and aurora in order to find the horizontal winds, temperatures and airglow emission rates in the MLT region. Two viewing angles were used in order to resolve both components of the horizontal wind over the 70 – 310 km altitude range. The instrument had a resolution of 1.5 km vertical, 20 km along track, 6° wide and 8 second integration time per image. It observed two emission lines from atomic oxygen (including the 557.7 nm greenline emission), as well as two lines from OH, one from O^+ and the 762.0 nm band from O_2 in order to obtain measurements over this whole height range during both day and night. The orbit of the spacecraft meant that all local solar times can be sampled in 33 days

Data from this instrument will be used throughout this Chapter as it was capable of measuring both the dynamical and compositional effects of the diurnal tide throughout the equatorial MLT region, where the effects of the tide are strongest. Yudin *et al.* (1998) have shown that the WINDII measurements of the airglow and tidal winds are consistent with each other. The average night-time WINDII observations of $O(^1S)$ airglow in the equatorial MLT region, which will be used in Section (VI.4) have been shown to be in good agreement with ground based observations (e.g. Wiens *et al.*, 1999) and the horizontal wind measurements are in good overall agreement with ground based radar observations (with the exception of winds measured by MF radars above ~85 km, Burrage *et al.*, 1996; McLandress *et al.*, 1996).

VI.2 PREVIOUS MODELLING AND OBSERVATIONAL STUDIES OF THE INTERACTION BETWEEN GRAVITY WAVES AND ATMOSPHERIC TIDES IN THE MLT REGION

Gavrilov *et al.* (1981) first observed a correlation between the intensity of gravity waves and the tidal winds, and Fritts and Vincent (1987) proposed a simple model by which the tidal wind shears could modulate the gravity wave spectrum in agreement with their observations. The authors showed that the vertical divergence of the flux of

horizontal momentum associated with the dissipating gravity waves (gravity-wave drag), acted to advance the phase fronts of the tide and decrease its amplitude. They concluded that this could account for much of the temporal variability they observed in the diurnal tide using an MF radar at Adelaide (35 °S, 138 °E). This result has been supported by other observational campaigns (e.g. Wang and Fritts, 1991; Nakamura *et al.*, 1997; Thayaparan *et al.*, 1995) and some idealised models such as Eckermann and Marks (1996).

Attempts to study gravity wave / tidal interactions using global-scale models have yielded some contradictory results regarding the impact of parameterised gravity-wave drag on the propagating diurnal tide. Miyahara and Forbes (1991), Forbes *et al.* (1991), McLandress and Ward (1994) and Meyer (1999) used parameterisations based on the work of Lindzen (1981) and Matsuno (1982) to study the effect of gravity-wave drag on tidal amplitudes and found the parameterised gravity-wave drag acted to decrease the amplitude of the tide and advance the tidal phase. In the same paper, Meyer (1999) showed the relative effects of parameterised gravity-wave drag and eddy diffusion on the amplitude of the diurnal tide using the Global Scale Wave Model (GSWM) (Hagan *et al.*, 1995), and demonstrated that vertical eddy diffusion significantly damps the amplitude of the diurnal tide. However, using spectral parameterisations of gravity-wave drag, a number of authors have shown that parameterised gravity-wave drag can increase the amplitude of the tide (e.g. Lu and Fritts, 1993; Mayr *et al.*, 1998). McLandress (1998) compared the effects of several different gravity wave parameterisations in the same mechanistic tidal model and found that the diurnal tidal amplitude could be either increased or decreased depending on the parameterisation used, but all were found to advance the phase of the tide by varying amounts. Manson *et al.* (2002) compared observations from a number of MF radars at different latitudes with results from the Canadian Middle Atmosphere Model (Beagley *et al.*, 1997) using both the Hines and the MK2000 gravity wave parameterisations. They reported a complex pattern of tidal amplitude and phase variations with the two parameterisations, dependant upon season, latitude and frequency of the tide (diurnal or semi-diurnal). In general the phases of the tide modelled with MK2000 were more consistent with low latitude MF radar observations than those modelled with the Hines parameterisation. It is worth noting that many of the results they presented were for 78 km altitude, which the authors noted was not the ideal altitude to assess the impact of the parameterised gravity-wave drag on the tides as the amplitude of the tides is still relatively small at this height. The authors suggested that a height of 90 km would be better for such a study and did present their model results up to 88 km altitude, however these lay within the Rayleigh friction sponge layer at the top of their model.

Following a similar method to Roble and Shepherd (1997) who used the National Center for Atmospheric Research Thermosphere – Ionosphere – Mesosphere – Electrodynamics General Circulation Model (NCAR TIME-GCM; Roble and Ridley, 1994), Harris used the Coupled Middle Atmosphere and Thermosphere (CMAT) general circulation model to study the effect of the diurnal tide on the dynamics, energetics and composition of the low to mid latitude mesosphere and thermosphere during equinox. The authors noted that when using the diurnal tidal fields from the GSWM to provide a lower boundary tidal forcing, the CMAT model underestimated the diurnal tidal amplitude in the MLT region (see figures VI.1 and I.11). The authors also noted that the model did not reproduce the observed zonal mean zonal wind structure in the subtropical mesosphere. The authors hypothesised that both of these discrepancies between the model and observations were due to the representation of gravity waves in the model, which were parameterised using a hybrid Matsuno-Lindzen parameterisation developed by Meyer (1999; Meyer hereafter). Furthermore, the authors noted that the CMAT model underestimated the magnitude of the atomic oxygen 557.7 nm $O(^1S)$ green line volume emission rates in the equatorial MLT region due to an underestimation of the number densities of atomic oxygen in this region. Both Harris and Roble and Shepherd (1997) suggested that this may be in part due to an overestimation of the vertical eddy diffusion in the two models.

In this Chapter the CMAT model is used to compare the effects of the Meyer and MK2000 gravity wave parameterisations on both the dynamics (horizontal winds, tidal amplitude and phase) and the chemical composition (through the $O(^1S)$ green line volume emission rates) during equinox at solar minimum conditions. These effects are examined in detail over the low latitude upper mesosphere and lower thermosphere region, between 70 km and 120 km altitude, where the amplitude of the diurnal tide maximises. Section (IV.3) will briefly outline the CMAT model and the simulations performed in this study. Section (IV.4) will present the results of these simulations, followed by a discussion of the effects of the two gravity wave parameterisations on the dynamics and green line airglow emission rates in the low latitude mesosphere and lower thermosphere. As the gravity-wave drag in the model impacts on both of these, both data sets can be used to assess the representation of parameterised gravity-wave drag in the model. The model results are primarily compared against the observations of the UARS WINDII instrument (Shepherd *et al.*, 1993), as this simultaneously measures both the $O(^1S)$ green line emission rates and the horizontal winds over the region of interest, and these observations have been shown to be consistent with each other (Yudin *et al.*, 1998). The implications of the results presented here on the work of Harris are also discussed.

VI.3 THE COUPLED MIDDLE ATMOSPHERE AND THERMOSPHERE GENERAL CIRCULATION MODEL

VI.3.1 CMAT MODEL DESCRIPTION

CMAT is an extension of the University College London time dependent 3 – Dimensional Coupled Thermosphere Ionosphere Plasmasphere (CTIP) model (e.g. Fuller-Rowell *et al.*, 1996; Millward *et al.*, 1996). This extended model has a vertical range from 10 mbar (~ 30 km) to 7.6×10^{-9} mbar ($\sim 300 - 600$ km depending on thermospheric temperature) and couples the upper stratosphere and mesosphere to the thermosphere / ionosphere system. The model is ideal for exploring the impact of the lower atmosphere on the dynamics and electrodynamics in the upper atmosphere, and *vice versa*. A detailed description of CMAT is given by Harris (2001), and a summary of its key points is given below.

The model solves the primitive equations using Eulerian finite difference calculations on a grid which has a 1/3 pressure scale height vertical grid spacing, 2° latitudinal and 18° longitudinal resolution. This corresponds to a grid of 63 vertical pressure levels, 91 latitudinal grid points (pole to pole) and 20 longitudinal grid points. To avoid numerical instabilities in the diffusive thermospheric region, a time-step of 60 s is used in the CMAT model (Harris, 2001).

Thermospheric heating, photodissociation and photoionisation due to the absorption of solar X-ray, EUV and UV radiation between 0.1-180 nm are calculated using fluxes from the SOLAR2000 empirical model (Tobiska *et al.*, 2000). Absorption and ionisation cross Sections between 0.1 and 105 nm are based on Henke (1993) and Fennelly and Torr (1992), and between 105 – 194 are provided by Rodney Viereck (2003; private communication). The EUV heating efficiencies from Roble (1987) are used.

Heating due to the dissociation of O_3 in the Chappuis, Huggins and Hartley bands, and O_2 in the Schumann-Runge Bands are due to Strobel (1978). The effect of exothermic neutral chemical heating is also calculated in the middle atmosphere. Heating efficiencies given by Mlynczak and Solomon (1993) are used to account for loss due to airglow in this region.

The $5.3 \mu\text{m}$ emission of NO from Kockarts (1980), the $63 \mu\text{m}$ fine structure emission of O from Bates (1951) and Fuller-Rowell (1998; private communication), the LTE and non- LTE $15.6 \mu\text{m}$ emission of CO_2 , with an atomic oxygen collisional deactivation rate

coefficient of $3.5 \times 10^{-12} \text{ s}^{-1}$ (Fomichev *et al.*, 1998) and the $9.6 \mu\text{m}$ emission of O_3 from Fomichev and Shved (1985) are used to calculate the IR heating and cooling of the atmosphere.

A mesosphere-thermosphere neutral chemical scheme (Allen *et al.*, 1984; Solomon *et al.*, 1985; Fuller-Rowell, 1992) is included, incorporating JPL-97 reaction rate coefficients from DeMore *et al.* (1997). The effect of transport on major constituents (O_2 , N_2 and $\text{O}_x = \text{O} + \text{O}_3$) is solved using Fuller-Rowell *et al.* (1996). Minor species transport is solved for $\text{N}(^4\text{S})$, $\text{N}(^2\text{D})$, NO , $\text{HO}_x = \text{OH} + \text{HO}_2 + \text{H}$, H_2O , H_2 , CO , CO_2 , CH_4 , NO_2 , $\text{O}(^1\text{D})$, and He . H_2O_2 and $\text{O}(^1\text{D})$ are assumed to be in a state of photochemical equilibrium. The calculation of $\text{O}(^1\text{S})$ green line emission assumes that $\text{O}(^1\text{S})$ is excited by the Barth mechanism (Barth, 1964), using the excitation rates given by Murtagh *et al.* (1990).

The high-latitude ionospheric parameters are obtained from the Sheffield University Coupled High Latitude Ionosphere Model (Quegan *et al.*, 1982; Fuller-Rowell *et al.*, 1996). This model calculates electron density, electron and ion temperatures, field aligned velocities and distribution of ion species assuming photochemical equilibrium. High latitude auroral precipitation from the TIROS / NOAA auroral precipitation statistical model are included (Fuller-Rowell and Evans, 1987; Codrescu *et al.*, 1997) along with electric field strengths from Foster *et al.* (1986). The effect of high latitude small-scale electric field variability is included (Codrescu *et al.*, 2000).

Lower boundary seasonal forcing from MSIS-E90 (Hedin 1991) and lower boundary tidal forcing from GSWM output (Hagan, private communication 1999) are used with a diurnal (1,1) oscillation of geopotential height of 14 m and a phase of 20.8 hours. The vertical gradients in the temperature and wind fields are set to zero across the upper boundary of the model.

Gravity-wave drag can be calculated from a hybrid Matsuno-Lindzen parameterisation of Meyer or using the non-linear diffusion parameterisation of MK2000. Details of the source spectra used in these two parameterisations are given in Section (VI.3.3). The vertical eddy diffusion coefficients are taken as a height dependent global mean based on the climatology calculated by Garcia and Solomon (1985).

VI.3.2 CMAT MODEL CONFIGURATION

All the CMAT model simulations presented here are for 40 days in steady state at the March equinox during solar minimum conditions ($10.7 \text{ cm solar flux} = 76 \times 10^{-22}$

$\text{Js}^{-1}\text{m}^{-2}\text{Hz}^{-1}$) and low geomagnetic activity ($K_p = 2+$, auroral power = 8 GW, cross polar cap electric potential = 36 kV). Three simulations were performed, which shall be referred to as A, B and C hereafter. The global mean height dependent, vertical thermal eddy coefficients based on the climatology of Garcia and Solomon (1985) are used for all model runs.

- A.** As a control test, the CMAT model was run without any parameterised gravity-wave drag. However, it is worth noting that this simulation does not represent the atmosphere in the absence of gravity waves (as the diffusion created by gravity waves is still present), but rather the atmosphere in the absence of gravity-wave drag. This allows the effects of parameterised gravity-wave drag to be found.
- B.** Gravity-wave drag calculated using the Meyer parameterisation was used in the CMAT model, while all other inputs remained the same as for simulation A. This test essentially repeats the standard run from Harris, but includes updates made to the thermospheric heating routine.
- C.** Gravity-wave drag calculated by the MK2000 parameterisation was used, while all other inputs remained the same as simulations A and B. This allows a direct comparison of the effects of the Meyer and MK2000 gravity-wave drags on the background atmosphere to be made.

VI.3.3 CONFIGURATION OF THE GRAVITY WAVE PARAMETERISATIONS USED IN THE CMAT MODEL

This Section details the source spectra and implementation of the Meyer and MK2000 gravity wave parameterisations used within the CMAT model.

The Meyer parameterisation uses a source spectrum which consists of 19 waves in both zonal and meridional directions. The distribution of wave amplitudes is Gaussian in shape, maximising for waves with horizontal phase speeds of 0 ms^{-1} and is constant with time. Wave phase velocities range between $\pm 60 \text{ m s}^{-1}$ with the zero velocity wave having an amplitude 20 times that of the $\pm 60 \text{ m s}^{-1}$ waves. The spectrum is filtered between $\sim 20 \text{ km}$ and the lower boundary at $\sim 30 \text{ km}$ using wind data from the Horizontal Wave Model (Hedin *et al.*, 1993).

The MK2000 gravity wave parameterisation used in the CMAT model is essentially the same as the implementation used in the STEVE and STEVE-3D models, outlined in Section (III.3.4), although an increased mean source amplitude and a different value

of m_* is used in the CMAT model compared with that given in STEVE to account for the difference in the altitude of the source (these values were found by propagating the gravity wave spectrum between 100 mb and 10 mb using global mean background conditions). As the MK2000 gravity wave parameterisation has only been fully tested using a ~ 2 km grid spacing (Medvedev *et al.*, 1997; McLandress, 1998; Medvedev and Klaassen, 2000; Manson *et al.*, 2002; Chapters III, IV and V of this thesis), the gravity-wave drags in the CMAT model are calculated on a quarter pressure scale height resolution grid (equivalent to that used by the STEVE and STEVE-3D models). The heights of these grid points are defined using a global mean profile from MSIS-E90 and all model inputs are extrapolated onto this grid using a cubic spline interpolation. As with the STEVE and STEVE-3D models, no account is taken of the effects of filtering of the spectrum below the bottom boundary of the CMAT model in this parameterisation.

VI.4 CMAT MODEL RESULTS AND DISCUSSION

VI.4.1 THE EFFECT OF PARAMETERISED GRAVITY-WAVE DRAG ON THE ZONAL MEAN ZONAL WINDS

Figure (VI.3) shows the zonal mean zonal winds for the March equinox from CMAT model simulations A, B and C. As might be expected, the mid-latitude eastward jets in the middle atmosphere remain open in simulation A (the jet closure above 120 km is the result of ion-neutral collisions which dominate over the thermal winds in the model at these altitudes) and the amplitude of these jets greatly exceeds that observed by UARS (figure VI.9). A strong westward flow is seen in the equatorial MLT region in simulation A, which is driven by the dissipating diurnal tide in this region. The eastward flow in the equatorial middle atmosphere, shown in the combined WINDII and HRDI zonal winds between 60 – 75 km is not present in this simulation.

The effects of parameterised gravity-wave drag on the zonal mean winds in the middle atmosphere have been well documented (see Chapters I to III). Figure (VI.4) shows the zonal mean zonal gravity-wave drag calculated by the CMAT model for simulations B and C. The effects of this parameterised gravity-wave drag on the zonal mean zonal winds can be seen by comparing model simulations B and C (figure VI.3b and VI.3c) to the control simulation A in figure (VI.3a). In simulation B, the westward gravity-wave drag acts to weaken the modelled mid-latitude eastward jets. Compared with simulation A, these mid-latitude jets are in better agreement with the UARS zonal winds. However, as noted by Harris, the jets are still stronger than those observed and the model fails to reproduce the jet reversal seen in the climatology in the lower

thermosphere. In simulation B there is almost no zonal mean zonal wind in the equatorial middle atmosphere. The westward zonal mean zonal wind, seen in simulation A, has been reduced to almost zero. The eastward flow seen in the mesosphere in the UARS observations is still absent.

In the mid-latitude stratosphere and mesosphere, the effects of the parameterised gravity-wave drag in simulation C is similar to the effect seen in simulation B, although the jets are not reduced as much. However, at around 90 km altitude, a jet reversal is seen in simulation C, which was absent in B. The zonal mean zonal gravity-wave drags in simulations B and C differ at mid-latitudes. The Meyer parameterisation produces a smaller westward drag, which peaks at lower altitudes than that produced by the MK2000 parameterisation. This accounts for the weaker eastward winds in the stratosphere and lower mesosphere in simulation B. In the upper mesosphere and lower thermosphere, the higher value drags produced in simulation C cause the jet reversal in the model. In the equatorial MLT region, the parameterised gravity-wave drags produced by the Meyer and MK2000 schemes differ greatly. While the drags in simulation B are almost zero in this region, the MK2000 parameterisation produces a strong eastward drag. This acts to reduce the tidally induced westward flow seen in simulation A, and to create a weak eastward flow of around 20 m s^{-1} above 90 km. The detailed structure of the flow in the equatorial middle atmosphere in simulation C does not agree fully with the UARS observations, but this simulation is in better agreement than simulations A and B.

VI.4.2 THE EFFECT OF PARAMETERISED GRAVITY-WAVE DRAG ON THE MERIDIONAL WINDS

The diurnal tidal component of the meridional wind often exceeds the mean meridional wind in the middle atmosphere. This makes the tidal amplitudes and phases clearly visible in the model data. Figure (VI.5) shows the amplitude of the diurnal oscillation in the meridional winds in the equatorial MLT region for simulations A, B, and C, which can be compared to the WINDII observations (figure VI.10). The peak amplitude of the meridional diurnal tidal winds in simulation A is around 65 m s^{-1} , which is in good agreement with the WINDII measurements of around 70 m s^{-1} . However, there is a discrepancy between the model and observations in the location of this peak. The model simulation A shows it to be centred on $\pm 25^\circ$ latitude, 115 km altitude, while the observation shows it to be centred on $\pm 20^\circ$ latitude at 100 km altitude.

In simulation B the peak diurnal amplitude in the meridional winds is around 25 m s^{-1} , which occurs between $\pm 30 - 45^\circ$ latitude and between 110 – 120 km altitude. While the lower altitude of this peak brings the model into better agreement with the WINDII

measurements than simulation A, the peak amplitude has been decreased to around one third of that observed by WINDII. Figure (VI.6) shows the phase fronts of the meridional wind and meridional gravity-wave drag as a function of height and local time at 20° South for simulations B and C. In simulation B, the meridional winds and meridional gravity-wave drags are almost completely out of phase, which leads to the reduction in tidal amplitude and is thought to be the cause of the reduced tidally driven westward flow seen in the zonal mean zonal winds in simulation B. This agrees with the results of Meyer (1999), who used the same gravity wave parameterisation to study the interaction between the diurnal tide and gravity waves. The out of phase relationship between the gravity-wave drag and meridional winds can be explained by the simple model proposed by Fritts and Vincent (1987), which the authors used to explain their MF radar observations at Adelaide. This model proposes that gravity waves with phase velocities in the same direction as the tidal wind are preferentially reduced in amplitude or removed by critical layer absorption, while those which propagate in the opposite direction experience an increase in amplitude with altitude, resulting in an increase in acceleration of the background wind.

From figure (VI.5c) it can be seen that the MK2000 gravity-wave drag has increased the peak amplitude of the meridional tidal winds in the CMAT model to around 85 m s^{-1} at 110 km and $\pm 20^\circ$ latitude. This peak amplitude is higher, but still comparable with that measured by WINDII and the location of the peak is in better agreement with the observations for this model simulation than either A or B. The increase in amplitude of the diurnal tidal winds can be attributed to the nearly in-phase relationship between the meridional gravity-wave drag and meridional winds, shown in figure (VI.6b). This enhancement of the tidal amplitude has been reported by Manson *et al.* (2002), who used the MK2000 gravity wave parameterisation in the Canadian Middle Atmosphere Model, and several other authors who have used spectral gravity wave parameterisations (e.g. Mayr *et al.*, 1998; McLandress, 1998). The gravity wave parameterisations used in these studies included wave-wave as well as wave-mean flow interactions, whereas the Lindzen (1981) based parameterisations and the conceptual model of Fritts and Vincent (1987), which show a reduction in tidal amplitude due to gravity-wave drag, only account for gravity wave-background atmosphere interactions. However, it is worth noting that a clear distinction between these two classes of gravity wave parameterisations cannot be made as the Meyer parameterisation includes some wave-wave interaction as the eddy diffusion coefficient used to calculate the breaking height of one wave takes into account the turbulent diffusion created by the breaking of other waves in the spectrum. The validity of using a broad spectrum of gravity waves in parameterisations is still a matter of some debate as there is observational evidence that a spectrum of just a few waves may

better describe the gravity wave field (see Fritts and Alexander, 2003 for a recent review) and further work in this area will be needed to constrain gravity wave-tidal interactions in global-scale models.

The difference in the phase relationship between the meridional winds and the gravity-wave drag calculated by the Meyer and MK2000 schemes can be accounted for. The Meyer parameterisation uses a hybrid of Lindzen (1981) and Matsuno (1982) gravity wave parameterisations to calculate gravity-wave drag. By considering equations (II.4), it can be seen that within a region of dissipation (above the breaking height) the drag calculated by the Lindzen scheme will tend to maximise in the opposite direction to the background wind. This accounts for the out of phase relationship between the meridional gravity-wave drag and winds in experiment B.

Using a winter-time CIRA-86 background atmosphere, Medvedev and Klaassen (2000) showed that MK2000 calculates the gravity-wave drag to be several km below that calculated by the Lindzen scheme (while this background atmosphere does not match that calculated by CMAT, it seems reasonable that this result would apply here). This altitude difference accounts for the different phase relationships between the meridional gravity-wave drag and winds found in experiments B and C. Further, by using a non-interacting version of the MK95 theory (which included only wave-mean flow interactions), Medvedev and Klaassen (2000) showed that the MK2000 scheme predicted the gravity-wave drag to occur at this lower altitude only when wave-wave interactions were included. This explanation may also account for the similarity between the tidal-gravity wave interactions calculated by the Hines and MK2000 schemes, which both include wave-wave and wave-mean flow interactions (albeit in different forms), and the contrast with Lindzen-based schemes, which consider only wave-mean flow interactions.

Figure (VI.7) shows the phase fronts of the meridional winds at 20° South for model simulations A, B and C. The introduction of parameterised gravity-wave drag causes the phase fronts of the meridional winds to advance (downwards) by around 5 km in both simulations B and C. Similar advancement has been noted by many authors (e.g. Miyahara and Forbes, 1991; McLandress, 1998; Meyer, 1999) and is caused by the phase fronts of the meridional gravity-wave drag preceding the phase fronts of the meridional wind. This advancement brings the simulated tidal phase into better agreement with the UARS HRDI measurements (Yudin *et al.*, 1997).

From figure (VI.7) it can also be seen that the vertical wavelength of the diurnal tide in the MLT region is affected by the introduction of parameterised gravity-wave drag. In

model simulation A the vertical wavelength of the diurnal tide is found to be ~ 30 km, which is in agreement with the theoretical value, but larger than that typically observed (e.g. Hays *et al.*, 1994). In simulations B and C, the vertical wavelength of the diurnal tide in the MLT region is reduced to ~ 25 km, which is in good agreement with the observed wavelength. This shorter vertical wavelength, and hence stronger molecular diffusion in the lower thermosphere, is likely to be responsible for the fall in amplitude of the diurnal tide above 110 km in simulation C. This, along with changes in the background atmosphere due to the introduction of gravity-wave drag is the likely reason for the change in the location of the peak amplitude of the tidal winds.

VI.4.3 THE EFFECT OF PARAMETERISED GRAVITY-WAVE DRAG ON ATOMIC OXYGEN GREEN LINE EMISSION RATES

The $O(^1S)$ green line volume emission rate is related to the atomic oxygen number density and the background atmospheric density. Previous studies have shown that it is strongly affected by the diurnal tide (e.g. Shepherd *et al.*, 1995; Roble and Shepherd, 1997; Yudin *et al.*, 1998; Harris) as the tide influences both of these parameters. It might be expected therefore, that the parameterised gravity-wave drag, which affects the diurnal tide in the CMAT model, will in turn affect the $O(^1S)$ green line volume emission rates, although the details of this effect have not previously been shown. Figure (VI.8) shows the night-time green line emission rates at the equator, calculated by the CMAT model for simulations A, B and C, which can be compared with the WINDII observations (figure VI.11). The modelled volume emission rates in simulation A show a strong local time dependence. The peak value in the model of $160 \text{ cm}^{-3}\text{s}^{-1}$ at ~ 97 km is in reasonable agreement with the WINDII observed peak of $180 \text{ cm}^{-3}\text{s}^{-1}$ between 85 – 95 km. However, the phase of the local time variation is not in good agreement with the observations, presumably because the phase of the diurnal tide in this simulation lags the observed tide. In simulation B, there is very little local time variation in the volume emission rates, associated with the low amplitude tide in this simulation. The peak emission rate of $140 \text{ cm}^{-3}\text{s}^{-1}$ at ~ 98 km is lower than the peak seen in simulation A, again as a consequence of the low amplitude tide in simulation B. In simulation C, the high amplitude tide reproduces the local time dependence of the $O(^1S)$ volume emission rates reasonably well. The peak emission rate of $240 \text{ cm}^{-3}\text{s}^{-1}$ at 95 km, 18 LT is higher than that observed by WINDII, presumably because the tidal amplitude is larger than that observed. As with simulations A and B, the peak green line emission rate in the model is at a higher altitude than that observed and the emission decreases rapidly below 90 km, whereas the observations show relatively strong emissions down to 80 km. However, this model simulation reproduces many features seen in the observations which the

previous simulations do not. The descending peak in high emission rates between 18 and 22 LT seen in the WINDII observations is present in the model results, as is the local minimum around 0 – 4 LT associated with the low atomic oxygen concentrations (Shepherd *et al.*, 1995). The secondary peak around 4 – 6 LT is associated with the downward advection of atomic oxygen (Shepherd *et al.*, 1995), although the phase of the variation appears to lag the observations by ~1 hour. This is likely to be due to a slight phase difference in the modelled and observed diurnal tide. A shift of the modelled emission rates by 1 hour would decrease the peak night-time emission rates and increase the local peak at 6 LT, bringing them closer to the observed values.

One further result comes to light through comparing the simulations presented here to the CMAT simulations of Harris, in which the authors reported the modelled night time volume emission rates to be almost a factor of 2 lower than those observed. The authors suggested that this was related to an overestimation of the vertical eddy diffusion coefficients used in the model. However, reasonable night-time volume emission rates have been achieved here using the CMAT model in simulation C, in which the same eddy diffusion profile is used. The proposed reason for the improved green line emission rate profiles is the combined effect of the updated thermospheric heating routine and the MK2000 gravity wave parameterisation. The heating routine implemented in CMAT since Harris (2002) uses the latest values of photoionization and absorption rates along with high resolution empirical solar flux data. The result is an improved representation of photochemistry and associated atomic oxygen morphology. This, combined with the MK2000 gravity wave scheme has lead to the more realistic green line emission profiles presented herein.

VI.5 SUMMARY

In this Chapter, two different gravity wave parameterisations have been used to study the interaction between gravity-wave drag and the diurnal tide in the CMAT model. Parameterised gravity-wave drag has been seen to affect the amplitude and phase of the diurnal tide and the diurnal pattern of O(¹S) greenline emissions. These results have been compared with observed winds and airglow emission rates and several areas requiring further study have been identified. The next Chapter will bring together the results of Chapters (IV) to (VI) to draw conclusions and outline future work which has been highlighted by this study.

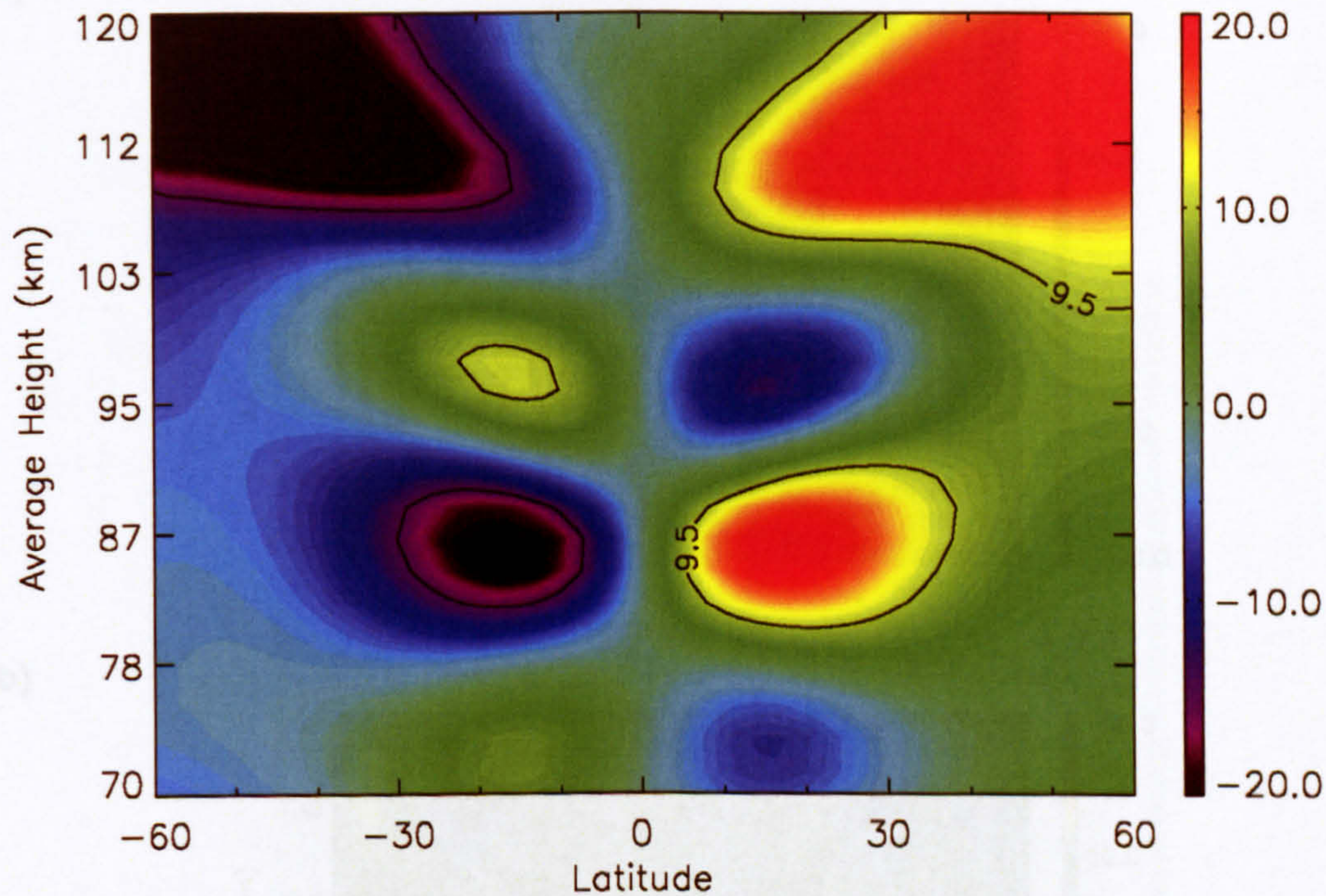


Figure VI.1, Meridional winds (m s^{-1}) at 12 noon calculated by the CMAT model during equinox. Positive values denote northward winds. Data from Harris *et al.* (2002) using the normal lower boundary tidal forcing.

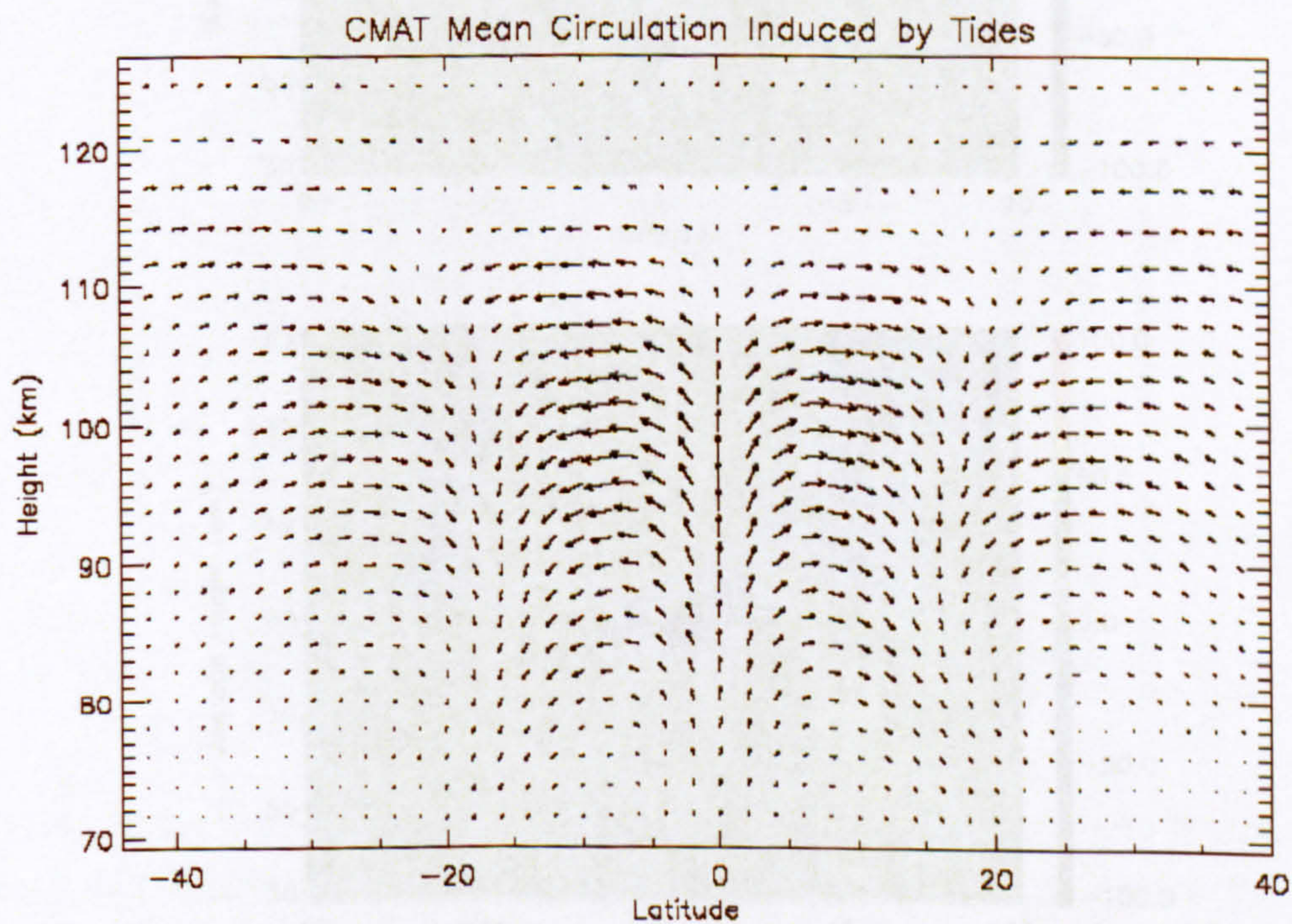
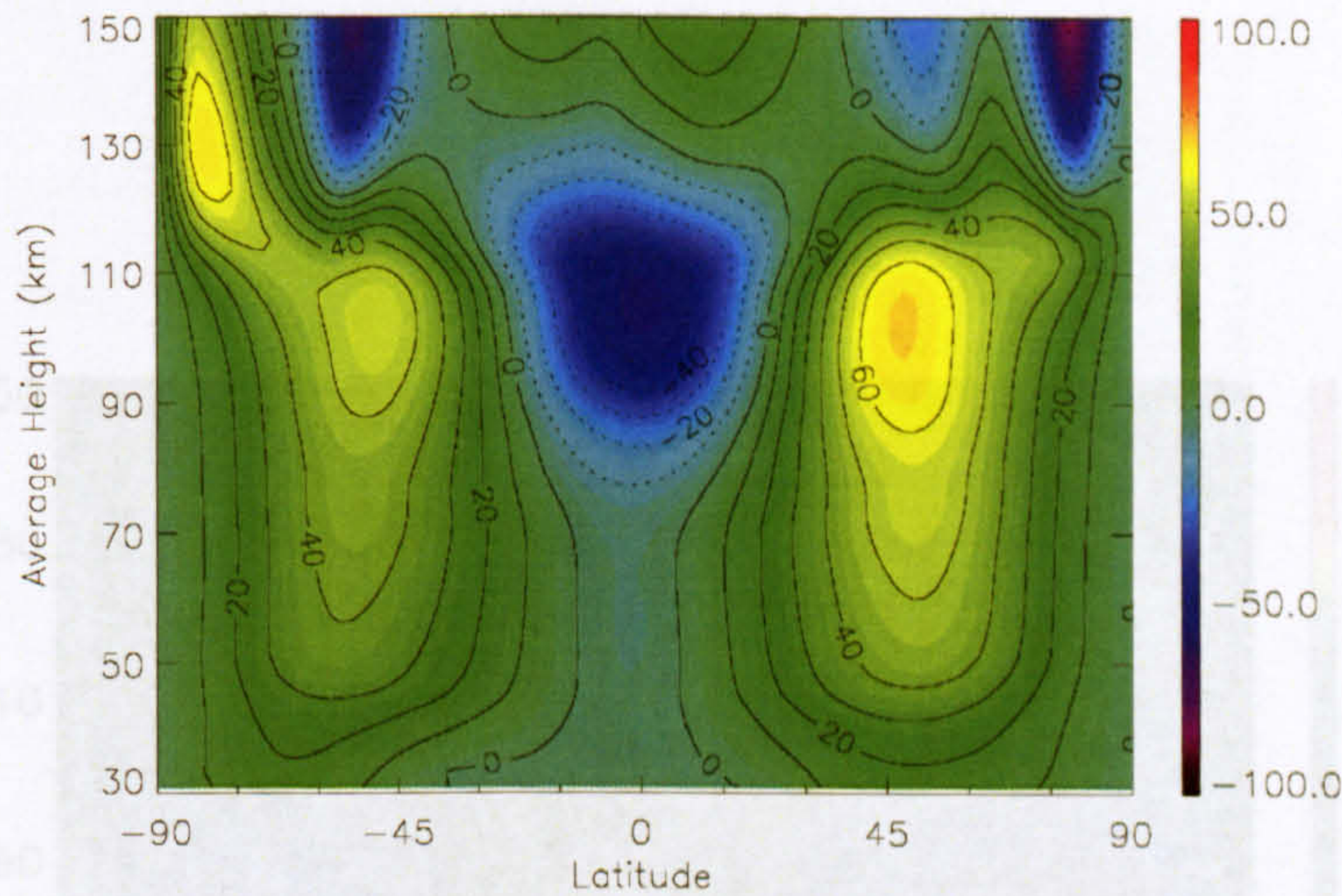
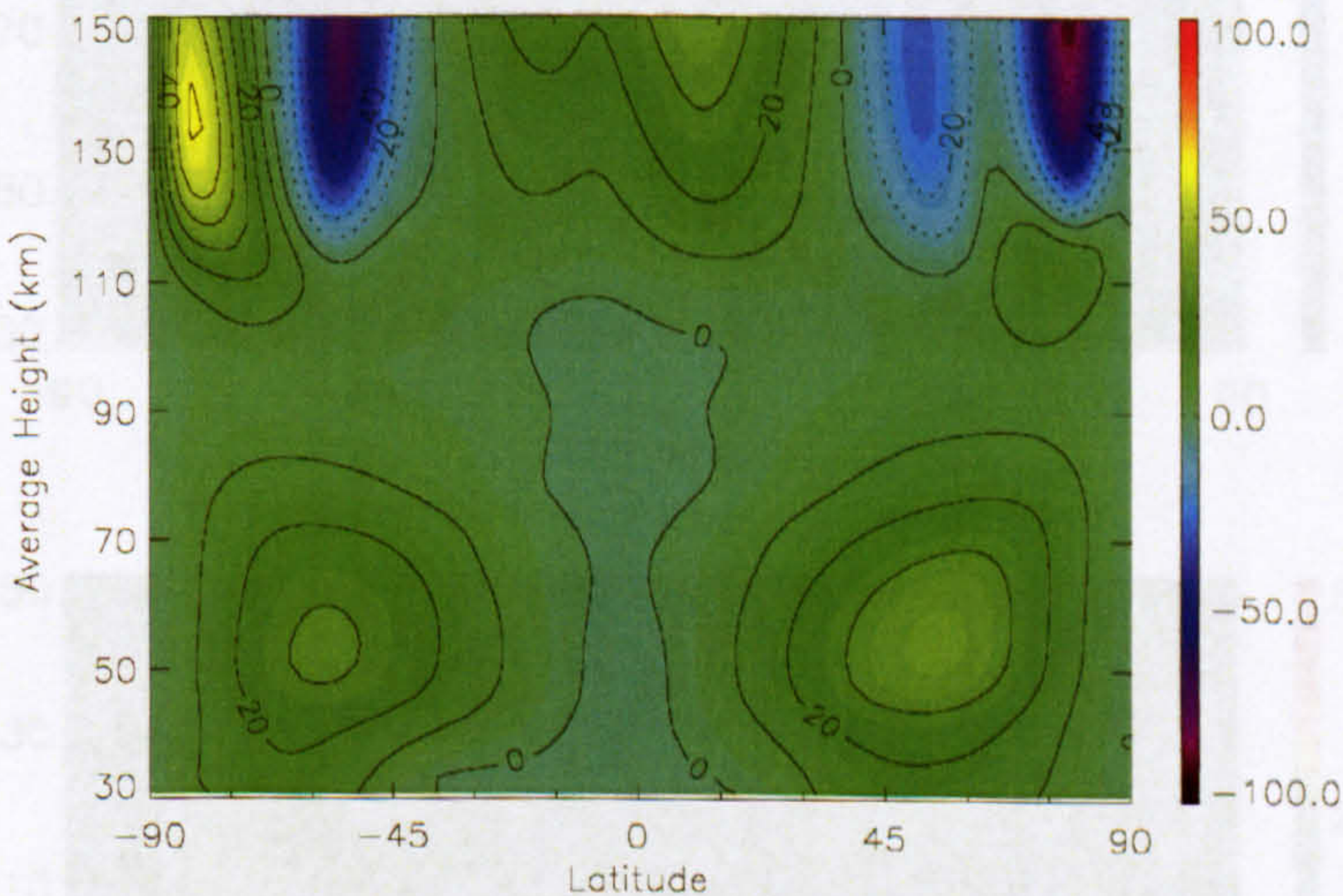


Figure VI.2, Mean meridional circulation induced by the diurnal tide in the CMAT model from Harris *et al.* (2002) using the an increased lower boundary tidal forcing. The vertical winds have been scaled up by a factor of 80 compared with the horizontal winds. The maximum vertical winds are 3 cm s^{-1} and the maximum meridional wind is 7.5 m s^{-1} .

(a)



(b)



(c)

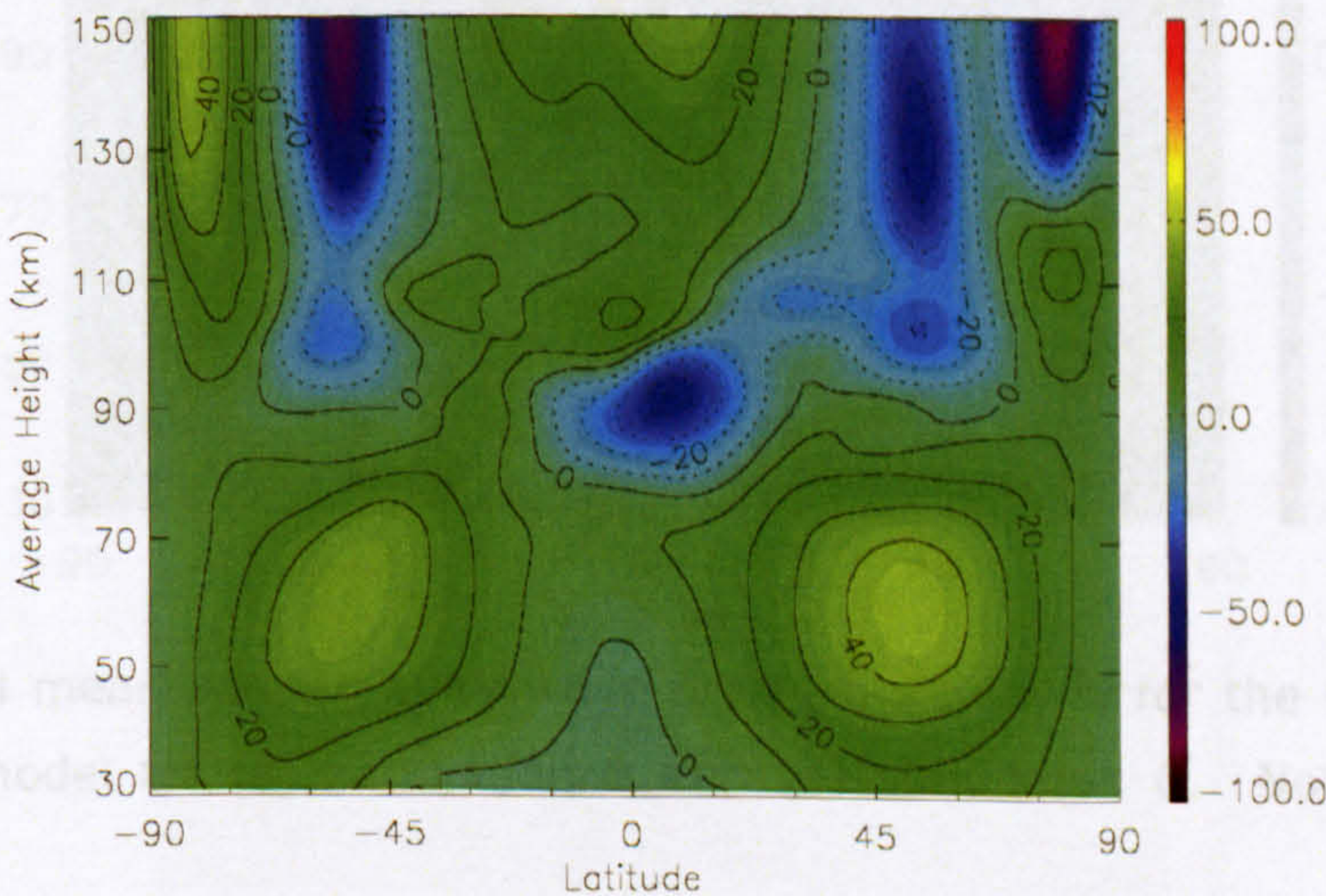


Figure VI.3, Zonal mean zonal winds (m s⁻¹) for the March equinox from CMAT model for (a) simulation A, (b) simulation B and (c) simulation C. Positive values denote eastward winds.

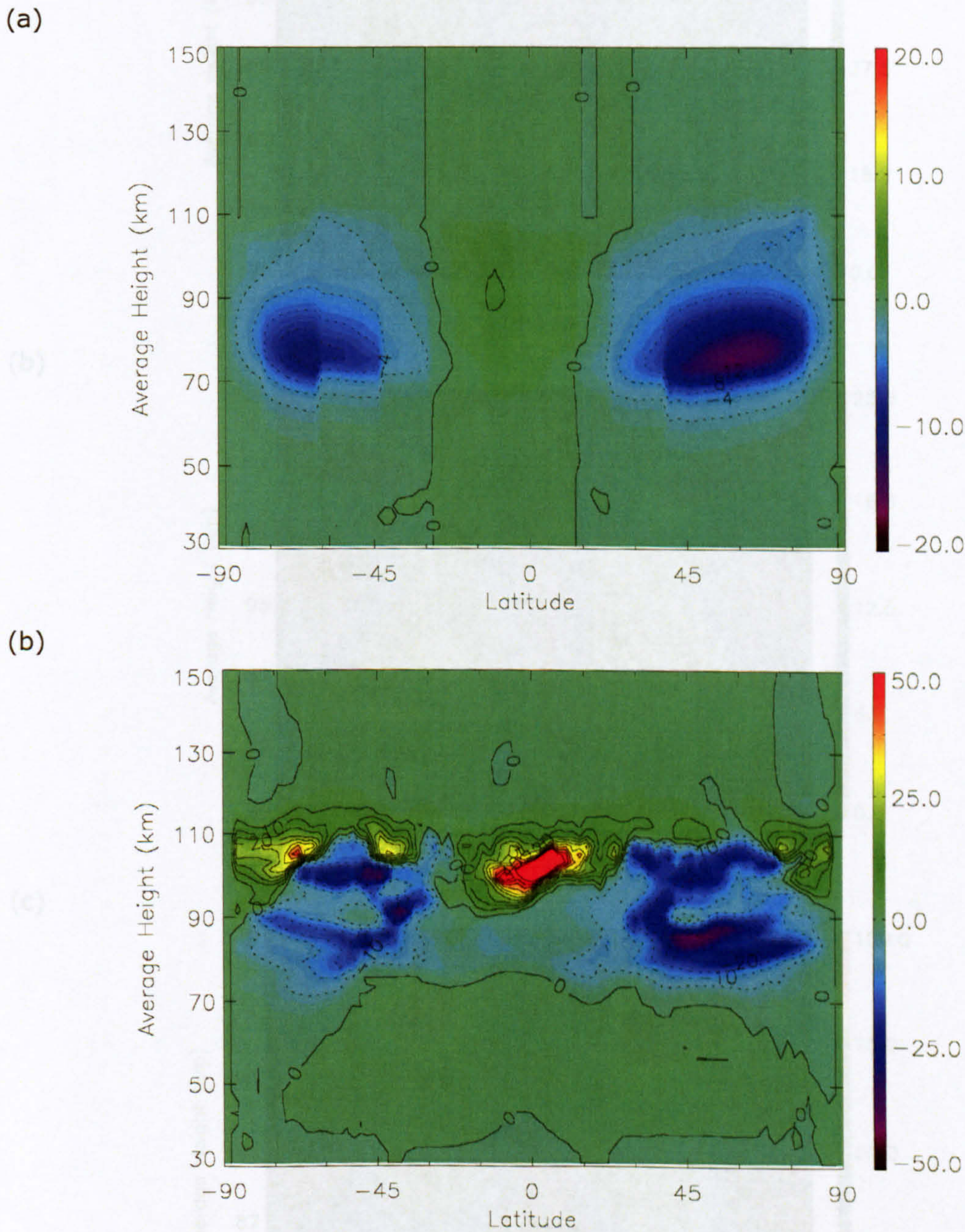
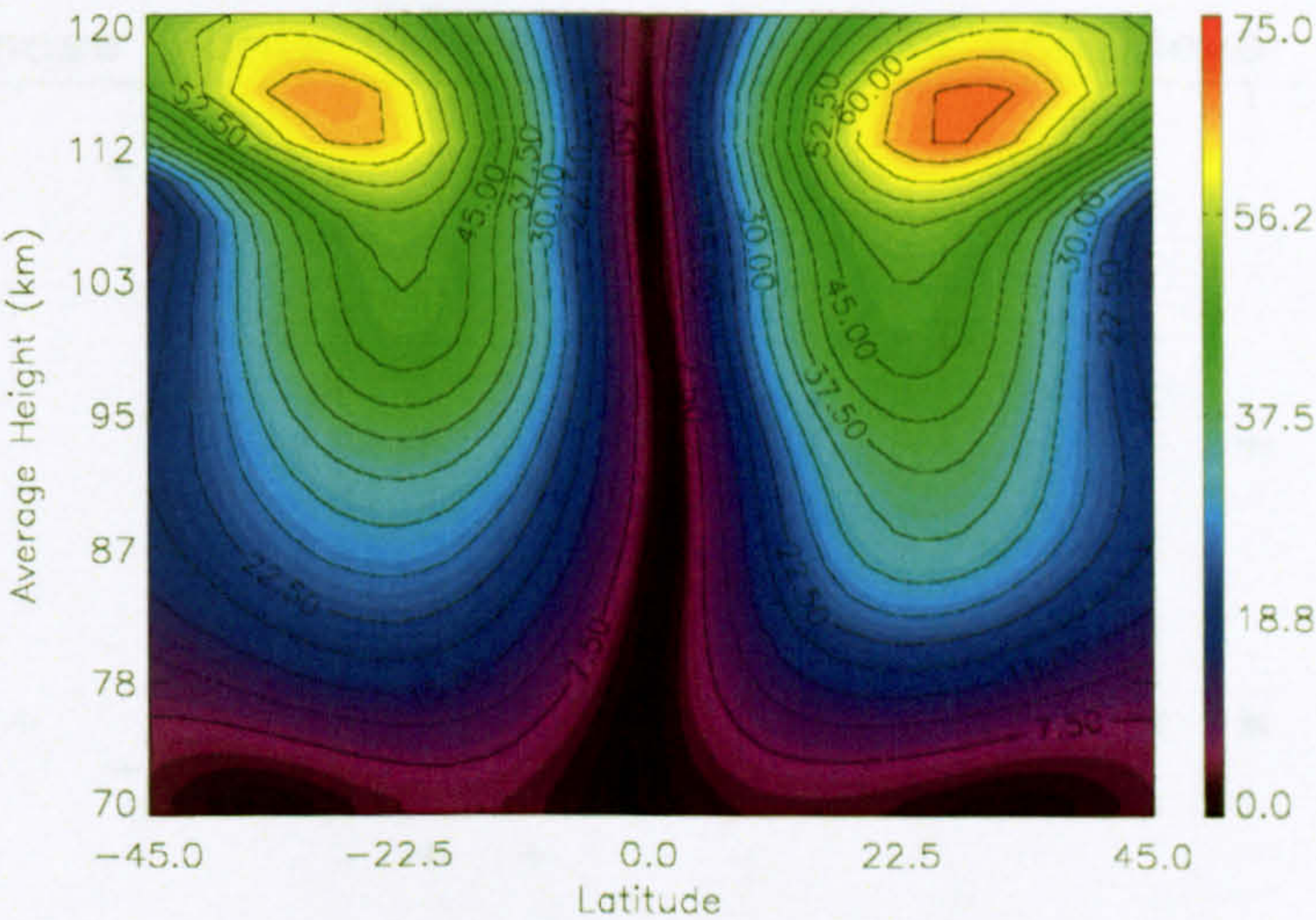
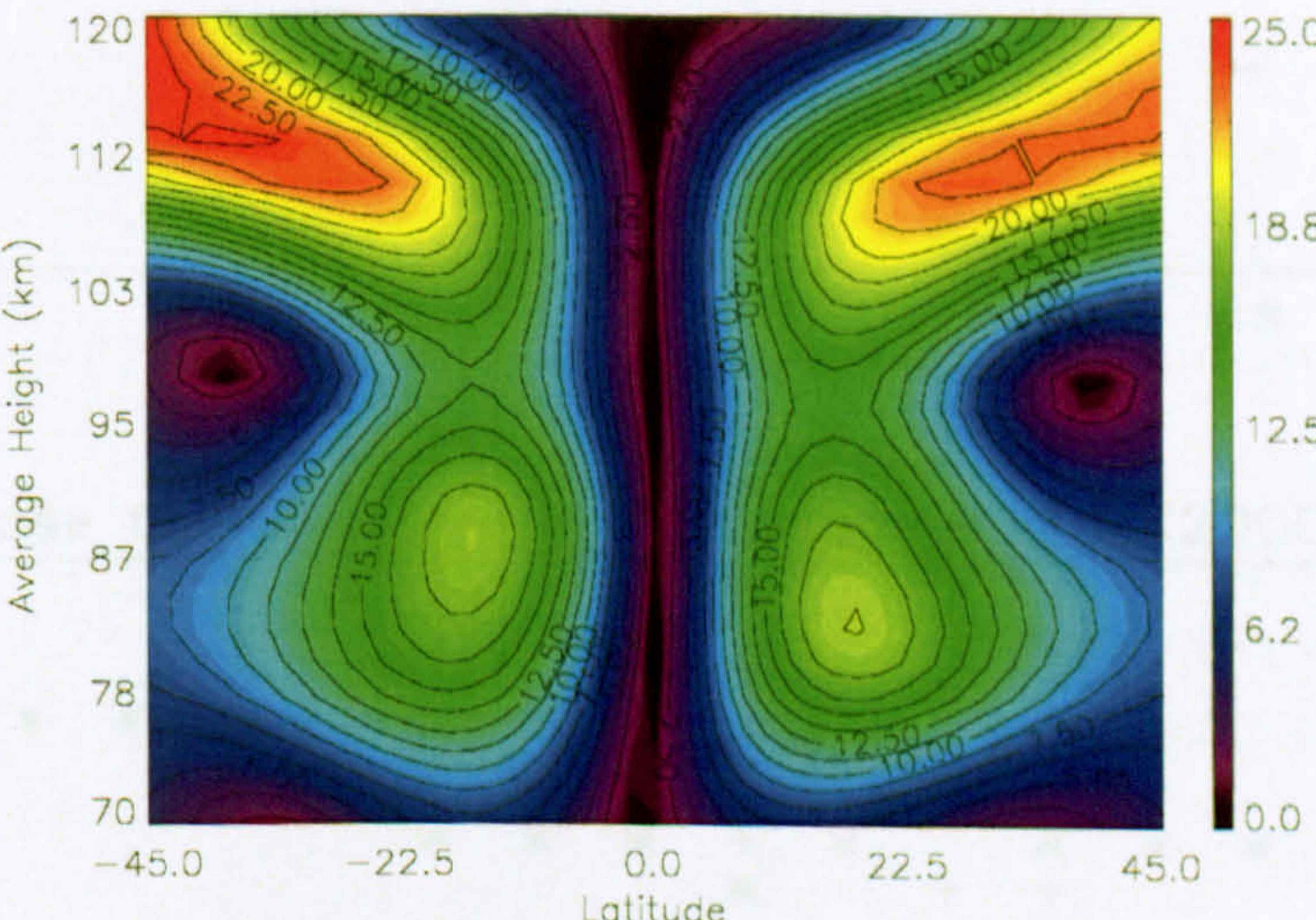


Figure VI.4, Zonal mean zonal gravity-wave drag ($\text{m s}^{-1} \text{ day}^{-1}$) for the March equinox from the CMAT model for (a) simulation B and (b) simulation C. Note the contour intervals used.

(a)



(b)



(c)

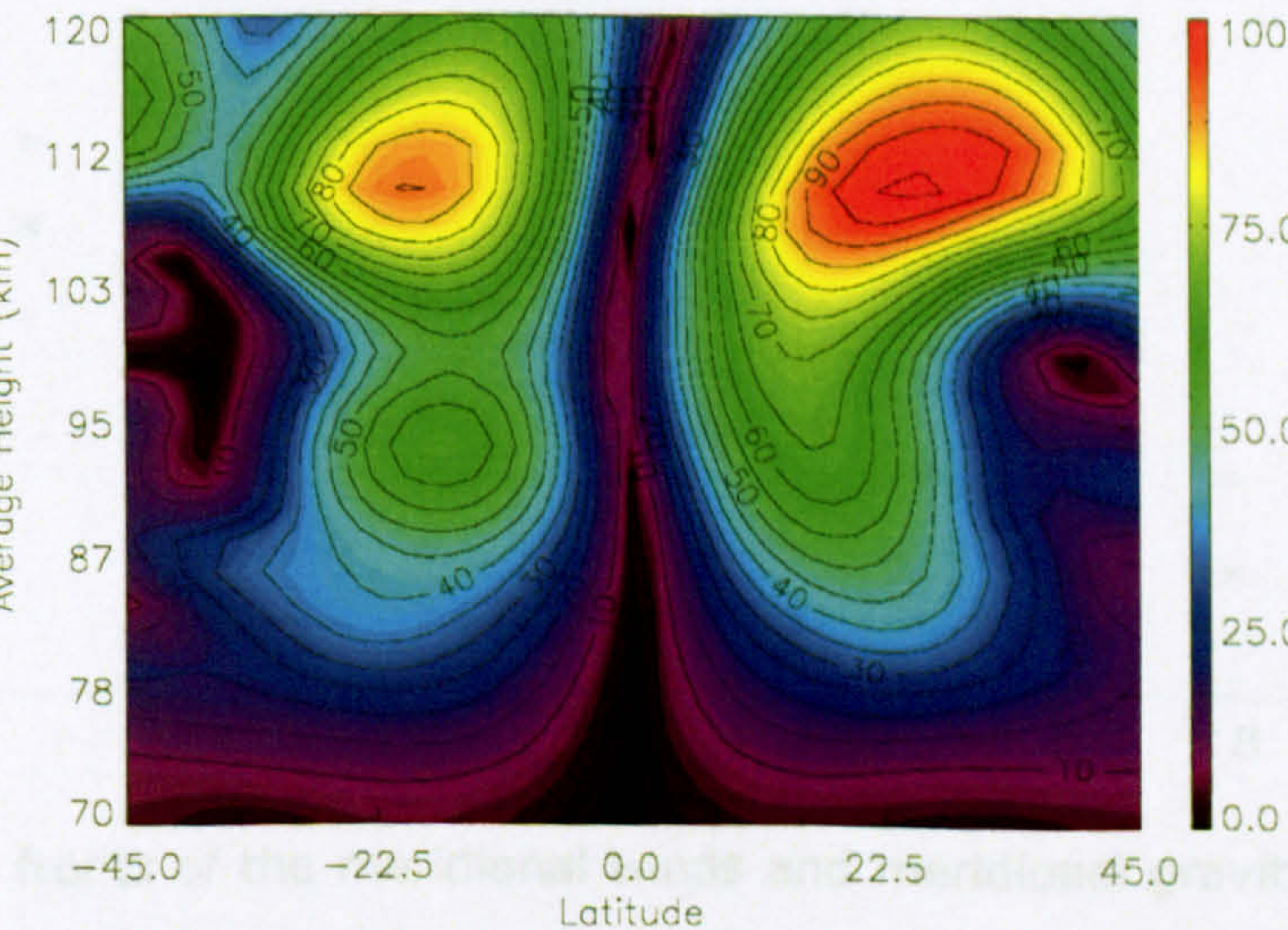


Figure VI.5, Diurnal amplitude (m s^{-1}) of the meridional winds for the March equinox from the CMAT model for (a) simulation A, (b) simulation B and (c) simulation C. Note the different contour intervals used.

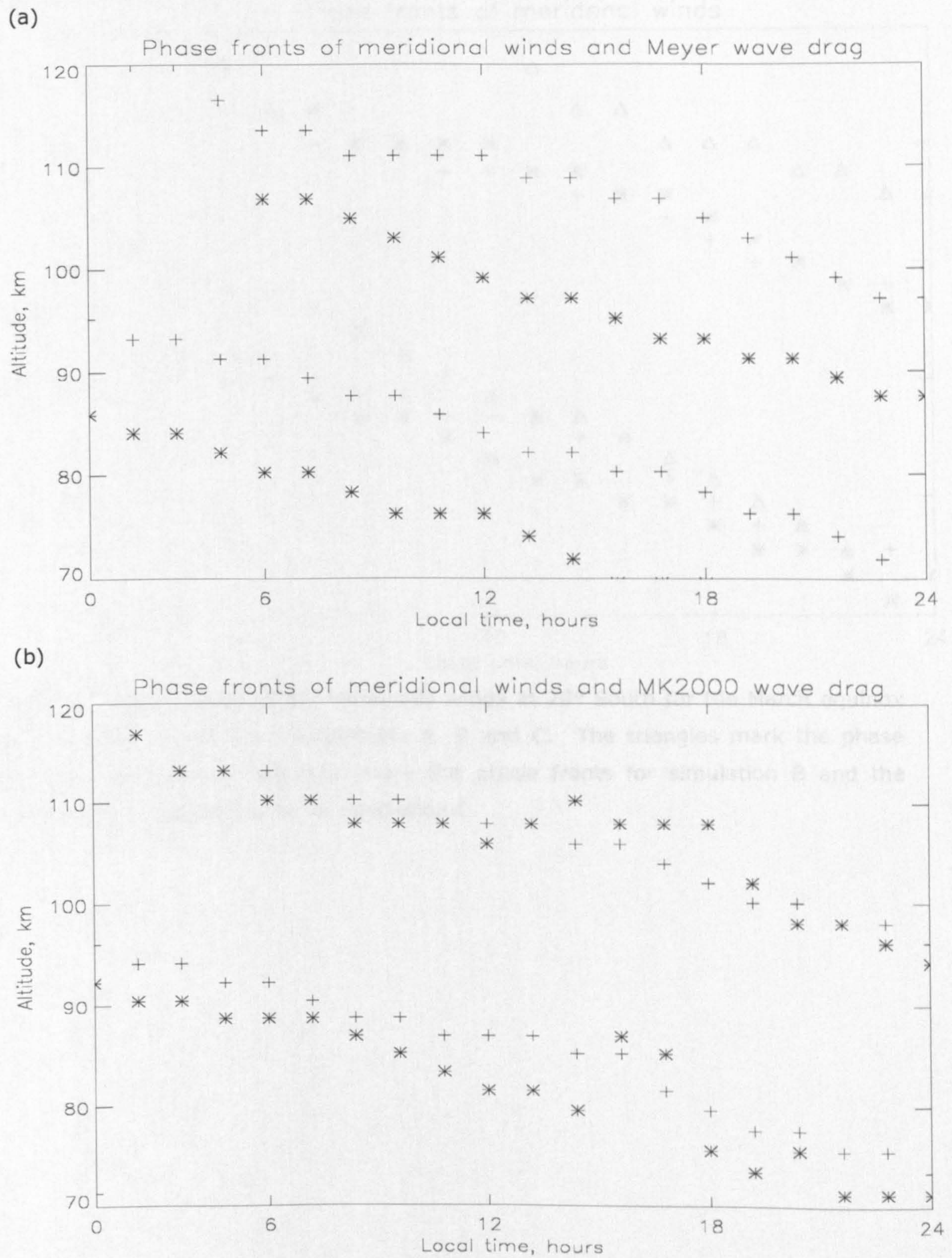


Figure VI.6, Phase fronts of the meridional winds and meridional gravity-wave drag at 20° South for the March equinox from the CMAT model from (a) model simulation B and (b) model simulation C. The crosses mark the phase fronts of the meridional winds and the stars mark the phase fronts of the meridional wave drag.

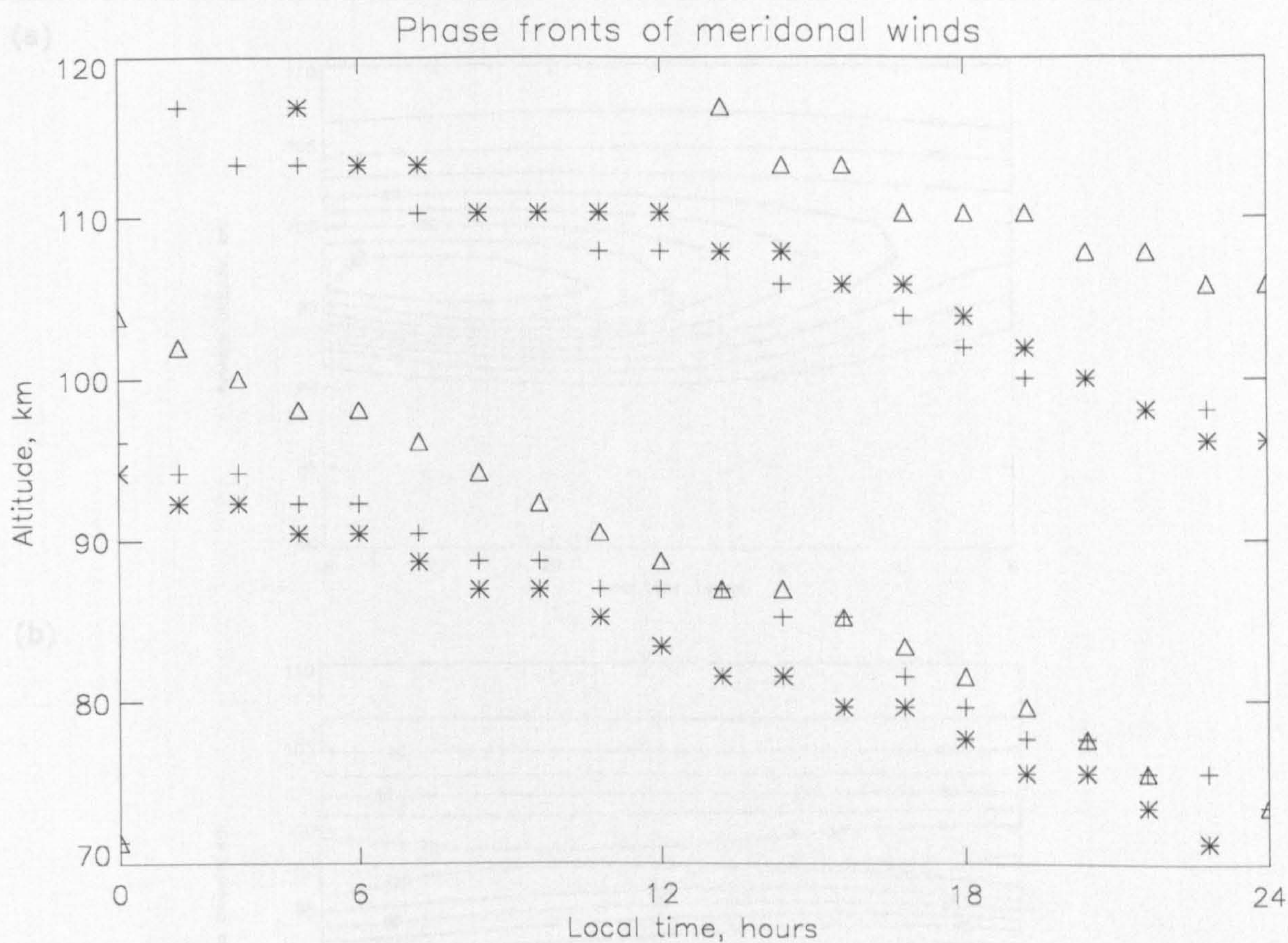
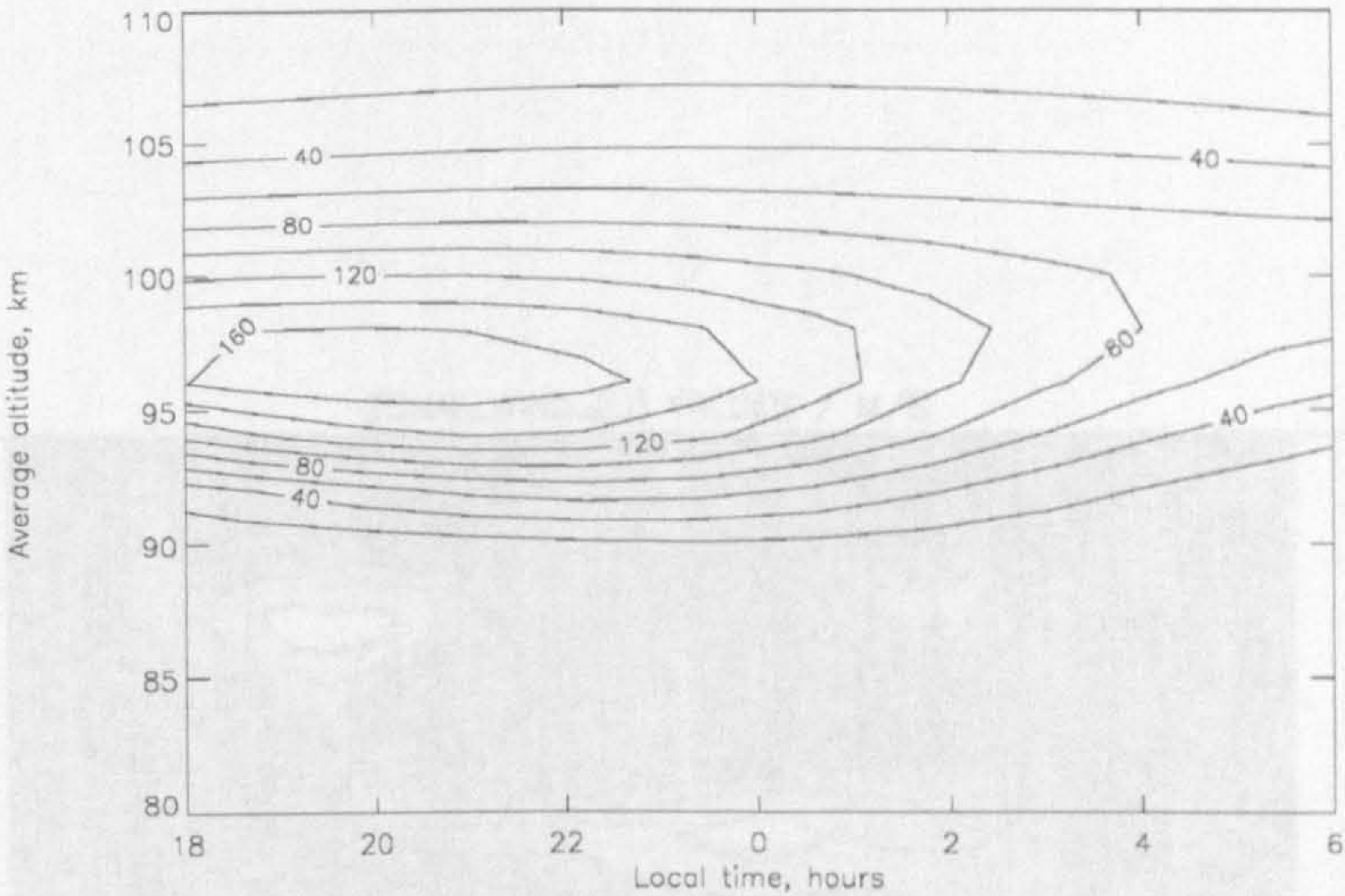


Figure VI.7, Phase fronts of the meridional winds at 20° South for the March equinox from the CMAT model from simulations A, B and C. The triangles mark the phase fronts for simulation A, the stars mark the phase fronts for simulation B and the crosses mark the phase fronts for simulation C.

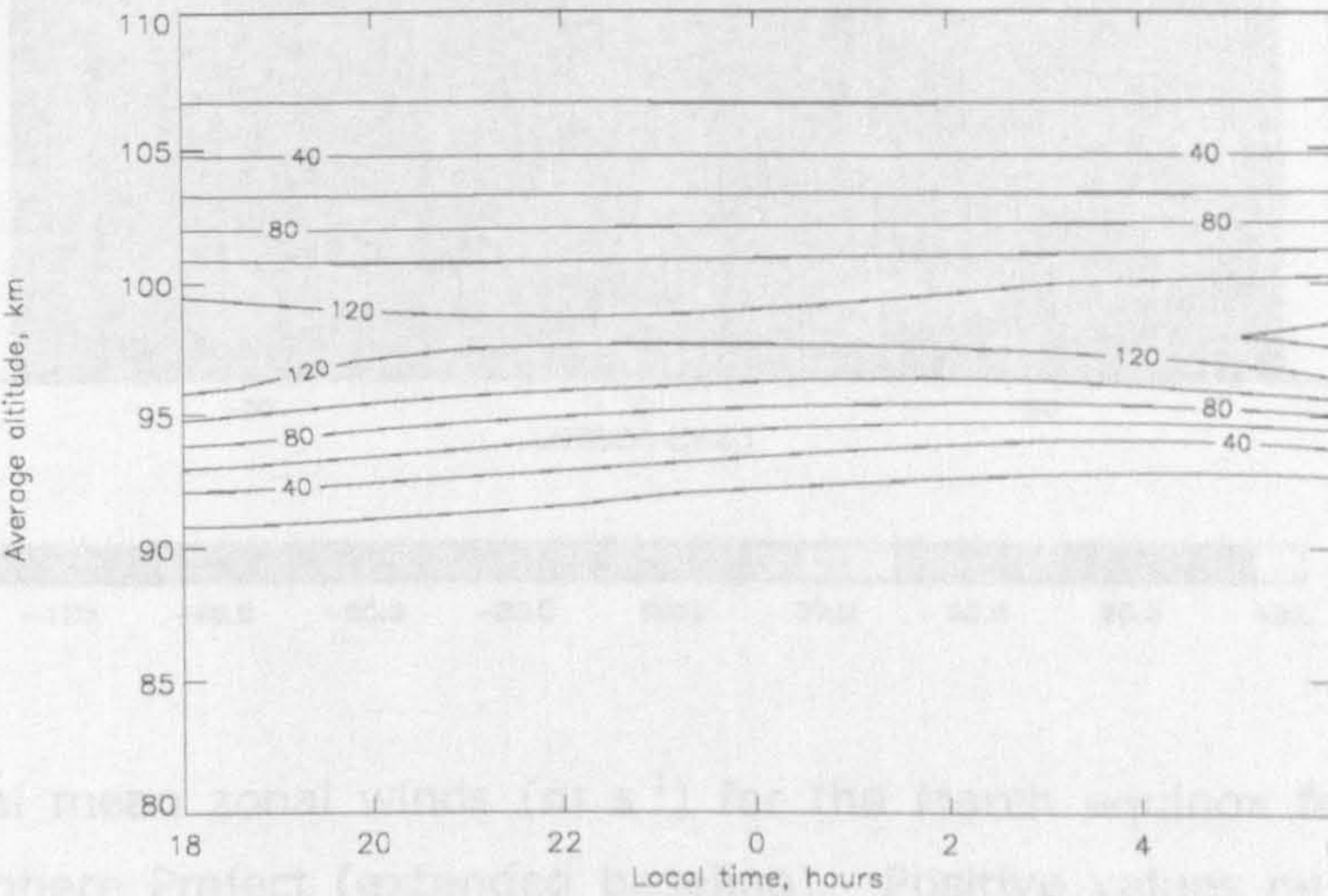


Figure VI.8, Local time variation of the 157.3 km high time-averaged volume emission rate (m^{-3}) calculated by the CMAT model for the March equinox at the equator for (a) model simulation A, (b) model simulation B and (c) model simulation C.

(a)



(b)



(c)

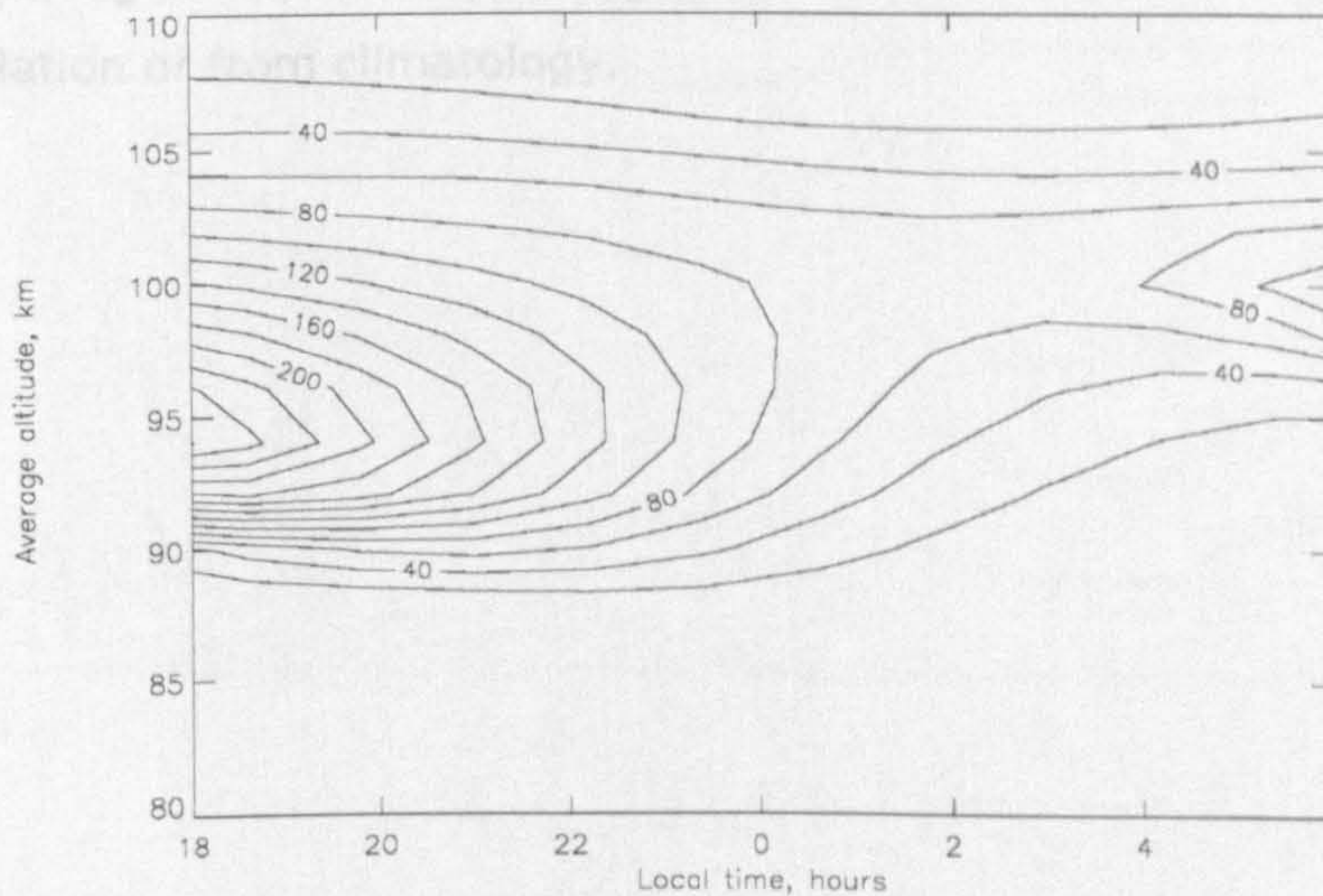


Figure VI.8, Local time variation of the 557.7 nm night time green line volume emission rate ($\text{m}^{-3}\text{s}^{-1}$) calculated by the CMAT model for the March equinox at the equator for (a) model simulation A, (b) model simulation B and (c) model simulation C.

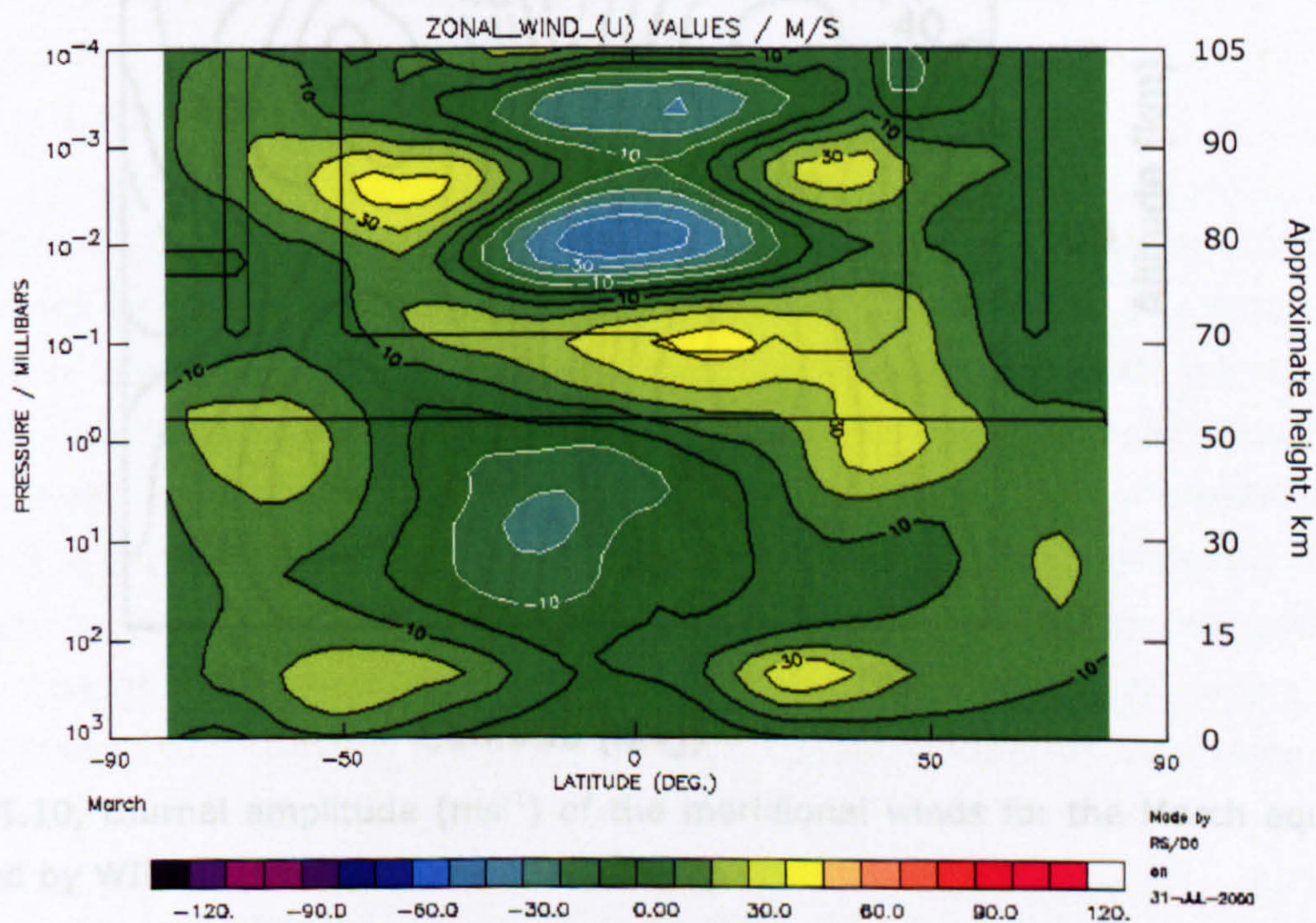


Figure VI.9, Zonal mean zonal winds (m s^{-1}) for the March equinox from the UARS Reference Atmosphere Project (extended baseline). Positive values denote eastward winds. The boxed region in the boxed regions indicate where the data are derived mainly by interpolation or from climatology.

WINDII (March/April 1992/93)

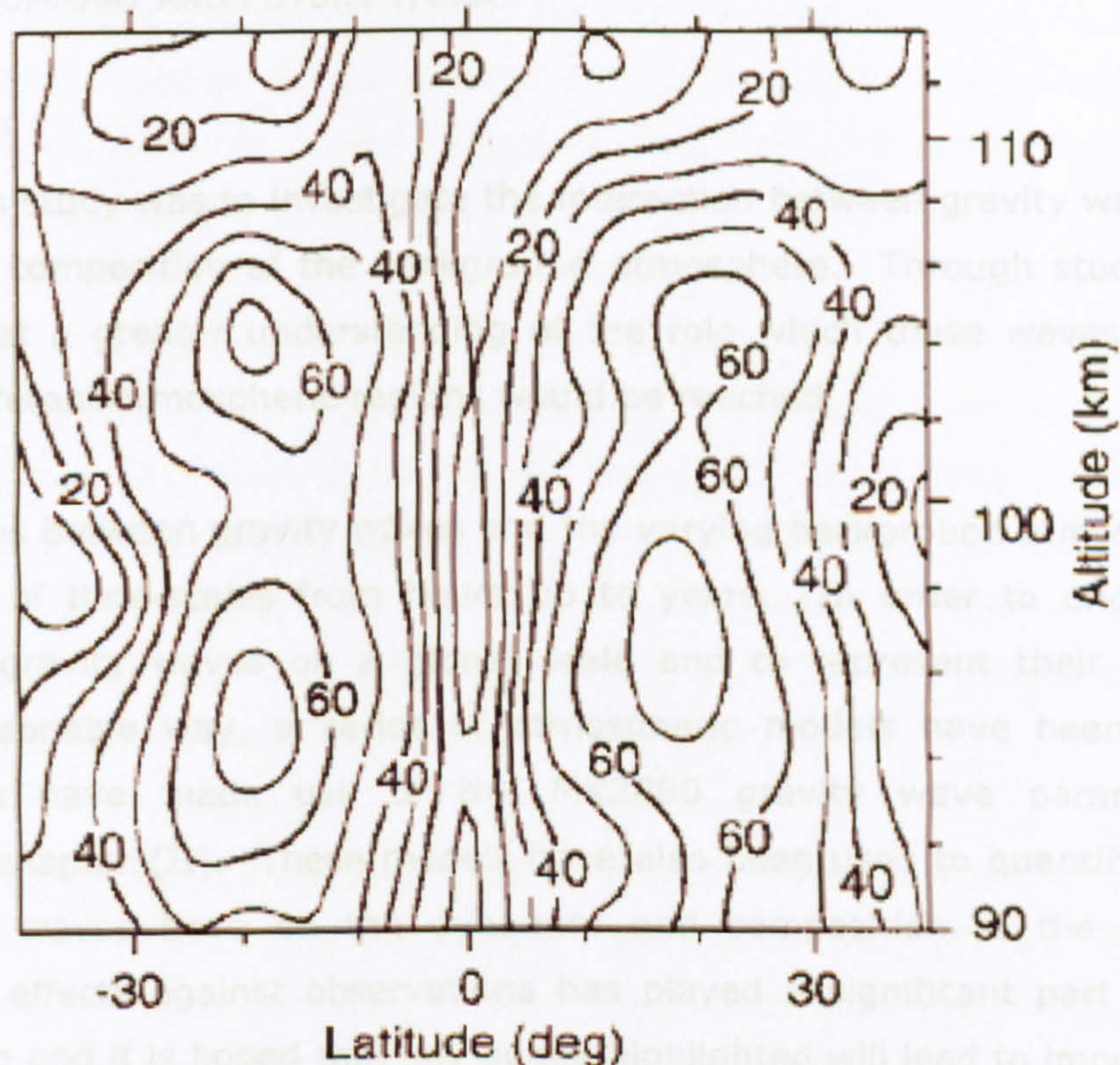


Figure VI.10, Diurnal amplitude (ms^{-1}) of the meridional winds for the March equinox measured by WINDII (from Akmaev *et al.*, 1997).

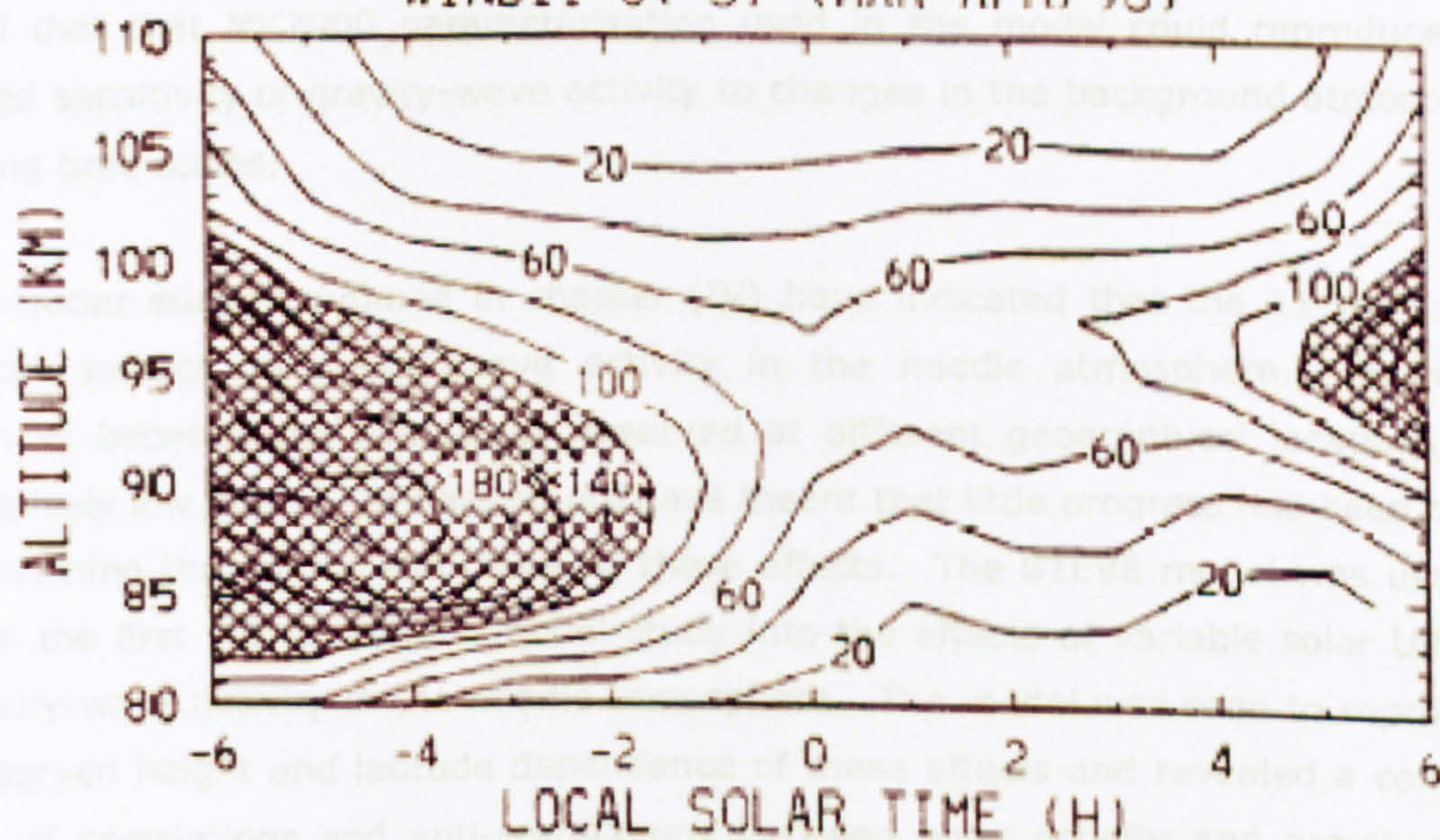
VOLUME EMISSION RATE AT EQ
WINDII 0(1S) (MAR-APR/93)

Figure VI.11 Local time variation of the 557.7 nm night time green line volume emission rate ($\text{s}^{-1}\text{m}^{-3}$) observed by WINDII for March / April 1993 (from Shepherd *et al.*, 1995). The shaded region indicates values over $60 \text{ s}^{-1}\text{m}^{-3}$.

CHAPTER VII SUMMARY AND FUTURE WORK

VII.1 SUMMARY

The aim of this study was to investigate the interaction between gravity waves and the dynamics and composition of the background atmosphere. Through studying this, it was hoped that a greater understanding of the role which these waves play in the coupling of different atmospheric regions would be reached.

The interactions between gravity waves and the varying background atmosphere occur over a range of time-scales from hours up to years. In order to understand the behaviour of gravity waves on a global scale and to represent their effects in a physically reasonable way, a series of atmospheric models have been developed. These models have made use of the MK2000 gravity wave parameterisation, introduced in chapter (II). These models have also been used to quantify the effects which gravity waves have on the dynamics and composition of the atmosphere. Testing these effects against observations has played a significant part in the work presented here and it is hoped that key points highlighted will lead to improvements in the representation of gravity waves in future models.

In chapter (IV), the STEVE model was used to simulate the annual and semi-annual cycles in gravity-wave activity in the upper stratosphere and mesosphere. This showed that that MK2000 parameterisation used in the model could reproduce the observed sensitivity of gravity-wave activity to changes in the background atmosphere over long time scales.

Several radar studies outlined in chapter (IV) have indicated that the 11-year solar cycle can impact on gravity-wave activity in the middle atmosphere. However, differences between the signatures observed at different geographical locations and the relatively low number of data points have meant that little progress has been made in establishing the nature and cause of these effects. The STEVE model was used to perform the first self-consistent model study into the effects of variable solar UV flux on gravity-wave activity in the middle atmosphere. The model was seen to reproduce the observed height and latitude dependence of these effects and revealed a complex picture of correlations and anti-correlations between solar activity and gravity-wave activity, which had been postulated by Laštovička (1999). This picture helps to explain the apparent discrepancies between the observed signatures.

In general, a smaller change in the horizontal zonal velocity associated with changes in solar forcing is seen in the model than in the observations (up to 10 % in the model and up to 100 % reported by Gavrilov *et al.*, 2002). This may be explained by a longitudinal dependence of these effects or the lack of simulated changes in ionospheric and tropospheric conditions.

The variations in gravity-wave activity were shown to affect the eddy diffusion coefficient and gravity-wave drag calculated by the model. Arnold and Robinson (2003) have shown that even small (10 %) changes in the gravity-wave drag in the middle atmosphere may help to significantly amplify the effects of solar variability on the winter polar stratosphere. As changes in the dynamics of this region can affect conditions at the Earth's surface (e.g. Thompson *et al.*, 2002), this result highlights the role which gravity waves play in solar-climate connections.

Outputs from the STEVE model were then used to investigate possible causes of the effects seen in the model. While other factors were seen to contribute, the most likely cause of the variable gravity-wave activity seen in the model were changes in the *in situ* winds in the mesosphere. Further tests also revealed that the nonlinear feedback between the parameterised gravity-wave drag and the *in situ* winds was necessary to create the latitude and height dependence which was seen. This further highlights the need for a realistic representation of gravity waves in global scale models. The model results were also used to provide quantitative evidence that a change in the lower atmosphere gravity-wave source could play an important role in the observed solar signature in gravity-wave activity in the middle atmosphere, as speculated by several previous authors.

The filtering of the gravity wave spectrum by stratospheric winds was also shown to play an important role in producing changes in gravity-wave activity in the mesosphere. This showed the need to study gravity wave interactions with the stratosphere in a more realistic way than was possible in the STEVE model. The STEVE-3D model was developed to simulate planetary-wave activity in the middle atmosphere and to quantify how changes in the middle and upper atmosphere, such as changes gravity-wave drag or solar variability may influence the dynamics of the stratosphere.

Consideration the order of magnitude of the energy and momentum content of the mesosphere and stratosphere may suggest that changes in the mesospheric dynamics should not have a large impact on the stratosphere under normal circumstances. However, modelling studies such as Pawson (1997), Pawson *et al.* (1998) and Arnold

and Robinson (2003) have shown that stratospheric sudden warmings are sensitive to such changes. In chapter (V) the STEVE-3D model was used to study the effects of gravity-wave drag on the winter polar stratosphere around the time of a minor warming. This was the first study to consider the effect of varying the tropospheric gravity-wave source strength over a wide range of values, using a nonlinear ray-tracing gravity wave parameterisation to quantify the impacts which this has on the winter polar stratosphere.

The study in chapter (V) showed that around the time of a sudden warming, variations in the strength of the lower atmosphere gravity-wave source could alter the monthly average temperature in the stratospheric winter polar region by up to 3.5 K. This compares to a change of only 1 K when the same tests were conducted in the absence of a clear stratospheric sudden warming. The impact on the sudden warming was attributed to changes in the preconditioning of the winter stratosphere through both *in situ* stratospheric gravity-wave drag and through changes in the mesospheric circulation. The inclusion of both of these in the model was shown to be crucial in producing the 3.5 K change in stratospheric temperature.

Stratospheric warmings have been shown to be important in understanding the impact of solar variability on climate (e.g. Arnold and Robinson, 1998) as their impacts can be seen throughout the entire atmosphere (e.g. Lauter and Schminder, 1976; Thompson *et al.*, 2002). The study in chapter (V) therefore highlights the importance of constraining the global-scale source of tropospheric gravity waves, their temporal variability, propagation and dissipation throughout the entire atmosphere. One such observational study reported recently by Ratnam *et al.* (2004) has shown that the Southern Hemisphere major warming in September 2002 was accompanied by an increase in the gravity-wave activity potential energy of up to three times around the edge of the polar vortex in the lower and middle stratosphere. Incorporating the effects of such changes into weather prediction models may be important in predicting any future Southern Hemisphere major warming events, as well as for studying the coupling between different regions of the atmosphere.

Having investigated the long-term effects of gravity-wave drag in the mesosphere, chapter (VI) considered the interaction between parameterised gravity-wave drag and the equatorial diurnal tide, which is a dominant feature of the equatorial MLT region. The CMAT model was adapted for this study to investigate the dynamical and compositional effects of the interaction between parameterised gravity-wave drag and the diurnal tide in the equatorial atmosphere between 70 and 120 km during equinox.

It was found that the amplitude, phase, location of the peak amplitude and vertical wavelength of the diurnal tide were altered by the introduction of parameterised gravity-wave drag within the model. These changes were seen to be dependant upon the parameterisation used. Both the Meyer and MK2000 parameterisations were found to advance the phase fronts and reduce the vertical wavelength of the diurnal tide in the MLT region. The Meyer parameterisation was found to significantly reduce the amplitude of the diurnal tide, in agreement with the results of Meyer (1999), whereas the MK2000 parameterisation was found to slightly increase its amplitude, in agreement with the results from Manson *et al.* (2002). These differences were explained in terms of the difference in the altitude of the drags calculated when only wave-mean flow interactions are considered, compared with calculations which include both wave-mean flow and wave-wave interactions. The model was found to produce the most realistic zonal mean zonal winds and diurnal tide when the MK2000 parameterisation was used.

While the MK2000 parameterisation has been seen to provide a good representation of gravity-wave drag throughout the middle atmosphere over long time scales, the enhancement of the tidal amplitude by the MK2000 gravity-wave drag appears to be in contradiction with some observational studies detailed in chapter (VI). This may be a consequence of the lack of representation of intermittency of all, or part of the gravity wave spectrum over the period of a day or because the parameterisation does not account for the propagation time of gravity waves through the middle atmosphere. Both of these simplifications in the parameterisation would be important on short time-scales, as would any general flaw with the nonlinear gravity wave dissipation theory used in this scheme.

The coupling between parameterised gravity waves and the diurnal tide was shown to have a large impact on the $O(^1S)$ green line volume emission rates calculated by the CMAT model. Changing the gravity wave parameterisation was seen to affect the peak value, location of the peak value and the local time dependence of the emission rates. As with the tidal winds, the model produced the most realistic green line emissions when the MK2000 gravity wave parameterisation was used in the model. While it was not addressed in this study, it seems likely that not only the gravity wave parameterisation used, but the gravity-wave source spectrum will impact on the modelled composition of the MLT region and airglow emissions. This may provide a further means of constraining the gravity wave spectra used in global scale models.

The results of this study have shown that gravity waves and the circulation they produce play a vital role in coupling different regions of the atmosphere. Waves

produced in the lower atmosphere have been shown to affect the dynamics, energetics and composition of the middle and upper atmosphere. By altering phenomena in the middle atmosphere such as stratospheric warmings, their vital role in linking changes in the middle and upper atmosphere to the lower atmosphere climate system has been shown. Their representation in global scale numerical models has been investigated and improvements necessary for future work have been highlighted.

VII.2 FUTURE WORK

While this study has provided new insight into several aspects of the coupling of gravity waves and the background middle atmosphere, it has also highlighted several pieces of work which need to be done in order to further understand these processes. This section will detail work which can build upon the results presented here, as well as other results which may be needed to better represent gravity-wave activity in global scale models.

VII.2.1 THE EFFECTS OF SOLAR VARIABILITY ON GRAVITY-WAVE ACTIVITY IN THE MIDDLE ATMOSPHERE

The study in chapter (IV) considered several aspects of the impact of variable solar UV on gravity-wave activity in the mesosphere, but this work can be taken further to address a number of issues which could not be included in this study.

The observations by Laštovička (1999) suggest that there may be a seasonal dependence in the effects of solar variability on gravity-wave activity in the MLT region. This could be investigated further by sorting the observations presented by other authors by season. Given the other discrepancies between the observed effects of solar variability on gravity-wave activity, it seems unlikely that the seasonal dependence at each location would agree, but such a study may resolve whether or not such a dependence exists. The results from the STEVE model did not reveal a seasonal dependence in the zonally averaged change in gravity-wave activity with changes in solar UV, but it is possible that a fully 3D version of the STEVE-3D model might be able to show whether or not such effects are reproduced in the model.

A fully 3D version of the STEVE-3D model would also allow the effects of the stratosphere on the gravity-wave activity in the mesosphere to be included in such a study. Chapter (IV) showed that this may play an important role, but the STEVE model is not able to simulate the conditions in the stratosphere well, primarily due to the lack of planetary-wave activity in the model. This also means it cannot simulate

the observed variations in the stratosphere with variable solar activity (e.g. Labitzke and van Loon, 1988), which may have further impacts on gravity-wave activity in the mesosphere.

Using a modified STEVE-3D model, it may also be possible to resolve the longitudinal dependence in the results discussed in chapter (IV). While this is unlikely to match the observed variations, as no consideration of localised gravity-wave sources is included in the model, such an experiment may provide some predictions on the horizontal scale of the structures seen in figure (IV.19) which can be tested against future observations.

Both Gavrilov *et al.* (2002) and Bošková and Laštovička (2001) showed that the effects of solar variability on gravity-wave activity were dependent on the frequency of gravity waves. The STEVE model uses a spectrum of 15 frequencies of gravity waves, so it is possible to sort the calculated horizontal wind variance by wave frequency to see if this effect is reproducible in the model. Similarly, it is possible to find the vertical flux of horizontal momentum carried by the gravity waves and it may be possible to compare the effects of solar variability on the model calculation of this quantity with that reported by future radar studies.

The coupling between gravity-wave activity and the chemical composition of the atmosphere was explored in chapter (VI) using the CMAT model. Including the effects of chemical changes in a future 11-year model simulation would allow both the changes in ozone, which have been shown to be important in simulating the effects of the solar cycle on the stratosphere (e.g. Haigh, 1996) and any possible feedback between changes in the composition through variable gravity wave driven circulation and diffusion in the mesosphere on gravity-wave activity in that region to be explored.

As discussed in chapter (II), gravity wave dissipation produces localised momentum and energy deposition. The STEVE model simulates only the coupling between variable gravity wave momentum deposition and the background atmosphere, and did not include any representation of the heating produced by gravity-wave activity. Including a self-consistent calculation of the heating due to gravity-wave activity (Medvedev and Klaassen, 2003) would allow any feedback between this and changes in the gravity-wave activity associated with solar variability to be explored.

The analysis of the model in chapter (IV) considered changes in the gravity-wave activity at each constant pressure level used by the model. This was useful in considering the nature of the effects seen in the model and their impact on other

model parameters. However, as noted in chapter (IV), a comparison of the changes in gravity-wave activity at fixed height levels may be more suitable for a comparison with the radar observations. This can not be performed with the current version of the STEVE model as it does not self-consistently calculate variations in the height of each constant pressure surface in the middle atmosphere with variable solar activity. Gavrilov *et al.* (1994) have shown that the variations in pressure on constant height surfaces in the mesosphere is insufficient to account for the observed solar cycle signature in gravity-wave activity in this region, so it is unlikely that the small variation in the height of each pressure level will significantly change the results presented in chapter (IV), but future model work should take these variations into account.

So far, it has been possible to demonstrate that variations in gravity-wave activity in the middle atmosphere with variable solar activity can be simulated using the MK2000 gravity wave parameterisation in the STEVE model. Future studies using STEVE and other atmospheric models, implementing other gravity wave parameterisations, such as the Hines schemes would help to add confidence to this model result and the observations which it sought to replicate.

There is a clear need for further observational studies of gravity-wave activity, on a time-scale of order 11 years, especially at low latitudes and in the Southern Hemisphere. Such observations would allow the global picture of changes in gravity-wave activity with solar activity, predicted in chapter (IV), to be fully tested.

Chapter (IV) considered only the effects of variations in solar UV flux over the 11-year solar sunspot cycle on gravity-wave activity in the middle atmosphere. However, larger variations in solar UV flux are sometimes observed around the 27-day solar rotation period, especially during high solar activity. Bošková and Laštovička (2001) and Eckermann *et al.* (1997) failed to find any link between this and gravity-wave activity in their observations, but it is possible that such a signature may be simulated by atmospheric models. Such discrepancies may help to highlight flaws in the gravity wave parameterisations used in such models.

Finally, only the effects of variations in solar UV have been considered using the STEVE model. The effects of geomagnetic changes on the middle atmosphere have been considered by Arnold and Robinson (2001), who showed that these may have a comparable impact to solar UV changes on the winter polar stratosphere. Gavrilov *et al.* (1995) have also shown that there was a possible correlation between the geomagnetic activity and gravity-wave activity. Using models such as CMAT, which include the effects of geomagnetic activity on the neutral atmosphere and a good

representation of atmospheric gravity waves, the contribution which geomagnetic changes make to variability in the middle and lower atmosphere can be assessed.

VII.2.2 THE EFFECT OF GRAVITY-WAVE DRAG ON THE WINTER POLAR STRATOSPHERE

So far the STEVE-3D model has only been fully validated using the 90 day SSU lower boundary geopotential height data detailed in chapter (III). This has allowed a minor warming to be simulated during January, and the effects of gravity-wave drag on this to be explored. Work is currently ongoing to allow the SMM model, and in the future the STEVE-3D model, to use geopotential height data from the NCEP reanalysis data set (Berg, 2004; private communication). This will allow the effect of gravity-wave drag on a variety of minor and major warmings to be investigated in the future, enabling a more thorough statistical analysis to be performed.

At present the STEVE-3D model uses zonally averaged gravity-wave drag and radiative transfer calculations. By extending these in three dimensions, the interaction between gravity waves and planetary waves can be examined in more detail (similar to the study in chapter VI). This will also provide a more accurate simulation of the impact of mesospheric gravity-wave drag on SSW. Also, by including a diurnal heating signature and by altering the lower boundary to include tropospheric tidal forcing, the interplay between atmospheric tides, gravity waves and planetary waves can be simulated. As described in chapter (VI), the inclusion of atmospheric tides in the model is especially important in simulating the structure and variability of the mesosphere and thermosphere. Further, Pawson *et al.* (1998) noted that including these effects in their model had a significant impact on the mean state and variability of their perpetual January simulation, and this may be important in accurately simulating the preconditioning required for a SSW.

As discussed in chapter (V), orographic gravity waves have been shown to have an effect on the lower stratosphere. The MK2000 parameterisation does not simulate these waves, and so they were not included in the study in chapter (V). Variability of these waves may have a large impact on SSW. The relationship between these effects and the effect of non-orographic gravity waves could be explored by including an orographic gravity wave scheme in the STEVE-3D or another model.

Arnold and Robinson (2001) have shown that variable geomagnetic activity can affect SSW and Arnold and Robinson (1998, 2000, 2003) have shown that variable solar UV flux can affect SSW. In a more general study, the effect of the upper atmosphere on SSW could be simulated in a model which includes the effects of gravity waves,

planetary waves, tides, Kelvin waves, interactive chemistry, variable solar UV and ion-neutral collisions. Further development of the CMAT model to create a new model (CMAT2), which may simulate all of these effects is currently underway (Harris, 2004; private communication). In the future it may also be possible to include the effect of the QBO on SSW (e.g. Holton, 1994) to explore how all of these factors may influence SSW and therefore the tropospheric climate system.

VII.2.3 THE EFFECT OF GRAVITY WAVE / TIDAL COUPLING

In chapter (VI), the Meyer gravity wave parameterisation, as implemented by Harris *et al.* (2002), was compared against the MK2000 parameterisation, implemented in a similar way to Medvedev *et al.* (1997). These were chosen as they had both been tested successfully in the model and shown to be reasonable by other authors. However, this means that the way in which the model interacts with these two parameterisations is different. The Meyer scheme accounts for the filtering of the gravity wave spectrum below 10 mb (the lower boundary of the model) by eliminating any waves of horizontal phase speed (Eulerian) which were equal to the wind speeds (to account for critical-level absorption) encountered between the tropopause (taken as the source region) and this level using wind data from the HWM (which may not necessarily match with CMAT winds at the 10 mb pressure level). The implementation of the MK2000 scheme used in chapter (VI) made no such attempt to account for any wave filtering below the CMAT lower boundary, so an isotropic spectrum is used as the gravity-wave source (as with the STEVE and STEVE-3D models).

Medvedev *et al.* (1998) demonstrated that using an anisotropic gravity-wave source with the MK2000 parameterisation improved the zonal mean wind structure in the CMAM model, so it is important to consider this in the CMAT model as well. Attempts to introduce filtering of the gravity wave spectrum, using HWM horizontal wind data between the tropopause and CMAT's lower boundary, in a manner which would make the implementation of the MK2000 scheme more consistent with that of the Meyer scheme and that used by Medvedev *et al.* (1998) is currently underway (Dobbin, 2004; private communication). The initial results show that while this does have a small impact on the zonal mean wind structure, the interaction between the diurnal tide and the parameterised gravity-wave drag remains the same as that reported in chapter (VI), although further testing and validation of the model is required before these results could be used.

Harris *et al.* (2002) argued that the vertical eddy diffusion coefficients used in the CMAT model were too large, and that this led to a depletion of atomic oxygen in the

airglow region. In chapter (VI), the deficiency in the $O(^1S)$ airglow reported by Harris *et al.* (2002) was shown to be solved by improving the solar spectrum used in the CMAT model. However, this does not necessarily mean that the global mean eddy diffusion profile used in the model is satisfactory. The MK2000 gravity wave parameterisation calculates a self-consistent eddy diffusion coefficient. A global mean profile of this has been implemented in the CMAT model, and shown to have some small effect on the tidal amplitude and $O(^1S)$ airglow emissions (Dobbin, 2004; private communication). While this is more self-consistent, it was not used in chapter (VI) as it would have made the cause of any differences in the coupling between the two parameterisations and the diurnal tide more ambiguous.

As the gravity-wave drag is calculated at every grid point in the middle atmosphere, a more ideal cause would consider the eddy diffusion coefficient calculated at each grid point, rather than a global mean profile. This may have a larger impact on the diurnal tide as local time variations in the eddy diffusion (associated with local time variations in the gravity wave dissipation) may act to alter the damping of the tidal amplitude. At present, this is not possible to test with the CMAT model, but it is hoped that modifications to this, or future models may allow this test to be conducted.

In chapter (VI), the coupling between the diurnal tide and gravity-wave drag during the March equinox in the equatorial MLT region was explored. An obvious further step would be to study the coupling between these two during other months, when the diurnal tidal amplitude and the zonal mean wind profile is different. This has not been considered in this study as the CMAT model has been shown to be numerically unstable around solstice conditions (Dobbin, 2003; private communication). It is hoped that the CMAT2 model can be used to further this study.

In the extra-tropical MLT region, the semi-diurnal tide dominates over the diurnal component. The interaction between this tide and parameterised gravity waves may reveal further discrepancies between the parameterisations used in global scale models and observations. At present, this cannot be satisfactorily tested in the CMAT model as it does not represent the semi-diurnal tide well. Tests with the CMAT2 model have confirmed that this is because the ozone rich excitation region in the model is shorter than the wavelength of the semi-diurnal tide. By lowering the level of the lower boundary of the model from that used in CMAT from ~ 30 km to ~ 10 km, the CMAT2 model has been able to generate a more realistic semi-diurnal tide (Harris 2004; private communication). It is hoped that further development of this new model will enable the interaction between this tide and parameterised gravity waves to be studied.

The lack of time-variation in the gravity-wave source is one possible explanation for the discrepancies between the interaction between the gravity-wave drag calculated by the MK2000 scheme and the diurnal tide with that observed. Including a time-varying source spectrum in the model may help to improve this. However, in the absence of a observed diurnal variation in the global tropospheric gravity-wave source, and given the computational expense associated with running the CMAT model, it may prove difficult to include a time varying source within the model against observations made at a variety of altitudes. Perhaps an appropriate intermediate step would be to study the day to day variation in gravity-wave activity and include this in the model, to see to what extent the day to day variability in the observed amplitude of the tide can be explained by gravity wave interactions.

The study in chapter (VI) showed that there is an apparent discrepancy between the interaction of the gravity-wave drag calculated by the MK2000 scheme and the diurnal tide and that reported by observations. It is hoped that observations of gravity-wave activity on a global scale (e.g. Ern *et al.*, 2004), along with specific case studies which focus on specific regions using networks of instruments (e.g. Ratnam *et al.*, 2004), along with other observed parameters such as horizontal winds and airglow emission rates and further theoretical work can be used to refine both the gravity-wave sources which are used in global scale models and the gravity wave parameterisations. This will allow a better understanding of the role which gravity waves play in determining the structure and variability of the atmosphere over a wide variety of time-scales.

APPENDIX I, DOWNWARD CONTROL THEORY

Haynes *et al.* (1991) provided a rigorous proof of the principle of downward control (McIntyre, 1987) for a steady state atmosphere. This principle describes the long term effect of an eddy-induced zonal mean force (such as gravity wave drag) acting on the extratropical middle atmosphere. This appendix will briefly outline this principle and its importance in middle atmosphere dynamics.

Starting from the steady state, transformed Eulerian Mean equations of motion, the zonal mean momentum is written in the form,

$$\frac{\partial \bar{u}}{\partial t} + \bar{v}^* \left[\frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} (\bar{u} \cos \phi) - 2\Omega \sin \phi \right] + \bar{w}^* \frac{\partial \bar{u}}{\partial z} = \mathbf{F}, \quad \text{A.1}$$

and the mass continuity equation,

$$\frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} (\bar{v}^* \cos \phi) + \frac{1}{\rho_0} \frac{\partial}{\partial z} (\rho_0 \bar{w}^*) = 0, \quad \text{A.2}$$

Using the stream function of the residual circulation, $\Psi(\phi, z)$, defined by,

$$\bar{v}^* = -\frac{1}{\rho_0 \cos \phi} \frac{\partial \Psi}{\partial z}, \quad \text{A.3}$$

and,

$$\bar{w}^* = \frac{1}{a \rho_0 \cos \phi} \frac{\partial \Psi}{\partial \phi}, \quad \text{A.4}$$

the zonal mean momentum equation can be rewritten as,

$$\frac{\partial(\Psi, \bar{m})}{\partial(\phi, z)} = \frac{\partial \Psi}{\partial \phi} \frac{\partial \bar{m}}{\partial z} - \frac{\partial \Psi}{\partial z} \frac{\partial \bar{m}}{\partial \phi} = \rho_0 a^2 \mathbf{F} \cos^2 \phi, \quad \text{A.5}$$

where \bar{m} is the angular momentum per unit mass.

The principle of downward control comes from solving equation (A.5) to find Ψ , and so finding the residual circulation produced by a zonal mean force in the middle atmosphere. Numerical solutions to the TEM equations and observations of tracer transport had already shown that the residual circulation in the middle atmosphere had the form shown in figure (I.11), but Haynes *et al.* (1991) demonstrated that, for any given realistic atmosphere, the circulation would always take this form. They showed that by constraining $\rho_0 \bar{w}^* \rightarrow 0$ and therefore $\Psi \rightarrow 0$ as $z \rightarrow \infty$, equation (A.5) could be integrated with respect to z , to give,

$$\Psi(\phi, z) = \int_z^\infty \left\{ \frac{\rho_0 a^2 \mathbf{F} \cos^2 \phi}{\bar{m}_\phi} \right\}_{\phi=\phi(z')} dz', \quad \text{A.6}$$

and the corresponding equation,

$$\overline{w^*} = \frac{1}{\rho_0 \cos \phi} \frac{\partial}{\partial \phi} \left[\int_z^\infty \left\{ \frac{\rho_0 a F \cos^2 \phi}{\overline{m}_\phi} \right\}_{\phi=\phi(z')} dz' \right] \quad \text{A.7.}$$

From equations (A.6 and A.7) it can be seen that the meridional circulation at any given height is controlled exclusively by the eddy forcing F at that location and above it; the eddy forcing below that level plays no part. This means that, for any realistic atmosphere, the streamlines of the long-term residual circulation in the middle atmosphere close via the lower atmosphere, and not via the thermosphere (figure I.11).

Haynes *et al.* (1991) showed that, in the transient case, the stream function can extend above the level of the eddy forcing, but these streamlines decay to zero over a short time period. This is because the upper atmosphere has less ability to 'take up' the angular momentum while still remaining in thermal wind balance than the lower atmosphere.

Haynes *et al.* (1991) went on to show that, in the case of gravity wave drag in the mesosphere, the eddy forcing term can be simplified to,

$$F = \frac{1}{\rho_0} \frac{\partial (\rho_0 \overline{u'w'})}{\partial z}, \quad \text{A.8}$$

and using the quasi-geostrophic approximation and assuming that $\rho_0 \overline{u'w'} \rightarrow 0$ as $z \rightarrow \infty$, equation (A.7) could be rewritten as,

$$\overline{w^*} = -\frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} \left\{ \frac{\overline{u'w'} \cos \phi}{2\Omega \sin \phi} \right\}. \quad \text{A.9}$$

This shows that, in this case, $\overline{w^*}$ at any height level is controlled by the gravity wave momentum flux through that level and its latitudinal derivative and that $\overline{w^*}$ is independent of the height at which the wave dissipation occurs. Further, McIntyre (1992) argued that even in the general case, the detailed structure of F is not a dominant factor in controlling the residual circulation in the long term.

It is worth noting that the theory of downward control applies only to the extratropical atmosphere, where the contours of constant \overline{m} are approximately vertical. In the equatorial regions, this control is not locally important. However, through the global-scale residual circulation, the equatorial regions are affected by extratropical forcing.

From a consideration of the typical height distribution of \mathbf{F} and ρ_0 , McIntyre (1987) argued that this downward control principle means that the residual circulation at any given location is dominated by the eddy forcing in the two to three scale heights above that level.

APPENDIX II, CAUSALITY IN MIDDLE ATMOSPHERE DYNAMICS

Chapman (1930) proposed the first mechanism to explain the peak in the ozone mixing ratio in the stratosphere (the ozone layer). The Chapman mechanism describes the photochemical production and loss of ozone in the middle atmosphere. The appendix will outline the photochemical reactions involved in this mechanism, in the form set out by Brasseur and Solomon (1984).

The first step in the Chapman mechanism involves the photolysis of O_2 , given

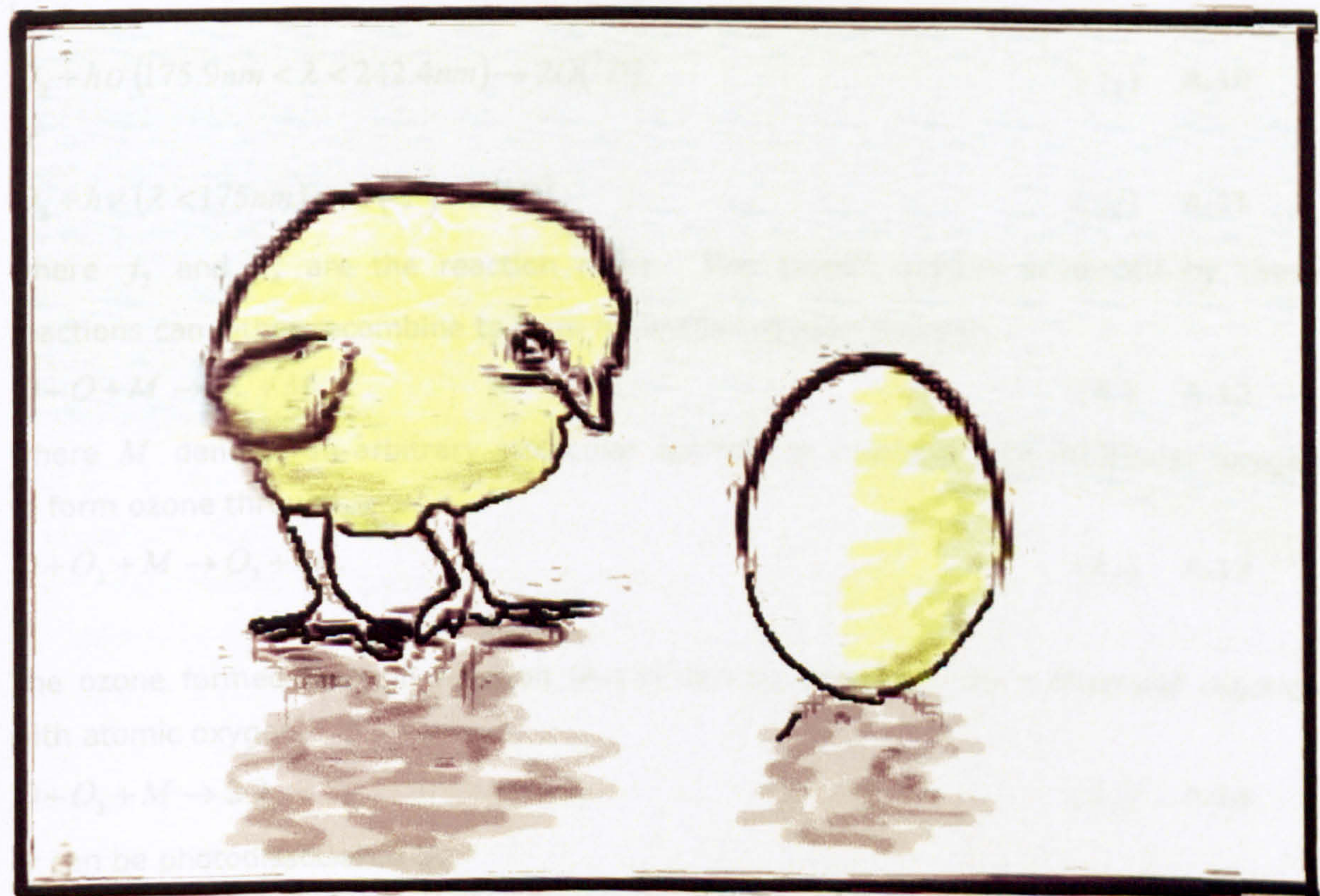


Figure A.1, Chicken and egg.



The net effect of reactions (A.10), (A.11) and (A.12) is to produce ozone from molecular oxygen and atomic oxygen:



and the net effect of reactions (A.10), (A.11) and (A.12) is to produce atomic oxygen from molecular oxygen:



which represents the destructive step of the Chapman mechanism.

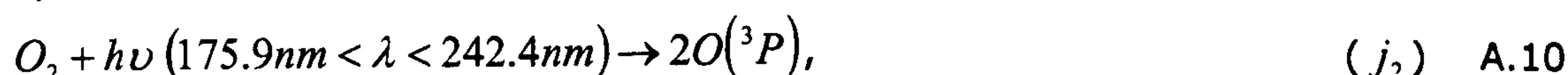
By considering the rate coefficients for reactions (A.10), (A.11) and (A.12) it can be shown that the net effect of reactions (A.10), (A.11) and (A.12) is to produce ozone from molecular oxygen and atomic oxygen:

at 30 km altitude can be found in Table A.1.

APPENDIX III, CHAPMAN CHEMISTRY

Chapman (1930) proposed the first mechanism to explain the peak in the ozone mixing ratio in the stratosphere (the ozone layer). The Chapman mechanism describes the photochemical production and loss of ozone in the middle atmosphere. This appendix will outline the photochemical reactions involved in this mechanism, in the form set out by Brasseur and Solomon (1984).

The first step in the Chapman mechanism involves the photodissociation of O_2 , given by,



or



where j_2 and j_4 are the reaction rates. The atomic oxygen produced by these reactions can either recombine to form molecular oxygen through,



where M denotes an arbitrary molecular species, or combine with molecular oxygen to form ozone through,



The ozone formed through reaction (A.13) can be destroyed by a chemical reaction with atomic oxygen,



or can be photodissociated by,



or



The net effect of reactions (A.10, A.11 and A.13) is the production of ozone,



and the net effect of reactions (A.14 – A.16) is,



which represents the destruction of ozone.

By considering the rate coefficients in the above equations, the concentration of ozone ~30 km altitude can be found by,

$$[O_3] = [O_2] \left(\frac{j_2 k_2 [M]}{j_3 k_3} \right)^{1/2}, \quad \text{A.19}$$

where the square brackets denote concentrations (in the upper stratosphere and mesosphere, other relations hold; Brasseur and Solomon, 1985).

Using equation (A.19), it is possible to show that there should be a peak in ozone concentration in the middle stratosphere, in overall agreement with observations. However, it is worth noting that the Chapman mechanism tends to overestimate the concentration of ozone in the middle atmosphere as no account of catalytic destruction of ozone is taken into account. These catalytic reactions have the form,



where X is the catalytic species, and



which have the net effect,



In the middle atmosphere, the principle catalytic species are believed to be HO_x , NO_x and Cl_x , where x is an odd number. HO_x concentrations are observed to peak in the mesosphere, Cl_x peaks between 40 – 50 km and NO_x peaks below 40 km. Estimates of the concentration of these three species (either through observation or chemical modelling) is therefore important in order to accurately estimate the production and loss of ozone in the middle atmosphere.

Through considering the reaction rates associated with equations (A.14-A.16), the photochemical lifetime of ozone can be found from,

$$\tau_{O_3} = \frac{1}{j_3 + j_3^* + k_3[O]}. \quad \text{A.23}$$

Brasseur and Solomon (1984) calculated the average chemical lifetime of ozone in the stratosphere at 30 km altitude to be ~ 2000 s, although this number varies considerably, especially in the polar night region. As this is short compared with global advection and diffusion timescales, a balance between local production and loss mechanisms and transport is needed to calculate the global distribution of ozone.

References

REFERENCES

Akmaev, R.A., 2001a, Simulation of large-scale dynamics in the mesosphere and lower thermosphere with the Doppler-spread parameterization of gravity waves, 1, Implementation and zonal mean climatologies, *J. Geophys. Res.*, 106, 1193-1204.

Akmaev, R.A., 2001b, Simulation of large-scale dynamics in the mesosphere and lower thermosphere with the Doppler-spread parameterization of gravity waves, 2, Eddy mixing and the diurnal tide, *J. Geophys. Res.*, 106, 1205-1213.

Akmaev, R.A., and Shved, G.M., 1980, Modelling of the composition of the lower thermosphere taking account of the dynamics with application to tidal variations of the (OI) 5577 Å airglow, *J. Atmos. Terr. Phys.*, 42, 705-716.

Akmaev, R.A., Yudin, V.A., and Ortland, D.A., 1997, SMLTM simulations of the diurnal tide: comparison with UARS observations, *Ann. Geophysicae*, 15, 1187-1197.

Alexander, M.J., 1997, A model of non-stationary gravity waves in the stratosphere and comparison to observations, from NATO ASI series, volume I, 50, Gravity wave Processes, their parameterization in global climate models, (Ed) Hamilton, K., p153-168.

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

Allen M., Lunini, J.I., and Lang, Y.L., 1984, The vertical distribution of ozone in the mesosphere and lower thermosphere, *J. Geophys. Res.*, 89, 4841-4872.

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

Andrews, D.G., 2000, An introduction to atmospheric physics, Cambridge University Press, Cambridge.

Andrews, D.G., and McIntyre, M.E., 1978a, An exact theory for nonlinear waves on a Lagrangian-mean flow, *J. Fluid Mech.*, 89, 609-646.

References

Andrews, D.G., and McIntyre, M.E., 1978b, Generalized Eliassen-Palm and Charney-Drazin theorems for waves on axisymmetric mean flows in compressible atmospheres, *J. Atmos. Sci.*, 35, 175-185.

Apruzese, J.P., Schoeberl, M.R., and Strobel, D.F., 1982, Parameterization of IR cooling at a middle atmosphere dynamics model 1. Effects on the zonally averaged circulation, *J. Geophys. Res.*, 87, 8951-8966.

Arnold, N.F., and Robinson, TR., 1998, Solar cycle changes to planetary wave propagation and their influence on the middle atmosphere circulation, *Ann. Geophysicae.*, 16, 69-76.

Arnold, N.F., and Robinson, TR., 2000, Amplification of the influence of solar flux variations on the winter stratosphere by planetary waves, *Space Sci. Rev.*, 94, 279-286.

Arnold, N.F., and Robinson, TR., 2001, Solar magnetic flux influences on the dynamics of the winter middle atmosphere, *Geophys. Res. Lett.*, 28, 2381-2384.

Arnold, N.F., and Robinson, TR., 2003, Solar cycle modulation of the winter stratosphere: the role of atmospheric gravity waves, *Adv. Space Res.*, 31, 2121-2126.

Austin, J., and Butchart, N., 1992, A three-dimensional modelling study of the influence of planetary wave dynamics on polar ozone photochemistry, *J. Geophys. Res.*, 97, 10 165-10 186.

Balsley, B.B., and Carter, D.A., 1982, The spectrum of atmospheric velocity fluctuations at 8 and 86 km, *Geophys. Res. Lett.*, 9, 465-468.

Barth, C.A., 1964, Three body reactions, *Ann. Phys.*, 5, 329.

Bates, D.R., 1951, The temperature of the upper atmosphere, *Proc. Phys. Soc. London*, 64B, 805.

Beagley, S.R., de Grandpre, J., Koshyk, J.M., McFarlane, N.A., Shepherd, T.G., 1997, Radiative dynamical climatology of the first generation Canadian middle atmosphere model, *Atmosphere-Ocean*, 35, 293-331.

References

Bochníček, J., and Hejda, P., 2002, Association between extraterrestrial phenomena and weather changes in the northern hemisphere in winter, *Surveys in Geophysics*, 23, 303-333.

Bošková, J., and Laštovička, J., 2001, Seasonal variation of gravity wave activity in the lower ionosphere in central Europe, *Studia Geoph. et Geod.*, 45, 85-92.

Brasseur, G., and Solomon, S., 1984, *Aeronomy of the middle atmosphere*, Reidel, Dordrecht, Netherlands.

Brewer, A.W., 1949, Evidence for a world circulation provided by measurements of helium and water vapour distribution in the stratosphere, *Q. J. Roy. Met. Soc.*, 75, 351.

Broutman, D., Macaskill, C., McIntyre, M.E., and Rottman, J.W., 1997, On Doppler-spreading models of internal waves, *Geophys. Res. Lett.*, 24, 2813-2816.

Brown G.M., and John, J.I., 1979, Solar cycle influences in tropospheric circulation, *J. Atmos. Terr. Phys.*, 41, 43-52.

Burrage, M.D., Skinner, W.R., Gell, D.A., Hays, P.B., Marshall, A.R., Ortland, D. A., Manson, A.H., Franke, S.J., Fritts, D.C., Hoffman, P., McLandress, C., Niecejewski, R., Schmidlin, F.J., Shepherd, G.G., Singer, W., Tsuda, T., Vincent, R. A., 1996, Validation of mesosphere and lower thermosphere winds from the high resolution Doppler imager on UARS, *J. Geophys. Res.*, 101, 10 365-10 392.

Butchart, N., Clough, S.A., Palmer, T.N., and Trevelyan, P.J., 1982, Simulations of an observed stratospheric warming with quasigeostrophic refractive index as a model diagnostic, *Q. J. R. Meteorol. Soc.*, 108, 475-502.

Chapman, S., 1930, A theory of upper atmosphere ozone, *Mem. Roy. Met Soc.*, 3, 103-125.

Chapman, S., and Lindzen, R.S., 1970, *Atmospheric tides*, Reidel, Dordrecht, Netherlands.

Chao, W.C., and Schoeberl, M.R., 1984, On the linear approximation of gravity wave saturation in the mesosphere, *J. Atmos. Sci.*, 41, 1893-1898.

References

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

Charney, J.G., Fjoroft, R., and von Neumann, J., 1950, Numerical integration of the barotropic vorticity equation, *Tellus*, 1, 38-54.

Chen, P., 1992, Two-day oscillation of the equatorial ionization anomaly, *J. Geophys. Res.*, 97, 6343-6357.

Chimonas, G., and Grant, J.R., 1984, Shear excitation of gravity waves, part II, Upscale scattering from Kelvin-Helmholtz waves, *J. Atmos. Sci.*, 41, 2278-2288.

Christiansen, B., 1999, Stratospheric vacillations in a general circulation model, *J. Atmos. Sci.*, 56, 1858-1872.

Codrescu, M.V., Fuller-Rowell, T.J., Roble, R.G., and Evans, D.S., 1997, Medium energy particle precipitation influences on the mesosphere and lower thermosphere, *J. Geophys. Res.*, 102, A9, 19 997-19 988.

Codrescu, M.V., Fuller-Rowell, T.J., Foster, J.C., Holt, J.M., and Cariglia, S.J., 2000, Electric field variability associated with the Millstone Hill electric field model, *J. Geophys. Res.*, 105, A3, 5265-5274.

Coy, L., and Fritts, D.C., 1988, Gravity wave heat fluxes: a Lagrangian approach, *J. Atmos. Sci.*, 45, 1770-1780.

DeMore, W.B., Sander, S.P., Howard, C.J., Ravishankara, A.R., Golden, D.M., Kolb, C.E., Hampson, R.F., Kurylo, M.J., and Molina, M.J., 1997, Chemical kinetics and photochemical data for use in stratospheric modeling, NASA Jet Propulsion Laboratory, JPL Publication, 97-94.

Dickinson, R.E., 1969, Theory of planetary wave-zonal mean interaction, *J. Atmos. Sci.*, 26, 73-81.

Dobson, G.M.G., 1956, Origin and distribution of polyatomic molecules in the atmosphere, *Proc. Roy. Soc. London A.*, 236, 187.

Dunkerton, T.J., 1982, Wave transience in a compressible atmosphere, part 3: the saturation of internal gravity waves in the mesosphere, *J. Atmos. Sci.*, 39, 1042-1051.

References

- Dunkerton, T.J., Hsu, C.-P.F., and McIntyre, M.E., 1981, Some Eulerian and Lagrangian diagnostics for a model stratospheric warming, *J. Atmos. Sci.*, 38, 819-843.
- Ebel A., Manson, A.A., and Meek, C.E., 1987, Short period fluctuations of the horizontal wind measured in the upper middle atmosphere and possible relationships to internal gravity waves, *J. Atmos. Terr. Phys.*, 49, 385-401.
- Eckermann, S.D., 1995, On the observed morphology of gravity-wave and equatorial-wave variance in the stratosphere, *J. Atmos. Terr. Phys.*, 57, 105-134.
- Eckermann, S.D., 1997, Influence of wave propagation on the Doppler spreading of atmospheric gravity waves, *J. Atmos. Sci.*, 54, 2554-2573.
- Eckermann, S.D., and Marks, C.J., 1996, An idealized ray model of gravity wave-tidal interactions, *J. Geophys. Res.*, 101, 21 195-21 212.
- Eddy, J.A., 1976, The Maunder Minimum, *Science*, 192, 1189-1202.
- England, S.L., Arnold, N.F., and Robinson, T.R., 2004, Solar modulation of the winter stratosphere via dynamical coupling with the upper atmosphere, submitted to *J. Atmos. Sol. Terr. Phys.*, October 2004.
- Ern, M., Preusse, P., Alexander, M.J., and Warner, C.D., 2004, Absolute values of gravity wave momentum flux derived from satellite data, *J. Geophys. Res.*, 109, doi:10.1029/2004JD004752.
- Ertel, H., 1942, Ein neuer hydrodynamischer Wirbelsatz, *Meteorol. Z.*, 59, 271-281.
- Fauliot, V., Thuiller, G., and Vial, F., 1997, Mean vertical wind in the mesosphere-lower thermosphere region (80-120 km) deduced from the WINDII observations on board UARS, *Ann. Geophys.*, 15, 1221-1231.
- Fennelly, J.A. and Torr, D.G., 1992, Photon Cross Sections for O, N₂, O₂ and N. *Atomic Data and Nuclear Data Tables*, Vol 51, No. 2, 322-360.
- Fels, S.B., 1982, A parameterization for scale dependent radiative damping in the middle atmosphere, *J. Atmos. Sci.*, 39, 1141-1152.

References

- Fetzer, E.J., and Gille, J.C., 1996, Gravity wave variance in LIMS temperatures, part II: Comparison with the zonal-mean momentum balance, *J. Atmos. Sci.*, 53, 398-410.
- Fisher, M., 1987, The Met O 20 stratosphere-mesosphere model, Dynamical Climatology Technical Note No. 52, Meteorological Office, Bracknell, UK.
- Fomichev, V.I, and Blanchet, J.P., 1995, Development of the new CCC/GCM radiation model for extension into the middle atmosphere, *Atmos. Ocean*, 33, 513-529.
- Fomichev, V.I., Blanchet, J. -P., and Turner, D.S., 1998, Matrix parameterisation of the 15 μ m band cooling in the middle and upper atmosphere for variable CO₂ concentration, *J. Geophys. Res.*, 103, D10, 11 505-11 528.
- Fomichev, V.I. and Shved, G.M., 1985 , Paramaterization of the radiative flux divergence in the 9.6 micro-m O3 band, *J. Atmos. Terr. Phys.*, 47, No. 11, 1037-1049.
- Fomichev, V.I., Ward, W.E., Beagley, S.R., McLandress, C., McConnell, J.C., McFarlane, N.A., and Shepherd, T.G., 2002, Extended Canadian Middle Atmosphere Model: Zonal-mean climatology and physical parameterizations, *J. Geophys. Res.*, 107, doi: 10.1029/2001JD000479.
- Forbes, J.M., 1982, Atmospheric tides I, Model description and results for the solar diurnal component, *J. Geophys. Res.*, 87, 5222-5240.
- Forbes, J.M, Gu, J., and Miyahara, S., 1991, On the interactions between gravity waves and the propagating diurnal tide, *Planet. Space Sci.*, 39, 1249-1257.
- Foster, J.C., Holt, J.M, Musgrove, R.G., and Evans, D.S., 1986, Ionospheric convection associated with discrete levels of particle precipitation, *Geophys. Res. Lett.*, 13, 656-659.
- Fritts, D.C., 1984a, Gravity wave saturation in the middle atmosphere: a review of theory and observations, *Rev. Geophys.*, 22, 275-308.
- Fritts, D.C., 1984b, Shear excitation of atmospheric gravity waves, II: nonlinear radiation from a free shear layer, *J. Atmos. Sci.*, 41, 524-537.
- Fritts, D.C., and Alexander, M.J., 2003, Gravity wave dynamics and effects in the middle atmosphere, *Rev. Geophys.*, 41,doi:10.1029/2001RG000106.

References

- Fritts, D.C., and Chou, H.-G., 1987, An investigation of the vertical wavenumber and frequency spectra of gravity wave motions in the lower stratosphere, *J. Atmos. Sci.*, 40, 3610-3624.
- Fritts, D.C., and Dunkerton, T.J., 1985, Fluxes of heat and constituents due to convectively unstable gravity waves, *J. Atmos. Sci.*, 42, 549-556.
- Fritts, D.C., and Lu, W., 1993, Spectral estimates of gravity wave energy and momentum fluxes, II, parameterization of wave forcing and variability, *J. Atmos. Sci.*, 50, 3695-3713.
- Fritts, D.C., Tsuda, R., VanZandt, T.E., Smith, S.A., Sato, T., Fukao, S., and Kato, S., 1990, Studies of velocity fluctuations in the lower atmosphere using the MU radar, II, Momentum fluxes and energy densities, *J. Atmos. Sci.*, 47, 51-66.
- Fritts, D.C., and VanZandt, T.E., 1993, Spectral estimates of gravity wave energy and momentum fluxes, I, Energy dissipation, acceleration, and constraints, *J. Atmos. Sci.*, 50, 3685-3694.
- Fritts, D.C., and Vincent, R.A., 1987, Mesospheric momentum flux studies at Adelaide, Australia: observations and a gravity wave/tidal interaction model, *J. Atmos. Sci.*, 48, 1535-1544.
- Fukao, S., Yamanaka, M.D., Ao, N., Hocking, W.K., Sato, T., Yamamoto, M., Nakamura, T., Tsuda, T., and Kato, S., J., 1994, Seasonal variability of vertical eddy diffusivity in the middle atmosphere, 1: Three-year observations by the middle and upper atmosphere radar, *Geophys. Res.*, 99, 18 973-18 988.
- Fuller-Rowell, T.J., 1993, Modelling the solar cycle change in nitric oxide in the thermosphere and upper mesosphere, *J. Geophys. Res.*, 98, 1559-1570.
- Fuller-Rowell, T.J., Rees, D., Quegan, S., Moffett, R.J., Codrescu, M.V., and Millward, G.H., 1996, A coupled thermosphere ionosphere model, Solar terrestrial energy program (STEP), handbook of ionospheric models, (Ed) Schunk, R.W.
- Fuller-Rowell, T.J., and Evans, D.S., 1987, Height-integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data, *J. Geophys. Res.*, 92, 7606-7618.

References

Gage, K.S., McAfee, J.R., and Williams, C.R., 1996, On the annual variation of tropospheric zonal winds observed above Christmas Island in the central equatorial Pacific, *J. Geophys. Res.*, 101, 15 061-15 070.

Garcia, R.R., and Solomon, S., 1985, A numerical model of the zonally averaged dynamical and chemical structure of the middle atmosphere, *J. Geophys. Res.*, 90, 3850-3868.

Gardiner, B.G., 1988, Comparative morphology of the vertical ozone profile in the Antarctic spring, *Geophys. Res. Let.*, 15, 901-904.

Gary, B.L., 1989, Observational results using the microwave temperature profiler during the airborne Antarctic ozone experiment, *J. Geophys. Res.*, 94, 11 223-11 231.

Gavrilov, N.M., Fukao, S., Nakamura, T., Jacobi, C., Kürschner, D., Manson, A.H., Meek, C.E., 2002, Comparative study of interannual changes of the mean winds and gravity wave activity in the middle atmosphere over Japan, Central Europe and Canada, *J. Atmos. Sol. Terr. Phys.*, 64, 1003-1010.

Gavrilov, N.M., Kalchenko, B.V., Kashcheev, B.L., and Shved, G.M., 1981, Discovery of a connection between internal gravity wave intensity in the upper atmosphere and tidal phases, *Akademiia Nauk SSSR, Izvestiia, Fizika Atmosfery i Okeana*, 17, 680-689.

Gavrilov, N.M., Manson, A.H., and Meek, C.E., 1995, Climatological monthly characteristics of middle atmosphere gravity waves (10 min-10 h) during 1979-1993 at Saskatoon, *Ann. Geophysicae*, 13, 285-295.

Gavrilov, N.M., Richmond, A.D., Bertin, F., Lafeuille, M., 1994, Investigation of seasonal and interannual variations of internal gravity wave intensity in the thermosphere over Saint Santin, *J. Geophys. Res.*, 99, 6297-6306.

Geller, M.A., 1983, Dynamics of the middle atmosphere, *Sp. Sci. Rev.*, 34, 359.

Geller, M.A., 1988, Solar cycles and the atmosphere, *Nature*, 332, 584-585.

Gill, A.E., 1982, *Atmosphere-ocean dynamics*, Academic, San Diego, California, p662.

References

Gille, J.C., and Russell III, J.M., 1984, The limb infrared monitor of the stratosphere: experiment description, performance, and results, *J. Geophys. Res.*, 89, 5125-5140.

Gille, J.C., and Lyjak, L.V., 1984, An overview of wave-mean flow interactions during the winter of 1978-79 derived from LIMS observations, *Dynamics of the Middle Atmosphere*, (Ed) Holton, J.R., and Matsuno, T., Terrapub, Tokyo, p289-306.

Gossard, E., and Munk, W., 1954, On gravity waves in the atmosphere, *J. Atmos. Sci.*, 11, 259-269.

Glass, M., Fellous, J.L., Massebeuf, M., Spizzichino, A., Lysenko, I.A., and Portnyagin, Y.I., 1975, Comparison and interpretation of simultaneous wind measurements in the lower thermosphere at Garchy (France) and Obnisk (USSR) by meteor radar technique, *J. Atmos. Terr. Phys.*, 37, 1077-1087.

Hagan, M.E., 1996, Comparative effects of migrating solar sources on tidal signatures in the middle and upper atmosphere, *J. Geophys. Res.*, 101, 21 213-21 222.

Hagan, M.E., Forbes, J.M., and Vial, F., 1995, On modelling migrating solar tides, *Geophys. Res. Lett.*, 22, 893-896.

Haigh, J.D., 1996, The impact of solar variability on climate, *Science*, 272, 981-984.

Hall, C.M., 2002, On the influence of neutral turbulence on ambipolar diffusivities deduced from meteor trail expansion, *Ann. Geophys.*, 20, 1857-1862.

Hamilton, K., Wilson, R.J., and Hemler, R.S., 1999, Middle atmosphere simulated with high vertical and horizontal resolution versions of a GCM: Improvements in the cold pole bias and generation of a QBO-like oscillation in the tropics, *J. Atmos. Sci.*, 56, 22, 3829-3846.

Harris, M.J., 2001, A new coupled middle atmosphere and thermosphere general circulation model: Studies of dynamic, energetic and photochemical coupling in the middle and upper atmosphere, University College London, PhD Thesis.

Harris, M.J. Arnold, N.F., and Aylward, A.D., 2002, A study into the effect of the diurnal tide on the structure of the background mesosphere and thermosphere using the new coupled middle atmosphere and thermosphere (CMAT) general circulation model, *Ann. Geophysicae.*, 20, 225-235.

References

Haurwitz, B., 1961, Frictional effects and the meridional circulation of the mesosphere, *J. Geophys. Res.*, 66, 2381.

Haynes, P.H., Marks, C.J., McIntyre, M.E., Shepherd, T.G., and Shine, K.P., 1991, On the downward control of extratropical diabatic circulations by eddy-induced mean zonal forces, *J. Atmos. Sci.*, 48, 651-679.

Hays, P.B., Wu, D.L., and the HRDI Science Team, 1994, Observations of the diurnal tide from space, *J. Atmos. Sci.*, 51, 3077-3093.

Hedin, A.E., 1991, Extension of the MSIS thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, 96, A2, 1159-1172.

Hedin, A.E., Flemming, E.L., Manson, A.H., Schmidlin, F.J., Avery, S.K., and Franke, S.J., 1993, Empirical wind model for the middle and lower atmosphere 1: Local time average, NASA Tech. Memo. NASA TM-104 581.

Henke, B.L., Gullikson, E.M., and Davis, J.C., 1993, Atomic Data and Nuclear Data Tables Vol. 54 No. 2

Hines, C.O., 1960, Internal atmospheric gravity waves at ionospheric heights, *Can. J. Phys.*, 38, 1441-1481.

Hines, C.O., 1993, The saturation of gravity waves in the middle atmosphere, part IV, Cutoff of the incident wave spectrum, *J. Atmos. Sci.*, 50, 3045-3060.

Hines, C.O., 1996, Nonlinearity of gravity wave saturated spectra in the middle atmosphere, *Geophys. Res. Lett.*, 23, 3309-3312.

Hines, C.O., 1997a, Doppler spread parameterisation of gravity wave momentum deposition in the middle atmosphere, Part I, Basic formulation, *J. Atmos. Solar-Terr. Phys.*, 59, 371-386.

Hines, C.O., 1997b, Doppler spread parameterisation of gravity wave momentum deposition in the middle atmosphere, Part II, Broad and quasi monochromatic spectra, *J. Atmos. Solar-Terr. Phys.*, 59, 387-400.

References

Hines, C.O., 1999, Comments on paper by S.D. Eckermann, "Influence of wave propagation on the Doppler spreading of atmospheric gravity waves", J. Atmos. Sci., 56, 1094-1098.

Hines, C.O., 2002, Comments on the gravity wave theory of J. Weinstock concerning dissipation induced by nonlinear effects, J. Atmos. Sci., 59, 2024-2030.

Hirota, I, 1984, Climatology of gravity waves in the middle atmosphere, J. Atmos. Terr. Phys., 46, 767-773.

Holton, J.R., 1982, The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere, J. Atmos. Sci., 39, 791-799, 1982.

Holton, J.R., 1983, The influence of gravity wave breaking on the general circulation of the middle atmosphere, J. Atmos. Sci., 40, 2497-2507.

Holton, J.R., 1992, An introduction to dynamic meteorology, Academic, San Diego, California, p507.

Holton, J.R., and Mass., C., 1976, Stratospheric vacillation cycles, J. Atmos. Sci., 33, 2218-2225.

Holton, J.R., and Wehrbein, W.M., 1980, A numerical model of the zonal mean circulation of the middle atmosphere, Pure Appl. Geophys., 118, 284-306.

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

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

Holton, J.R., 1975, The dynamical meteorology of the stratosphere and mesosphere, Meteorological Monographs, 37, Am. Meteorol. Soc., Boston, Massachusetts.

Hood, L.L., Jirikowicz, J.L., and McCormack, J.P., 1993, Quasi-decadal variability of the stratosphere: influence of long-term solar ultraviolet variations, J. Atmos., Sci., 50, 3941-3958.

References

- Hoppel, K., Bevilacqua, R., Allen, D., Nedoluja, G, and Randall, C., 2003, POAM III observations of the anomalous 2002 Antarctic ozone hole, *Geophys. Res. Lett.*, 30, 47-51.
- Houghton, J.T., 1978, The stratosphere and mesosphere, *Q. J. R. Meteorol. Soc.*, 104, 1-29.
- Hoyt, D.V., and Schatten, K.H., 1997, *The role of the Sun in climate change*, Oxford University Press.
- Hsu, C.-P.F., 1980, Air parcel motions during a numerically simulated sudden stratospheric warming, *J. Atmos. Sci.*, 37, 2768-2792.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D., 1996, The NMC/NCAR 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc.*, 77, 437-471.
- Keating, G.M., and Young, D.F., 1985, Interim reference ozone models of the middle atmosphere, *Handbook for MAP*, 16, Atmospheric structure and its variation in the region 20-120 km: Draft of a new reference middle atmosphere, (Ed.) Labitzke, K. , Barnett, J.J. and Edwards, B., 205-209.
- Kivelson, M.G., and Russell, C.T., 1995, *Introduction to space physics*, Cambridge University Press, Cambridge, p11.
- Klenk, K.F., Bhartia, P.K., Hilsenrath, E., and Fleig, A.J., 1983, Standard ozone profiles from balloon and satellite data sets, *J. Clim. Appl. Meterol*, 22, 2012-2022.
- Klostermeyer, J., 1984, Observations indicating parametric instabilities of internal gravity waves at thermospheric heights, *Gophys. Astrophys. Fluid Dyn.*, 29, 117-138.
- Klostermeyer, J., 1990, On the role of parameteric instability of internal gravity waves in atmospheric radar observations, *Radio Sci.*, 25, 983-995.
- Kockarts, G., 1980, Nitric oxide cooling in the terrestrial thermosphere, *Geophys. Res. Lett.*, 7, 2, 137-140.

References

Kodera, K., 1997, Influences of the solar cycle on climate through stratospheric processes, In: The stratosphere and its role in the climate system, (Ed) G.P. Brasseur, NATO ASI Series, Vol. I 54, Springer-Verlag, Berlin Heidelberg.

Labitzke, K., 1981, The amplification of height wave 1 in January 1979: a characteristic precondition for the major warming in February, Mon. Weather Rev., 109, 983-989.

Labitzke, K, and van Loon, H., 1988, Association between the 11-year solar cycle, the QBO and the atmosphere. Part I: the troposphere and stratosphere in the Northern Hemisphere winter, J. Atmos. Terr. Phys., 50, 197-206.

Labitzke, K., and van Loon, H., 1995, Connection between the troposphere and stratosphere on a decadal scale, Tellus, 47A, 275-286.

Lang, K.R., 1995, Sun, Earth and Sky, Springer-Verlag, Berlin Heidelberg New York, p210.

Laplace, P., 1825, Traite de mecanique celeste, Vol. 4, Book X., Chapter VII, section 22, reprinted in Ouevres completes de Laplace, (Ed) Gauthier-Villars, Paris, 1891.

Laštovička, J., 1999, Solar activity effects on gravity wave activity inferred from radio wave absorption in the lower ionosphere, Studia Geoph. Et Geod., 43, 107-127.

Laštovička, J., Burešová, D., and Boška, J., 1998, Effect of the QBO and the Mt. Pinatubo volcanic eruption on the gravity wave activity in the lower ionosphere, Studia Geoph. et Geod., 42, 170-182.

Laštovička, J. ,and Sauli, P., 1999, Are planetary wave type oscillations in the F2 region caused by planetary wave modulation of upward propagating tides?, Advances in Space Research, 24, 1473-1476.

Lauter, E.A., and Schminder, R., 1976, Pressure and wind variations in the mesopause region, resulting from stratospheric warmings, Physica Solariterrestris, 1, 111-119.

Lawrence, B.N., 1997, The effect of parameterized gravity wave drag on simulations of the middle atmosphere during northern winter 1991/1992 – General evolution, Gravity Wave Processes: Their Parameterization in Global Climate Models, Springer, Berlin, (Ed) Hamilton, K., 291-307.

References

Leovy, C.B., 1964, Simple models of thermally driven mesospheric circulation, J. Atmos. Sci., 21, 327-341.

Lighthill, J., 1978, Waves in fluids, Cambridge University Press, New York, p504.

Liller, W. and F.L. Whipple, 1954, High-altitude winds by meteor train photography, J. Atmos. Phys. Special Supplement, 1, 112-130.

Lindzen, R.S., 1967, Thermally driven diurnal tide in the atmosphere, Q. J. R. Meteorol. Soc., 93, 18-42.

Lindzen, R.S., 1968, The application of classical atmospheric tidal theory, Proc. R. Soc. London Ser., A303, 299-316.

Lindzen, R.S., 1981, Turbulence and stress owing to gravity wave and tidal breakdown, J. Geophys. Res., 86, 9707-9714.

London, J., 1980, Radiative energy sources and sinks in the stratosphere and mesosphere, Atmos. Ozone, Proc. NATO Adv. Study Inst., (Ed) Aiken, A.C., p703-721, Fed. Aviat. Admin., U.S. Dep. Transp. Washington, D.C.

London, J., Rottman, G.J., and Woods, T.N., 1992, Variations of solar ultraviolet irradiance derived from SOLSTICE (UARS) observations, Proc. IRS, Tallinn, Estonia, Ed. Keevallik and Karner, Deepak, Hampton, V.A.

Longuet-Higgins, M.S., 1968, The eigenfunctions of Laplace's tidal equations over a sphere, Philos. Trans. R. Soc. London Ser., A262, 511-607.

Lu, W., and Fritts, D.C., 1993, Spectral estimates of gravity wave energy and momentum fluxes, III, Gravity wave-tidal interactions, J. Atmos. Sci., 50, 3714-3727.

Manabe, S., and Wetherald, R.T., 1975, The effect of doubling the CO₂ concentration of the climate of a general circulation model, J. Atmos. Sci., 32, 3-15.

Manson, A.H., and Meek, C.E., 1993, Characteristics of gravity waves (10 min – 6 hours) at Saskatoon (52 °N, 107 °W): observations by the phase coherent medium frequency radar, J. Geophys. Res., 98, 20 357-20 367.

References

Manson, A.H., and Meek, C.E., 1997, Gravity wave spectra, direction statistics and wave interactions as observed by radars and the global MLT network. Symp. 2.01, 8th Sci. Assembly of IAGA, Uppsala.

Manson, A., Meek, C., Hall, C., Nozawa, S., Mitchell, N., Pancheva, D., Singer, W., Hoffmann, P., 2004, Mesopause dynamics from the Scandinavian Triangle of radars within the PSMOS-DATAR Project, *Annals Geophys.*, 22, 367-386.

Manson, A.H., Meek, C.E., Koshyk, J., Franke, S., Fritts, D.C., Riggin, D., Hall, C.M., Hocking, W.K., MacDougall, J., Igarashi, K., and Vincent, R.A., 2002, Gravity wave activity and dynamical effects in the middle atmosphere (60-90km): observations from an MF/MLT radar network, and results from the Canadian Middle Atmosphere Model (CMAM), *J. Atmos. Solar-Terr. Phys.*, 64, 65-90.

Marks, C.J., and Eckermann, S.D., 1995, A three-dimensional nonhydrostatic ray-tracing model for gravity waves: formulation and preliminary results for the middle atmosphere, *J. Atmos. Sci.*, 52, 1959-1984.

Martin, D.F., 1950, Cellular atmospheric waves in the ionosphere and troposphere, *Proc. R. Soc.*, 201, 216.

Matsuno, T., 1971, A dynamical model of the stratospheric warming, *J. Atmos. Sci.*, 28, 1479-1494.

Matsuno, T., 1982, A quasi one-dimensional model of the middle atmosphere circulation interacting with internal gravity waves, *J. Meteorol. Soc. Japan*, 60, 215-226.

Maunder, E.W., 1894, A prolonged sunspot minimum, *Knowledge*, 17, 173.

Mayr, H.G., Mengel, J.G., Chan, K.L., and Porter, H.S., 1998, Seasonal variation of the diurnal tide induced by gravity wave filtering, *Geophys. Res. Lett.*, 25, 943-946.

McFarlane, N.A., 1987, The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere, *J. Atmos. Sci.*, 44, 1775-1800.

References

McIntyre, M.E., 1987, Dynamics and tracer transport in the middle atmosphere: an overview of some recent developments, *Transport Processes in the Middle Atmosphere*, (Ed) Visconti, G., and Garcia, R., Reidel, p267-296.

McIntyre, M.E., 1989, On dynamics and transport near the polar mesopause in summer, *J. Geophys. Res.*, 94, 14 617-14 628.

McIntyre, M.E., 1992, Polar cooling thought experiment, in *The use of EOS for studies of atmospheric physics*, Proc Internat School Phys Enrico Fermi CXV Course, (Ed) Gille, J.C. and Visconti, G., North Holland, Amsterdam, Oxford, New York, Toronto, 313-386.

McLandress, C., 1998, On the importance of gravity waves in the middle atmosphere and their parameterization in general circulation models, *J. Atmos. Solar-Terr. Phys.*, 60, 1357-1383.

McLandress, C., Shepherd, G.G., Solheim, B.H., Burrage, M.D., Hays, P.B., and Skinner, W.R., 1996, Combined mesosphere / thermosphere winds using WINDII and HRDI data from the Upper Atmosphere Research Satellite, *J. Geophys. Res.*, 101, D6, 10 441-10 453.

McLandress, C., and Ward, W.E., 1994, Tidal/gravity wave interactions and their influence on the large-scale dynamics of the middle atmosphere: Model results, *J. Geophys. Res.*, 99, 8139-8156.

Medvedev, A.S., and Klaassen, G.P., 1995, Vertical evolution of gravity wave spectra and the parameterization of associated wave drag, *J. Geophys. Res.*, 100, 25 841-25 853.

Medvedev, A.S., and Klaassen, G.P., 2000, Parameterization of gravity wave momentum deposition based on nonlinear wave interactions: basic formulation and sensitivity tests, *J. Atmos. Solar-Terr. Phys.*, 62, 1015-1033.

Medvedev, A.S., and Klaassen, G.P., 2003, Thermal effects of saturating gravity waves in the atmosphere, *J. Geophys. Res.*, 108, doi: 10.1029/2002JD002504

Medvedev, A.S., Klaassen, G.P., and Beagley, S.R., 1998, On the role of an Anisotropic gravity wave spectrum in maintaining the circulation of the middle atmosphere, *Geophys. Res. Lett.*, 25, 509-512.

References

- Medvedev, A.S., Klaassen, G.P., and Boville, B.A., 1997, The parameterization of gravity wave drag based on the nonlinear diffusion of wave spectra, *Gravity Wave Processes: Their Parameterization in Global Climate Models*, Springer, Berlin, (Ed) Hamilton, K., 309-325.
- Meek, C.E., Reid, I.M., and Manson, A.H., 1985, Observations of mesospheric wind velocities, 2, cross sections of power spectral densities for 48-8 hours, 8-1 hours, and 1 hour to 10 min over 60 – 110 km for 1981, *Radio Sci.*, 20, 1383-1402.
- Meyer, C.K., 1999a, Gravity wave / planetary wave interactions, *Journal of Geophysical Research*, 104, 28181-28196.
- Meyer, C.K., 1999b, Gravity wave interactions with the diurnal propagating tide, *Journal of Geophysical Research*, 104, 4223-4239.
- Miller, D.E., Brownscombe, J.L., Carruthers, G.P., Prick, D.P., and Stewart, K.H., 1980, Operational temperature sounding of the stratosphere, *Phil. Trans. R. Soc. Lond.*, A292, 65-71.
- Millward, G., Moffett, R.J., Quegan, S., and Fuller-Rowell, T.J., 1996, A Coupled Thermosphere-Ionosphere-Plasmasphere Model (CTIP), *Solar Terrestrial Energy Program (STEP) Handbook*, (Ed) Schunk, R.W.
- Mitchell, N.J., Middleton, H.R., Beard, A.G., Williams, P.J.S., and Muller, H.G., 1999, The 16-day planetary wave in the mesosphere and lower thermosphere, *Annales Geophysicae*, 17, 1447-1456.
- Mitchell, N.J., Thomas, L., and Marsh, A.K.P., 1991, Lidar observations of long-period gravity waves in the stratosphere, *Ann. Geophys.*, 9, 588-596.
- Miyahara, S. and Forbes, J.M., 1991, Interactions between gravity waves and the diurnal tide in the mesosphere and lower thermosphere, *J. Meteorol. Soc. Japan*, 69, 523-531.
- Mlynczak, M.G., and Solomon, S., 1993, A detailed study of heating efficiency in the middle atmosphere, *J. Geophys. Res.*, 98, 10 517-10 541.

References

- Murtagh, D.P., Witt, G., Stegman, J., McCade, I.C., Llewellyn, E.J., Harris, F. and Greer, R.G.H., 1990, An assessment of proposed $O(^1S)$ and $O_2(b^1\Sigma_b^+)$ nightglow excitation parameters, *Planet. Space. Sci.*, 38, 43-53.
- Mueller, H.G., and Nelson, L., 1978, A travelling quasi-2 day wave in the meteor region, *J. Atmos. Terr. Phys.*, 40, 761-766.
- Munro, G.H., 1948, Short-period disturbances in the ionosphere, *Nature*, 162, 886.
- Murgatroyd, R.J., 1971, Dynamical modelling of the stratosphere and mesosphere, Mesospheric models and related experiments, Proc. 4th ERSIN-ESLAB Symposium, Frascati, Italy 6-10 July 1970, p104-121.
- Murgatroyd, R.J., and Singleton, F., 1961, Possible meridional circulation in the stratosphere and mesosphere, *Q. J. Roy. Met. Soc.*, 87, 125.
- Nakamura, T., Fritts, D.C., Isler, J.R., Tsuda, T, and Vincent, R.A., 1997, Short-period fluctuations of the diurnal tide observed with low-latitude MF and meteor radars during CADRE: Evidence for gravity wave/tidal interactions, *J. Geophys. Res.*, 102, 26 225-26 238.
- Nakamura, R., Tsuda, T., Fukao, S., Manson, A.H., Meek, C.E., Vincent, R.A., and Reid, I.M., 1996, Mesospheric gravity waves at Saskatoon (52 °N), Kyoto (35 °N), and Adelaide (35 °S), *J. Geophys. Res.*, 101, 7005-7012.
- Nakamura, T., Tsuda, T., Miyagawa, H., Matsushita, Y., Fukunishi, H., Takahashi, Y., and Yamada, Y., 1998, Propagation directions of gravity wave patterns observed in OH CCD images during the SEEK campaign, *Geophys. Res. Lett.*, 25, 1793-1796.
- Nakamura, T., Tsuda, T., Yamamoto, M., Fukao, S., and Kato, S., 1993, Characteristics of gravity waves in the mesosphere observed with the middle and upper atmosphere radar, 1, Momentum flux, *J. Geophys. Res.*, 98, 8899-8910.
- O'Neill, A., and Pope, V.D., 1988, Simulations of linear and nonlinear disturbances in the stratosphere, *Q. J. R. Meteorol. Soc.*, 114, 1063-1110.
- Pancheva, D., Mitchell, N., Clark, R.R., Drobjeva, J., Laštovička, J., 2002, Variability in the maximum height of the ionospheric F2-layer over Millstone Hill (September 1998–March 2000); influence from below and above, *Ann. Geophysicae*, 20, 1807-1819.

References

Pancheva, D., Schminder, R., and Laštovička, J., 1991, 27-day fluctuations in the ionospheric D-region, *J. Atmos. Terr. Phys.*, 53, 1145-1150.

Philbrick, C.R., Grossmann, K.U., Henning, R., Lange, G., Krankowsky, D., Offermann, D., Schmidlin, F.J., and von Zahn, U., 1983, Vertical density and temperature structure over Northern Europe, *Adv. Space Res.*, 2, 121.

Phillips, N.A., 1956, The general circulation of the atmosphere: a numerical experiment, *Q. J. R. Meteor. Soc.*, 82, 123-164.

Palmer, T.N., 1981, Diagnostic study of a wavenumber-2 stratospheric sudden warming in a transformed Eulerian-mean formalism, *J. Atmos. Sci.*, 38, 844-855.

Palmer, T.N., Shutts, G.J., Swinbank, R., 1986, Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterisation, *Q. J. R. Meteorol. Soc.*, 112, 1001-1039.

Pawson, S., 1997, Effects of gravity wave drag in the Berlin troposphere-stratosphere-mesosphere GCM, In *Gravity Wave Processes, Their Parameterization in Global Climate Models*, (Ed) Hamilton, K., NATO ASI Series, Vol. I 50, Springer-Verlag, Berlin Heidelberg.

Pawson, S., Langematz, U., Radek, G., Schlese, U., and Strauch, P., 1998, The Berlin troposphere-stratosphere-mesosphere GCM: sensitivity to physical parameterizations, *Q. J. R. Meteorol. Soc.*, 124, 1343-1371.

Portnyagin, Y.I., Solovjova, T.V., Makarov, N.A., Merzlyakov, E.G., Manson, A.H., Meek, C.E., Hocking, W., Mitchell, N., Pancheva, D., Hoffmann, P., Singer, W., Murayama, Y., Igarashi, K., Forbes, J.M., Palo, S., Hall, C., Nozawa, S., 2004, Monthly mean climatology of the prevailing winds and tides in the Arctic mesosphere/lower thermosphere, *Ann. Geophys.*, 22, 10, 3395-3410.

Plumb, R.A., 1983, Baroclinic instability of the summer mesosphere: a mechanism for the quasi-two-day wave?, *J. Atmos. Sci.*, 40, 262-270.

References

- Plumb, R.A., and McEwan, A.D., 1978, The instability of a forced standing wave in a viscous stratified fluid: a laboratory analogue of the quasi-biennial oscillation, *J. Atmos. Sci.*, 35, 1827-1839.
- Preusse, P., Schaefer, B., Bacmeister, J.T., and Offermann, D., 1999, Evidence of gravity waves in CRISTA temperatures, *Adv. Space Res.*, 24, 1601-1604.
- Quegan, S., Bailey, G.L., Moffett, R.J., Heelis, R.A., Fuller-Rowell, T.J., Rees, D., and Spiro, A.W., 1982, A theoretical study of the distribution of ionisation in the high-latitude ionosphere and plasmasphere: First results of the mid-latitude trough and the light ion trough, *J. Atmos. Terr. Phys.*, 44, 619-640.
- Ramanathan, V., Picher, E.J., Malone, R.C., and Blackmon, M.L., 1983, The response of a spectral general circulation model to refinements in radiative processes, *J. Atmos. Sci.*, 40, 605-630.
- Rapp, M., Strelnikov, B., Müllemann, A., and Lübken, F.-J., 2004, Turbulence measurements and implications for gravity wave dissipation during the MaCWAVE/MIDAS rocket program, *Geophys. Res. Lett.*, 31, doi:10.1029/2003GL019235.
- Ratnam, V., Tsuda, T., Jacobi, C., and Aoyama, Y., 2004, Enhancement of gravity wave activity observed during a major Southern Hemisphere stratospheric warming by CHAMP/GPS measurements, *Geophys. Res. Lett.*, 31, doi:10.1029/2004GL019789.
- Rees, M.H., 1989, *Physics and chemistry of the upper atmosphere*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, p126.
- Reid, G.C., 2000, Solar variability and the Earth's climate: Introduction and overview, *Space Sci. Rev.*, 94, 1-11.
- Richmond, A.D., 1983, Thermospheric dynamics and electrodynamics, in *Solar Terrestrial Physics – principles and theoretical foundations*, (Ed) Carovillano, R.L. and Forbes, J.M., Reidel, Dordrecht, 523-607.
- Riggin, D.M., Fritts, D.C., Fawcett, C.D., Kudeki, E., Hitchman, M.H., 1997, Radar observations of gravity waves over Jicamarca, Peru during the CADRE campaign, *J. Geophys. Res.*, 102, 26 263-26 282.

References

Roble, R.G., 1987, The earth's thermosphere, The solar wind and the earth, Tokyo/Dordrecht, Terra Scientific Publishing, 244-264.

Roble, R.G., 1995, Energetics of the mesosphere and thermosphere, Upper mesosphere and lower thermosphere: A review of experiment and theory, Geophys. Monograph, 87, American Geophysical Union, p87.

Roble, R.G., 1977, The upper atmosphere and magnetosphere, National Academy of Sciences, Washington DC.

Roble, R.G., and Ridley, E.C., 1994, A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (TIME-GCM): Equinox solar cycle minimum simulation (30-500 km), Geophys. Res. Lett., 21, 417-420.

Roble, R.G., Ridley, E.C., and Dickinson, R.E., 1987, On the global mean structure of the thermosphere, J. Geophys. Res., 92, 8745-8758.

Roble, R.G., and Shepherd, G.G., 1997, An analysis of wind imaging interferometer observations of O(¹S) equatorial emission rates using the thermosphere-ionosphere-mesosphere-electrodynamics general circulation model, J. Geophys. Res., 102, 2467-2474.

Rocken, C., Anthes, R., Exner, M., Hunt, D., Sokolovsky, S., Ware, R., Gorbunov., M., Schreiner, W., Feng, D., Herman, B., Kuo, Y.-H., Zou., X., 1997, Analysis and validation of GPS/MET data in the neutral atmosphere, J. Geophys. Res., 102, 29 849-29 866.

Rosenfield, J.E, Schoeberl, M.R., and Geller, M.A., 1987, A computation of stratospheric diabatic circulation using an accurate radiative transfer model, J. Atmos. Sci., 44, 859-876.

Rossby, C.-G., 1939, Relations between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semipermanent centers of action, J. Marine Res., 2, 38-55.

Rossby, C.-G., 1940, Planetary flow patterns in the atmosphere, Q. J. R. Meteorol. Soc., 66, Suppl., 68-87.

References

- Rottman, G.J., 1985, Solar/stellar irradiance comparison experiment (SOLSTICE), Upper atmosphere research satellite (UARS) mission, (Ed.) Reber, C.A., Goddard Space Flight Center, Greenbelt, Maryland, p25-27.
- Sato, K., Hasegawa, F., and Hirota, I., 1994, Short-period disturbances in the equatorial lower stratosphere, J. Meteorol. Soc., Japan, 72, 859-872.
- Satoth, M., Tomita, H., Nasuno, T., Iga, S-I., Goto, K., Tanaka, Y., Tsugawa, M., Sakasita, M., Kogi, M., 2003, Development of super-high resolution atmospheric and oceanic general circulation models on quasi-uniform grids, Annual Report of the Earth Simulator Center, The Earth Simulator Center, Japan Agency for Marine-Earth Science and Technology, (Ed) Sato, T., p13-17.
- Scaife, A., Butchart, N., Warner, C.D., and Swinbank, R., 2002, Impact of a spectral gravity wave parameterization on the stratosphere in the Met Office Unified Model, J. Atmos. Sci., 59, 1473-1489.
- Scherhag, R., 1952, Die explosionsartige Stratosphärenerwärmung des Spätwinters 1951/52, Ber. Deut. Wetterdienst 38, 51-63.
- Shapiro, R., 1970, Smoothing, filtering and boundary effects, Rev. of Space Phys., 8, 359-387.
- Shepherd, G.G., 1985, Wind Imaging Interferometer (WINDII), in Upper Atmosphere Research Satellite (UARS) Mission, (Ed) Reber, C.A., Goddard Space Flight Center, Greenbelt, Maryland.
- Shepherd, G.G., Thuiller, G., Gault, W.A., Solheim, B.H., Hersom, C., Alunni, J.M., Brun, J.-F., Brune, S., Charlot, P., Cogger, L.L., 1993, WINDII, the wind imaging interferometer on the Upper Atmosphere Research Satellite, J. Geophys. Res., 98, 10 725-10 750.
- Shine, K.P., 1987, The middle atmosphere in the absence of dynamical heat fluxes, Q. J. R. Meteorol. Sol, 133, 603-633.
- Shine, K.P., 1989, Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation, Q. J. R. Meteorol. Soc., 115, 265-292.

References

- Solomon, S., Garcia, R.R., Olivero, J.J., Bevilacqua, R.M., Schartz, P.R., Clancy, R.T., and Muhleman, D.O., 1985, Photochemistry and transport of carbon monoxide in the middle atmosphere, *J. Atmos. Sci.*, 42, 1072-1083.
- Solomon, S., Kiehl, J.T., Garcia, R.R., and Grose, W., 1986, Tracer transport by the diabatic circulation deduced from satellite observations, *J. Atmos. Sci.*, 43, 1603-1617.
- Spiegel, E.A., 1957, The smoothing of temperature fluctuations by radiative transfer, *Astrophys. J.*, 126, 202-207.
- Strobel, D.F., 1978, Parameterization of the atmospheric heating rate from 15 to 120 km due to O₂ and O₃ absorption of solar radiation, *J. Geophys. Res.*, 82, 6225-6230.
- Strobel, D.F., Apruzese, J.P., and Schoeberl, M.R., 1985, Energy balance constraints on gravity wave induced eddy diffusion in the mesosphere and lower thermosphere, *J. Geophys. Res.*, 90, 13 067-13 072.
- Svensmark, H, and Friis-Christensen, E., 1997, Variation of cosmic ray flux and global cloud coverage – a missing link in solar-climate relationships, *J. Atmos. Sol. Terr. Phys.*, 59, 1225.
- Swinbank, R., and Ortland, D.A., 2003, Compilation of wind data for the Upper Atmosphere Research Satellite (UARS) Reference Atmosphere Project, *J. Geophys. Res.*, 108, doi: 10.1029/2002JD003135.
- Thayaparan, T., Hocking, W.K., and MacDougall, J., 1995, Observational evidence of tidal/gravity wave interactions using the UWO 2 MHz radar, *Geophys. Res. Lett.*, 22, 373-376.
- Thompson, D.W.J., Baldwin, M.P., and Wallace, J.M., 2002, Stratospheric Connection to Northern Hemisphere Wintertime Weather: Implications for Prediction, *J. Climate*, 15, 1421-1428.
- Tinsley, B.A., and Heelis, R.A., 1993, Correlation of atmospheric dynamics withs olra activity: evidence for a connection via the solar wind, atmospheric electricity, and cloud microphysics, *J. Geophys. Res.*, 98, 10 375 – 10 384.

References

- Tobiska, W.K, T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottman, and O.R. White, 2000, 'The SOLAR2000 empirical solar irradiance model and forecast tool,' J. Atmos. Solar Terr. Phys., 62, 1233-1250.
- Torr, M.R., Richards, P.G., and Torr, D.G., 1980a, A new determination of ultraviolet heating efficiency in the thermosphere, J. Geophys. Res., 85, 6819-6826.
- Torr, M.R., Richards, P.G., and Torr, D.G., 1980b, The solar ultraviolet heating efficiency in the mid-latitude thermosphere, Geophys. Res. Lett., 6, 673-676.
- Tsuda, T., Nishida, M., and Rocken, C., 2000, A global morphology of gravity wave activity in the stratosphere revealed by the GPS occultation data (GPS/MET), J. Geophys. Res., 105, 7257-7274.
- Tsuda, T., VanZandt, T.E., Mizumoto, M., Kato, S., and Fukao, S., 1991, Spectral analysis of temperature and Brunt-Väisälä frequency fluctuations observed by radiosondes, J. Geophys. Res., 96, 17 265-17 278.
- VanZandt, T.E., 1982, A universal spectrum of buoyancy waves in the atmosphere, Geophys. Res. Lett., 9, 575-578.
- Vincent, R.A., 1984, MF/HF radar measurements of the dynamics of the mesopause region – a review, J. Atmos. Terr. Phys., 46, 961-974.
- Vincent, R.A., and Fritts, D.C., 1987, A climatology of gravity wave in the mesosphere and lower thermosphere over Adelaide, Australia, J. Atmos. Sci., 44, 748-760.
- Vincent, R.A., and Reid, I.M., 1983, HF Doppler measurements of mesospheric momentum fluxes, J. Atmos. Sci., 40, 1321-1333.
- Wang, D.Y, and Fritts, D.C., 1991, Mesospheric momentum fluxes observed by the MST Radar at Poker Flat, Alaska, J. Atmos. Sci., 47, 1511-1521.
- Warner, C.D., and McIntyre, M.E., 1996, On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere, J. Atmos. Sci., 53, 3213-3235.
- Warner, C.D., and McIntyre, M.E., 2001, An ultra-simple spectral parameterization for non-orographic gravity waves, J. Atmos. Sci., 58, 1837-1857.

References

- Weinstock, J., 1982, Nonlinear theory of gravity waves: momentum deposition, generalized Rayleigh friction, and diffusion, *J. Atmos. Sci.*, 39, 1698-1710.
- Weinstock, J., 1990, Saturated and unsaturated spectra of gravity waves and scale-dependent diffusion, *J. Atmos. Sci.*, 47, 2211-2225.
- Weinstock, J., 1993, Lagrangian coordinates and their application to gravity wave spectra, *Coupling Processes in the Lower and Middle Atmosphere*, (Ed) Thrane, E.V., Bix, T.A. and Fritts, D.C., p241-260, Kluwer Academic Publishers, London.
- Wiens, R.H., Taylor, M.J., Zhang, S.P., and Shepherd, G.G., 1999, Coupled longitudinal and tidal variations of equatorial nightglow $O(^1S)$ zenith emission rates from WINDII and Christmas Island data, *Adv. Space Res.*, 24, 1577-1582.
- Wilson, R., Chanin, M.-L., and Hauchecorne, A., 1991, Gravity waves in the middle atmosphere observed by Rayleigh lidar, part 2, *Climatology*, *J. Geophys. Res.*, 96, 5169-5183.
- Wu, D.L, and Waters, J.W., 1996a, Satellite observations of atmospheric variances: A possible indication of gravity waves, *Geophys. Res. Lett.*, 23, 3631-3634.
- Wu, D.L, and Waters, J.W., 1996b, Gravity-wave-scale temperature fluctuations seen by the UARS MLS, *Geophys. Res. Lett.*, 23, 3289-3292.
- Yoden, S., 1987, Bifurcation properties of a stratospheric vacillation model, *J. Atmos. Sci.*, 44, 1723-1733.
- Yudin, V.A., Geller, M.A., Khattatov, B.V., Ortland, D.A., Burrage, M.D., McLandress, C. Shepherd, G.G., 1998, TMTM simulations of tides: comparison with UARS observations, *Geophys. Res. Lett.*, 25, 221-224.
- Yudin, V.A., Khattatov, B.V., Geller, M.A., Ortland, D.A., McLandress, C., and Shepherd, G.G., 1997, Thermal tides and studies to tune the mechanistic tidal model using UARS observations, *Ann. Geophysicae*, 15, 1205-1220.