Met Office Upper Stratospheric and Mesospheric Analyses: Validation and Improvement of Gravity Wave Drag Scheme

Submitted by

David John Long

to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Mathematics, February 2011.

I certify that all material in this thesis which is not my own work has been identified and that no material is included for which a degree has previously been conferred upon me.

. . . . . .

David John Long

Abstract

The global analysis fields of the Met Office stratospheric assimilated data set have been investigated. Systematic biases for select years were identified through validation with independent satellite observations. Particular attention was given to analyses from January 2005 to October 2009 produced from a 50 level (L50) configuration of the UM with model lid at ~0.1 hPa, and the impact on analyses fields from November 2009 to September 2010 when the middle atmospheric configuration of the UM was extended to ~0.01 hPa using 70 levels (L70).

Validation results from both the L50 and L70 analyses show that largest temperature biases occur at polar latitudes approaching the model lid in the mesosphere, exhibiting a clear seasonal cycle. Here cold biases in the winter season of the L50 analyses and warm biases in the summer season of the L70 analyses would strongly suggest that the mean meridional circulation in the mesosphere is underestimated, and that small scale gravity wave forcing supplied by the operational Ultra Simple Spectral Parameterisation (USSP) scheme is insufficient.

Based on the above validation results numerous experiments were conducted to investigate the temperature response in the mesosphere to increased gravity wave forcing. Such experiments concentrated on tuning the energy scale factor (β) in the USSP scheme and the application of a momentum conserving "opaque" lid. Furthermore, the impact of developing the USSP scheme to include direct heating from gravity wave induced turbulent dissipation was also investigated. Maximum temperature responses in the summer polar upper mesosphere of ~22

K were found when increasing the standard value of β=0.1 to β=0.14 combined with the application of an opaque lid. Magnitudes of direct heating rates due to gravity wave turbulent dissipation diagnosed via the USSP scheme were found to be consistent with previous estimates. However applying such heating would most likely have a negative impact on the L70 analyses, which already display warm biases in the upper mesosphere, strongly suggesting that additional physical processes such as eddy diffusion must also be accounted for when applying direct heating from gravity wave breaking.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Since the first attempted forecasts of numerical weather prediction (NWP) by Lewis Fry Richardson during World War I, considerable progress has been made due to the advances of human knowledge and the advent of state-of-the-art computing power. NWP forecasts are performed by running (integrating in time) complex computer models of the atmosphere, commonly known as general circulation models (GCMs), that can simulate the evolution of the atmosphere given it's initial state at a given point in time. Initially GCMs were restricted to the troposphere and did not extend far into the stratosphere and mesosphere, collectively known as the middle atmosphere. During the late 60's and 70's the pioneering work of Kasahara et al. [1973], Manabe and Hunt [1968], Manabe and Mahlman [1976] began to look at the middle atmosphere, however these models had relatively poor (low) vertical resolution and only extended into the middle stratosphere. In the early 1980's further interest in middle atmospheric models was stimulated by the need to understand the physical mechanisms behind observed phenomena such as the depletion of ozone in the stratosphere [WMO, 1986, Solomon et al., 1986]. Hence higher resolution, vertical and horizontal, models with extended vertical domains were developed. Since the 1980's middle atmospheric GCMs have received considerable attention from the scientific

community. While certain middle atmospheric GCMs have resolved all of the mesosphere and above into the thermosphere, for example those studied in Jackson and Gray [1994], McLandress et al. [2006], most have a lid no higher than 80 km (upper mesosphere), for example those detailed by Hamilton et al. [1999], Kiehl et al. [1996]. Reviews of modern middle atmospheric GCMs can be seen in Pawson et al. [2000], Randel et al. [2004].

Middle atmospheric GCMs are currently used over a wide range of research areas and time scales, from operational NWP forecasts (~1→15 days), seasonal forecasts (>30 days) and decadal (~10→50 years) and climate (>100 years) predictions. The increased vertical domain in such models allows for a more realistic representation of the feedback mechanisms between the troposphere and the middle atmosphere. This includes the impact of stratospheric variability on tropospheric forecasts [Baldwin and Dunkerton, 1999, Christiansen, 2001], the downward influence of the middle atmosphere on the troposphere through the downward propagation of solar influence [Matthes et al., 2006], impact of the mesospheric flow on the tropospheric circulation [Boville, 1984], influence of mesospheric dynamical forcing on lower stratospheric temperatures [Garcia and Boville, 1994] and the upward influence of tropospheric forcing on the middle atmosphere through the redistribution of momentum and energy by vertically propagating atmospheric waves [Fritts and Alexander, 2003, Holton and Alexander, 2000, Huang and Smith, 1991]. The development of middle atmospheric GCMs allowed the physical mechanisms responsible for the above phenomena to be investigated, thus providing a means of validating newly proposed mechanisms and hypotheses.

At the heart of any middle atmospheric GCM there is a set of governing equations which describe the motion of the atmosphere, typically based on, or a variation of, a simplified version of the Navier-Stokes equations for an ideal gas known as the "Primitive Equations". To solve these governing equations they

are numerically integrated (forward in time), typically on some form of finite grid which has a limited mesh size due to limits of computational power. Features of the atmospheric flow which are larger than the finite grid length are captured (resolved) in the model, while those smaller (sub-grid processes) must be separately parameterized. While sub-grid processes occur on smaller scales than the resolved flow, they can and do interact with the resolved scale to affect the large scale flow. For middle atmospheric GCMs the two parameterizations which have possibly the greatest impact on the large scale flow are the radiation scheme and the momentum forcing induced by the breaking of small scale gravity waves. As such, several radiation schemes [Edwards and Slingo, 1996, Fouquart and Bonnel, 1980, Ingram, 1990, Shine, 1987] and gravity wave schemes have been developed. Gravity waves generated from flow over topography are known as orographic and have zero phase speeds relative to the ground, while those generated from various other mechanisms (see Section 2.3.1) are known as non-orographic and generally have non-zero phase speeds relative to the ground. Schemes such as Palmer et al. [1986], McFarlane [1987] were developed to represent the influence of orographic gravity wave breaking, which dominates in the troposphere and lower stratosphere. Schemes such as Lindzen [1981], Hines [1997], Warner and McIntyre [2001] were developed to represent the influence of non-orographic gravity wave breaking, which dominates in the middle to upper stratosphere and mesosphere. Many middle atmospheric models which resolve above the lower stratosphere, such as the European centre for medium-range weather forecasts (ECMWF) integrated forecast system (IFS) [Gobiet et al., 2005], and the UK Meteorological Office (UKMO) global middle atmospheric configuration of the Unified Model (UM) [Jackson et al., 2001], contain both orographic and non-orographic gravity wave schemes which are run (uncoupled) together. A review of current gravity wave schemes and their general formulation is given in Fritts and Alexander [2003]. All gravity wave schemes contain significant uncertainty as they contain numerous simplifications, such as assuming that wave packets only propagate vertically and not horizontally which is physically un-

realistic. Furthermore, there is still significant uncertainty as to the amplitudes, phase speeds and exact breaking mechanisms of gravity waves, hence additional uncertainty in gravity wave schemes occurs since they contain free parameters which to some extent are arbitrarily set.

Increasingly, meteorological organisations have been using middle atmospheric GCM's that resolve the entire stratosphere and portions of the mesosphere for NWP, since the full representation of the stratosphere has been shown to improve tropospheric forecasts due to more accurate initial conditions (produced via increased assimilation of data, see below) and extended range forecasts [Baldwin and Dunkerton, 1999]. The shorter forecast length of NWP timescales allows models to be run at higher resolution, resulting in more of the atmospheric flow to be resolved, thus a more accurate representation of atmospheric dynamics and surface features. Furthermore, processes which are not thought to have a significant impact over NWP timescales such as deep ocean circulation, atmospheric chemistry and the carbon cycle can be neglected. Since NWP is an initial value problem, an accurate representation of the initial state of the atmosphere is vital to produce an accurate forecast. The accuracy of the initial conditions therefore affects the quality of the evolved forecast. Typically the initial conditions for a forecasting model are supplied by geophysical observations from various sources including but not exclusively buoys, ships, radiosonde networks, aircraft and various satellite instruments. The degrees of freedom in a modern NWP model is of the order 10^7. The total number of conventional observations is considerably less, of the order $10^4 - 10^5$, and they have a non-uniform distribution in space and time. Therefore observations on their own are insufficient to define the initial conditions of the atmosphere and must be combined with an initial first guess, also known as the background field [Bergthorsson and Doos, 1955]. In modern NWP models the background field is supplied by the previous short-range, usually 6 hours, forecast data in what is known as an 'analysis cycle'. The process of combining observational and short-range forecast data to

obtain an initial condition for NWP is known as atmospheric data assimilation. The data assimilation system therefore produces an optimal estimate of the state of the atmosphere, an "analyses", based on prior estimates of the error statistics in both the observational and forecast data. The major focus of this thesis is to assess and improve the accuracy of middle atmospheric analyses produced by the UKMO global middle atmospheric configuration of the UM. Such analyses are produced daily by the UKMO and achieved in the stratospheric assimilated data set.

1.2 Motivation

Previous studies at the Met Office have shown that improving the representation of the middle atmosphere has a positive impact on forecasts in the troposphere. This improvement occurs through increased accuracy of assimilated satellite radiance observations in the lower atmosphere, which are sensitive to biases in the middle atmosphere. The UKMO operational NWP model has recently been further extended in the vertical to include the upper mesosphere, up to ~85 km. This, therefore allows the addition of more satellite radiance channels. For this extension to have maximum impact on tropospheric forecasts it is therefore vital to determine, understand and improve the biases seen through the upper stratosphere and mesosphere. Additionally, there has been considerable research into the interaction between the stratosphere and troposphere, where improved representation of the stratosphere may lead to improved tropospheric forecasts for timescales over ~10 days. It is therefore expected that further resolving the mesosphere could possibly further reduce biases in the stratosphere and subsequently the troposphere. Here the positive impact on biases downward through the atmosphere maybe achieved by the numerous feedback mechanisms discussed above. This impact of accurately representing the mesosphere on stratospheric and tropospheric forecasts is expected to be small, particularly over the timescale of 1 day, since in the troposphere the analyses are highly constrained by assimilation of large quantities of observational data. However, there is the

possibility that knowledge gained from the issues addressed in this thesis could be applied to future chemistry and climate studies using UM configurations similar to those discussed in Hardiman et al. [2010]. There is also the possibility that improving the mesospheric representation will also improve the accuracy of space weather models being developed by the University of Bath and the Met Office, which could potentially use the mesospheric analyses as their lower boundary conditions [Bushell, 2010]. Finally, the stratospheric assimilated data set is widely used by the scientific community in numerous studies concerned with understanding physical processes such as sudden stratsopheric warmings and chemical transport modelling. Thus a detailed evaluation of biases within the data set is of importance for future users.

1.3 Thesis Structure

As noted above this thesis is concerned with the global daily analyses of the UKMO stratospheric assimilated data set. While the last evaluation of the middle atmospheric configuration of the UM was performed by Butchart and Austin [1998], Jackson et al. [2001] over a decade ago, there has been little previous study on the systematic biases contained in the daily analyses fields of the stratospheric assimilated data set. Over the past two decades there have been numerous changes to the operational system producing the daily analyses. These include changes to the dynamical core, data assimilation scheme, model resolution (both vertical and horizontal) and the physical parameterizations used. Also there was a vertical extension of the unified model (UM), used to populate the data set, from ~0.1 hPa to ~0.01 hPa in November 2009. With no definitive study of the stratospheric and mesospheric analyses over the past decade the questions we wish to address in this thesis are as follows.

1. Are there are any systematic biases in the temperature and wind fields of the stratospheric assimilated data set?

2. Through their magnitude, sign, location and seasonal evolution can we at-

tribute these systematic biases to specific aspects of the operational system?

3. What is the impact of extending the model lid on the temperature and wind biases? What do the similarities/differences between the two models tell us?

4. Through the tuning of operational parameterizations is it possible to alleviate previously defined systematic temperature biases?

5. What is the impact of applying previously neglected physical processes on the current temperature biases?

6. From what we have found above, what is the further logical approach to reduce currently operational temperature and wind biases?

The thesis structure is as follows. Chapter 2 introduces the observed distribution and seasonal cycle of middle atmospheric temperatures and winds, along with the physical mechanism behind their evolution. It is this distribution and evolution which must be accurately represented by the stratospheric assimilated data set. Chapter 3 details the data sets used in this thesis, namely the daily analyses fields from the stratospheric assimilated data set and observational data from the Earth Observing System (EOS) Microwave Limb Sounder (MLS), the Upper Atmosphere Research Satellite (UARS) MLS and the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) experiments, whose independent temperature profiles are used in the validation of the daily analyses. Chapter 4 details the temperature biases of the stratospheric assimilated data validated using the above independent satellite profiles, over selected periods based on the importance of changes made to the operational system and satellite temporal coverage. Here systematic biases are identified and attributed to specific aspects of the operational system. Over identical time periods, Chapter 5 compares the wind distributions of the stratospheric assimilated data set against those from the stratospheric processes and their role in climate (SPARC) climatology. Results from Chapters 4 and 5 suggest that there is scope to improve the operational gravity wave scheme, namely the ultra simple spectral

parameterization (USSP) of Warner and McIntyre [2001]. Thus full details of the USSP scheme and it's extension to include the additional physical process of direct dissipative heating due to gravity wave breaking is discussed in Chapter 6. Chapter 7 discusses off-line experiments using the USSP scheme, which based on the results of previous Chapters, were designed to estimate the sensitivity of gravity wave forcing and direct heating to perturbations of the free parameters in the USSP scheme. The results of Chapter 7 are then employed into free running experimental integrations of the UM designed to reduce the known biases of the operational analyses, full details and results of which can be seen in Chapter 8. Finally conclusions and suggested further work is given in Chapter 9.

Chapter 2

Background

This chapter introduces the temperature and wind fields of the middle atmosphere, discussing both their vertical and horizontal distribution and seasonal evolution. The relationship which exists between the temperature and wind fields is discussed along with the primary mechanisms which determine their distributions and seasonal evolution. Of these mechanisms the impact of dynamical forcing from numerous wave motions is detailed. Such impacts must be represented correctly in numerical models of the middle atmosphere to accurately reproduce temperature and wind fields. Since this project is primarily concerned with the upper stratosphere and mesosphere, where as detailed in later chapters the majority of biases in the Met Office analyses occur, particular attention is given to the dynamical forcing from small scale gravity waves which typically dominates in these regions.

2.1 Atmospheric Vertical Structure

Conventionally the atmosphere is divided into layers which can be distinguished on the grounds of how temperature varies with altitude. Each layer has a characteristic change of temperature with height known as the environmental lapse rate, given by the negative of the vertical temperature gradient $-dT/dz$. The first layer is known as the troposphere and typically extends vertically from

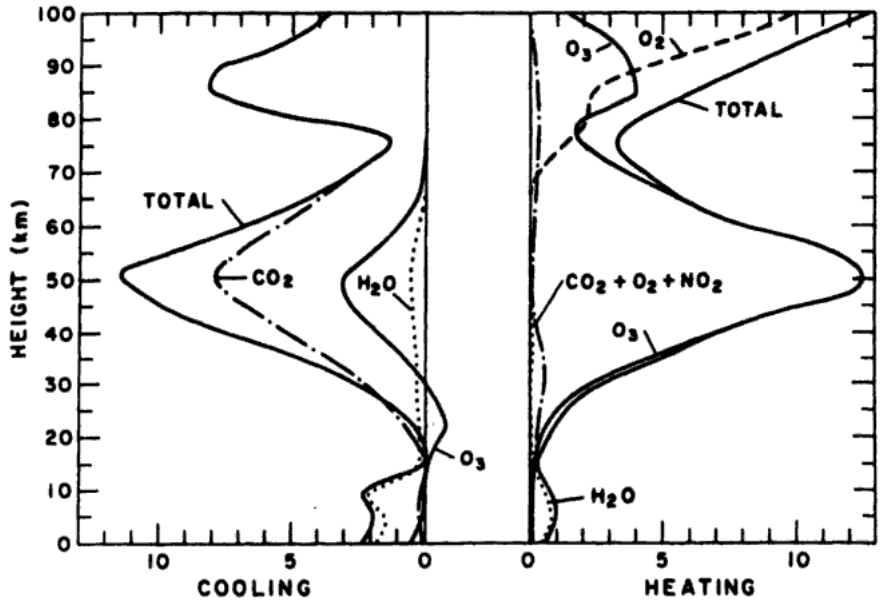

Figure 2.1: Vertical distribution of radiative heating and cooling from longwave region (left) and shortwave region (right). Units of K day^{-1}. Reproduced from [London, 1980].

the ground to an upper limit of ~15 km at tropical latitudes, and ~8 for extra-tropical latitudes, known as the tropopause. Above the tropopause the second layer of the atmosphere, known as the stratosphere, extends to ~50 km. At ~50 km there is a transition, known as the stratopause, from the stratosphere to the mesosphere which typically extends upwards to ~90 km. The middle atmosphere is conventionally described as the combination of the stratosphere and mesosphere. The mean vertical temperature structure of the middle atmosphere is primarily determined through absorption of shortwave ultraviolet radiation (heating), and thermal emission of longwave infrared radiation (cooling). The vertical distribution of this heating and cooling due to the most important radiatively active gases is shown in Figure 2.1, reproduced from London [1980].

In the shortwave region heating is primarily due to ozone, while in the longwave cooling is primarily due to carbon dioxide [Gille and Russell, 1984, Kiehl and Solomon, 1986]. Throughout most of the middle atmosphere radiative cooling is in the global mean balanced by radiative heating [Holton, 2004]. The temper-

ature distribution determined entirely by this net radiative transfer is referred to as the "radiative equilibrium" temperature T_r [Shine, 1987]. Since carbon dioxide is relatively well mixed, the vertical temperature structure of the middle atmosphere is predominantly determined by the strength of solar heating due to ozone.

The increase of ozone heating from ~15 km to it's maximum value at ~50 km, coupled with the influence of dynamical motions discussed below, results in the stratosphere having a characteristic negative environmental lapse rate. The stratospause is characterised by a change of the negative stratospheric environmental lapse rate to a positive value, where solar heating due to ozone and hence temperatures have maximum values. Above the stratopause the decrease of ozone concentration with altitude results in a decrease of the associated solar heating, hence the mesosphere is characterised by having a positive lapse rate.

2.2 The Observed Middle Atmosphere

Observational studies have shown that monthly zonal mean temperature and wind distributions in the middle atmosphere have large latitudinal variations which undergo a strong season cycle [Fleming et al., 1988]. The temperature and wind distributions are dynamically linked through approximate thermal wind balance [Hedin et al., 1996], where the vertical shear of the zonal wind (u) is related to the meridional temperature (T) gradient via,

$$\frac{\partial u}{\partial p} = \frac{R}{Pf}\left(\frac{\partial T}{\partial y}\right)_p . \tag{2.1}$$

Where R is the universal gas constant, f is the Coriolis parameter and P is the pressure. The subscript p of the term in brackets on the right hand side of equation (2.1) requires that the meridional temperature gradient is diagnosed under constant pressure. Figure 2.2 details the monthly zonal mean distributions

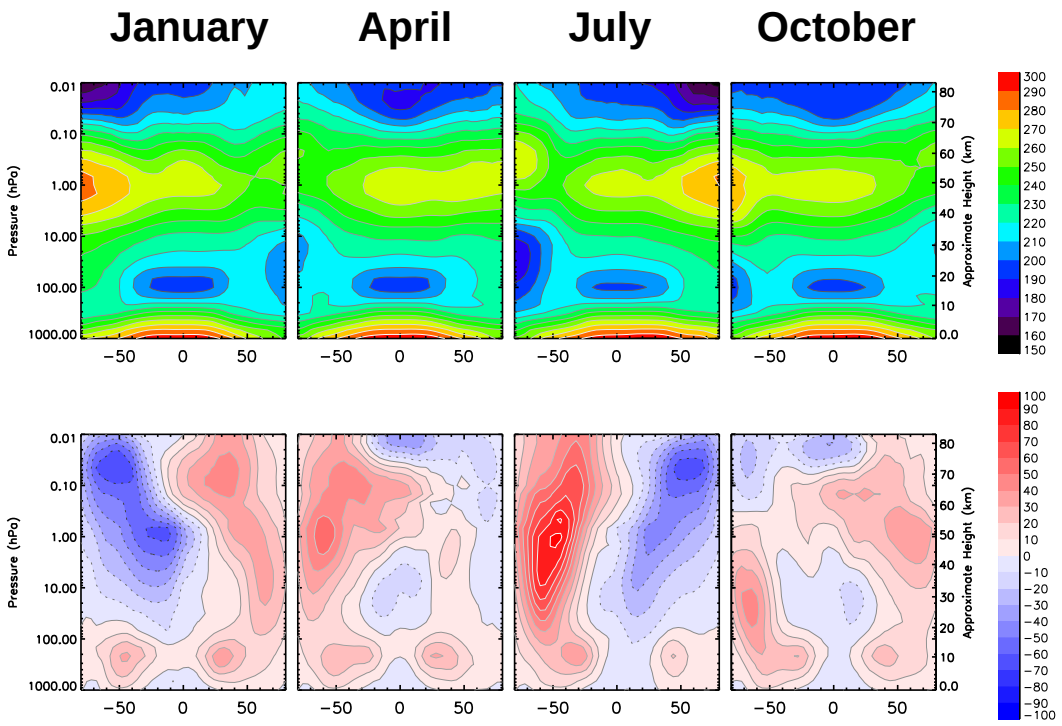

Figure 2.2: SPARC climatological zonal mean temperature (top) and zonal wind (bottom) for January, April, July and October. Contours intervals are respectively 10 K and 10 ms^{-1} for temperature and wind distributions.

of temperature and zonal wind for the months of January, April, July and October obtained from the stratospheric processes and their role in climate (SPARC) climatology detailed in Randel et al. [2004].

2.2.1 Seasonal Cycle of Zonal Mean Temperatures

Below is a description of the seasonal cycle in the atmospheric temperatures of the upper stratosphere and mesosphere based on the SPARC climatology, with reference to previous observational studies.

Upper Stratosphere (~40-55 km, ~10-1.0 hPa)

In the upper stratosphere maximum temperatures are located in the summer polar regions, with SH summer temperatures of ~280-285 K typically ~5-10

K warmer than values of ~270-275 K seen in the NH. For both months these temperature maxima decrease towards the equator with largest meridional temperature gradients seen between polar and mid-latitude regions. Unlike the middle stratosphere this decrease in temperature in the meridional plane does not extend to the winter pole. Minimum temperatures are now seen in the mid-latitude regions of the winter hemisphere where the meridional temperature gradient changes sign, resulting in increasing temperatures towards winter polar regions. In fact the winter polar stratopause of solstice months occurs at noticeably higher altitudes than in the remaining atmosphere. This elevated or "separated" stratopause feature was noted by Barnett [1974], Hitchman et al. [1989] and is most prominent in the SH winter, where temperatures of ~260-270 K are typically ~15-20 K warmer [Kanzawa, 1989] and located ~7-8 km higher than corresponding NH winter polar values.

In April the NH has relatively weak meridional temperature gradients with maximum temperatures of ~260-270 K seen at both polar and tropical latitudes. In the SH temperatures decrease by ~15-20 K from the equator to polar regions, with largest meridional gradients seen at mid-latitudes. In October temperature maximums of ~270-280 K are seen in the SH polar regions and typically decrease by ~10-15 K towards the equator. Here the largest meridional gradients are seen in the polar regions. In the NH temperatures continue to decrease by ~10-15 K from the equator to polar regions. Note that the separated stratopause feature described above for solstice months is typically absent for the months of April and October, except for the NH of October where maximum temperatures in the polar regions occur ~6→7 km higher than at all other latitudes.

Mesosphere (between ~55-85 km, ~1.0-0.01 hPa)

In the lower mesosphere during January and July tropical temperatures are near isothermal in the horizontal with relatively weak meridional gradients. For both hemispheres of winter and summer solstice months temperatures decrease

from the tropics towards mid-latitudes by ~5-10 K, where the meridional temperature gradient rapidly changes sign resulting in an increase of temperature towards polar regions. This feature is most prominent in the winter hemisphere of each month, with the SH winter having noticeably larger meridional temperature gradients in this region compared to the NH. From ~0.3-0.1 hPa these larger meridional gradients result in SH winter polar temperatures of ~250-260 K being ~20 K warmer than NH winter polar values of ~230-240 K [Fleming et al., 1990]. In the upper mesosphere during both January and July there is a near monotonic increase of temperature from the summer to winter poles, where pole to pole differences of ~60-70 K are seen for both months. Temperatures of the summer upper mesosphere are ~7 K warmer for the SH when compared to those from the NH [Huaman and Balsley, 1999, Woodman et al., 1999]. Meridional temperature gradients are largest at mid-latitudes in the summer hemisphere of both months.

In the lower mesosphere during both April and October temperature minima occur at tropical latitudes and increase towards both the northern and southern polar regions. For April the largest meridional temperature gradients are seen in the SH, while in October they are seen in the NH. For both of these months similar temperature distributions are also seen in the upper mesosphere, with relatively larger meridional temperature gradients.

2.2.2 Seasonal Cycle of Zonal Mean Zonal Wind

Below is a description of the seasonal cycle in the atmospheric winds of the upper stratosphere and mesosphere based on the SPARC climatology, with reference to previous observational studies.

Upper Stratosphere (between ~40-55 km, ~10-1.0 hPa)

For January and July zonal mean winds typically increase with altitude. For these months they are predominantly eastward (westward) for the winter (summer) hemispheres of both months, with noticeable inter-hemispheric similarities

and differences. For the summer hemispheres of both January and July, strongest winds occur between ~15-40^o latitude and decrease towards the equator and polar regions. At these latitudes and ~40 km altitude the SH westward winds of ~50-60 ms^{-1} are typically ~20 ms^{-1} stronger than the corresponding ~30-40 ms^{-1} values seen in the NH. For the winter hemispheres of both January and July, strongest values are seen between ~40-60^o latitude, and again decrease towards the equator and polar regions. Here inter-hemispheric differences in zonal wind strength are more pronounced than for the summer season, where SH summer winds of ~70-80 ms^{-1} at ~40 km are typically ~50 ms^{-1} stronger than the corresponding ~20-30 ms^{-1} values seen in the NH.

For solstice months the zonal winds in the upper stratosphere have a noticeable "tilt" with increasing altitude. In the winter hemisphere the winds tilt towards the equator with increasing altitude while in the summer hemisphere they tilt towards the polar regions.

In April the zonal mean winds are eastward for SH extra-tropical latitudes, with maximum values of ~40-50 ms^{-1} at ~40 km occurring between ~50-60^o. Throughout the middle stratosphere the tropical zonal winds are westward with maximum values of ~10-20 ms^{-1}. In the NH extra-tropics there are weak eastward winds at mid-latitudes, reversing direction to weak westward values for polar regions. In October tropical zonal winds throughout the middle stratosphere are westward with maximum values of ~10-20 ms^{-1}. In the SH eastward winds are confined pole-ward of ~40^o, where maximum values of ~40-50 ms^{-1} occur at ~25 km between ~60-70^o latitude and decrease with altitude above this region. In the NH weaker eastward winds occur pole-ward of ~30^o latitude, which increase with altitude to ~20-30 ms^{-1} at ~40 km.

Mesosphere (between ~50-85 km, ~1.0-0.01 hPa)

Unlike the majority of the stratosphere, zonal winds in the mesosphere do not continue to increase with altitude and reach a maximum value at specific altitudes. These maximum values are known as the "mesospheric jets" and are commonly described as "closed off", since they are contained within a specific region of the atmosphere above which zonal winds decrease with increasing altitude.

For both January and July the summer mesospheric jets occur at ~70 km centred between ~50-60^o. Magnitudes of these summer jets are typically ~60-70 ms^{-1}, however the previous study of Dowdy et al. [2001] has noted that summer zonal winds in the upper mesosphere of the NH show a larger degree of variance and are generally larger than those seen in the SH summer. It should be noted that for the SH summer there is a secondary jet maximum of similar magnitude located at a lower altitude of ~50 km between ~15-25^o latitude [Watanabe et al., 2009]. As for the zonal winds of the summer upper stratosphere, the summer mesospheric jets of both hemispheres have a characteristic tilt towards polar latitudes.

Compared to the summer the winter season mesospheric jets, commonly known as the polar night jets (PNJ), show a larger degree of inter-hemispheric difference. For the NH winter, jet values of ~40-50 ms^{-1} at ~65 km between ~30-40^o latitude, are considerably weaker than the SH winter values of ~85 ms^{-1} at ~48 km between ~40-50^o latitude. The PNJ's of each hemisphere have a characteristic tilt towards the equator with increasing altitude, again consistent with the zonal mean winds of the winter upper stratosphere.

In April SH winds are predominantly eastward and extend into NH tropical latitudes from ~50-75 km. SH jets of ~50-60 ms^{-1} are seen at ~50 km between ~50-60^o. In the NH there are week westward winds confined to extra-tropical latitudes in the lower mesosphere, which extend pole-ward with altitude to SH

tropical latitudes and increase to ~30-40 ms^{-1}. In October the zonal winds distribution in the lower and upper mesosphere are almost a mirror image of those seen in April. With typical wind values ~10 ms^{-1} weaker, except for the SH polar region at ~75 km. Where westward wind values of ~20-30 ms^{-1} are ~10 ms^{-1} stronger than those seen in the corresponding NH polar region in April.

2.3 Dynamical versus Radiative Forcing

As previous stated the middle atmosphere to first order is radiatively determined. However certain regions of the middle atmosphere undergo large perturbations from this radiatively determined state to produce the observed structure discussed in Sections 2.2.1 and 2.2.2. Departures from radiatively determined temperatures are predominantly forced by the transfer of momentum and energy from the lower to middle atmosphere through various wave motions [Andrews et al., 1987], hence it is commonly referred to as dynamical heating/cooling. Below we discuss the main atmospheric waves responsible for dynamically forcing the middle atmosphere, along with a description of the transformed Eulerian-mean (TEM) equations used to approximate the time mean transport. Following this we discuss the dynamical forcing and induced circulations specific to both the stratosphere and mesosphere.

2.3.1 Atmospheric Waves

The atmosphere is capable of supporting numerous forms of wave motions, with varying spatial and temporal characteristics. Such wave motions result from the combined effect of inertia and restoring forces acting on air parcels displaced from their equilibrium values. Under certain conditions atmospheric waves, primarily generated in the troposphere, are capable of propagating vertically through the atmosphere carrying energy and momentum. Vertically propagating waves continually have their characteristics altered by the changing atmospheric background state, which can lead to possible wave-mean flow interactions. The

most important wave-mean flow interaction that occurs in the middle atmosphere is non-linear dissipation or "wave breaking" [Holton and Alexander, 2000]. Vertically propagating atmospheric waves have amplitudes that increase with altitude, due to the decrease in pressure. At certain altitudes the amplitude growth is large enough that non-linear processes will dominate the evolution of the flow and the waves undergo breaking. Such wave breaking is also strongly influenced by the background flow and tends to occur when waves propagate into regions where their Doppler shifted (intrinsic) phase speeds, those relative to the mean flow, equal zero [Andrews et al., 1987]. Regions where the intrinsic phase speeds of propagating waves become zero are known as "critical layers". Upon breaking the atmospheric waves deposit their energy and momentum into the mean flow, which as described in further detail below results in a modification of the wind and temperature fields [Eliassen, 1951]. It is this response to the dynamical forcing from non-linear wave breaking that primarily results in the perturbations from a purely radiatively determined state. The atmospheric wave motions which have significant impacts on the temperature and wind distributions of the middle atmosphere are detailed below.

Rossby Waves

Rossby waves, or planetary waves, are large scale undulations of the atmospheric flow which have a restoring force due to the isentropic meridional gradient of potential vorticity, which acts to resist meridional displacements. Rossby waves have westward phase speeds relative to the mean flow and hence carry westward angular momentum.

Free travelling Rossby waves are weakly excited by the resonance properties of the atmosphere. Such free oscillations, known as a normal mode of the atmosphere, generally have relatively small amplitudes and hence play a secondly role in dynamically forcing the middle atmosphere compared to forced Rossby wave motions. Examples of such normal modes can be found in Pendlebury

et al. [2008].

Forced Rossby waves are primarily generated in the troposphere by flow over continental-scale topography and ocean-land heating contrasts (planetary scale) and through baroclinic instability (synoptic scale), where the westward phase speeds of such waves rapidly increases with increasing horizontal wavelengths [Gill, 1982]. Vertically propagating Rossby waves predominantly have zero phase speeds (quasi-stationary) relative to the ground [Holton and Alexander, 2000]. The theory of Charney and Drazin [1961] states that for stronger zonal mean eastward flow, longer wavelengths (smaller wavenumbers) are required to produce Rossby waves which are quasi-stationary to the ground. Here linear wave theory based on the conservation of potential vorticity shows that for waves quasi-stationary to the ground the vertical wavenumber m satisfies,

$$m^2 = \frac{N^2}{f^2}\left[\frac{\beta}{\bar{u}} - (k^2 + l^2)\right] - \frac{1}{4H^2}. \tag{2.2}$$

Where N is the buoyancy frequency, f is the Coriolis parameter and β it's meridional gradient, $\bar{u}$ is the zonal flow, k and l are the zonal and meridional wavenumbers respectively and H is the density scale height. Thus it can be show that vertical propagation of quasi-stationary waves (when $m^2 > 0$) is restricted to,

$$0 < \bar{u} < \beta\left[k^2 + l^2 + \frac{f^2}{(4N^2H^2)}\right]^{-1} \tag{2.3}$$

Hence the vertical propagation of forced Rossby waves is primarily restricted to the winter months, where the stratospheric flow is eastward, and in the upper stratosphere where the flow is stronger for waves of large horizontal wavelengths (zonal wavenumbers 1 and 2).

Gravity Waves

Gravity waves are vertical oscillations of the atmospheric flow. Their restoring force is supplied by the stable density stratification of the atmosphere, which results in fluid parcels displaced vertically from their equilibrium positions to undergo (buoyancy) oscillations. Such waves may propagate vertically throughout the atmosphere. Gravity waves occur over a broad spectrum of spatial and temporal scales which are determined by their source mechanism. Such mechanisms include flow over topography, convection and the adjustment of unbalanced flow. Further details on gravity wave generation is given in Section 6. Typically horizontal wavelengths range between ~10→1000 km with vertical wavelengths ranging from ~1→30 km. Gravity waves with small periods (compared to the local inertial period) are not affected by the rotation of the Earth, and are known as pure internal gravity waves. Waves with periods larger than a few hours and larger horizontal wavelengths (~1000 km) are affected by the rotation of the Earth, and are known as inertia-gravity waves. Vertical propagation of inertia-gravity waves is limited to waves which have intrinsic frequencies bounded by the local inertial frequency (lower limit) and the local static stability (upper limit).

Gravity waves can propagate in all horizontal directions, hence unlike Rossby waves are capable of carrying both eastward and westward momentum in the zonal direction. Horizontal phase speeds occur over a broad spectral range, with magnitudes approaching ~80 ms^{-1} in the mesosphere.

Kelvin Waves

Kelvin waves are large scale low frequency internal gravity waves which have zero meridional velocity. Such waves are trapped at the equator due to the change in sign of the Coriolis parameter between hemispheres, and hence have eastward (positive) phase speeds carrying eastward momentum. They are forced by oceanic-scale patterns of tropical convection in the troposphere and can propagate vertically through the middle atmosphere. Observed Kelvin waves typi-

cally have zonal wavenumbers 1 or 2, vertical wavelengths between ~6→10 km and periods ranging from ~3→20 days. For a given zonal wavenumber, Kelvin waves with longer periods (lower phase speeds) will not propagate to high altitudes as they encounter critical layers and dissipation lower in the atmosphere. Therefore "slow" Kelvin waves with periods of ~10→20 days typically break in the stratosphere, "fast" waves with periods of ~5→10 days typically break in the lower mesosphere and "ultra-fast" waves with periods of ~3→4 days typically break in the upper mesosphere and lower thermosphere as detailed by references within Ern et al. [2007].

Mixed Rossby-Gravity Waves

Mixed Rossby-Gravity waves, like Kelvin waves, are large scale low frequency internal gravity waves which are trapped near the equator. As for Kelvin waves they are also forced by patterns of tropical convection in the troposphere, however such waves have non-zero meridional velocities which rapidly decay away from the equator. Rossby-gravity waves have westward phase speeds w.r.t the mean flow and hence carry westward momentum. Such waves were first observed by Yanai and Maruyama [1966] and typically have periods of ~5 days, zonal wavenumber 4 and vertical wavelengths between ~4→8 km.

2.3.2 Atmospheric Solar Tides

The regular day/night cycle in radiative heating of the atmosphere produces global scale periodic oscillations of the background flow, known as solar tides. Such tidal oscillations are primarily generated in the troposphere and stratosphere, where radiative heating is strongest, and can conservatively propagate vertically through the atmosphere with increasing amplitudes due to the decreasing density. As for vertically propagating waves, tidal oscillations approaching the upper mesosphere can become unstable and under go non-linear breaking [Lindzen and Holton, 1968]. Solar tides can be separated into two components, migrating and non-migrating. Migrating tides are sun-synchronous oscillations,

forced by zonally symmetric thermal sources, that propagate westwards with the apparent motion of the sun. The periods of such oscillations are related to the sub-harmonics of the 24 hour solar day, where observations have shown the existence of tides with periods of 24 (diurnal), 12 (semi-diurnal), 8 (ter-diurnal) and even 6 hours [Zhang et al., 2006]. It is thought that the diurnal tide is primarily generated in the troposphere by solar heating due to water vapour, while the semi-diurnal tide is primarily generated in the stratosphere by solar heating due to ozone. Non-migrating tidal signals have identical periods to migrating tides, however they either do not propagate horizontally or propagate eastward/westward at a different speed to the motion of the sun. Such non-migrating solar tides require additional longitudinal dependent thermal forcing, such as latent heating from deep tropical convection [Hagan and Forbes, 2002]. In addition non-migrating tides can also be excited when migrating tides propagate trough a zonally asymmetric background state, i.e through non-linear interactions with both stationary planetary waves [Hagan and Roble, 2001] and distributions of gravity waves [McLandress and Ward, 1994].

In the extra-tropics the migrating semi-diurnal tide dominates with maximum amplitude values seen approaching and above the mesopause during the solstice seasons [Zhang et al., 2006]. Throughout the mesosphere the semi-diurnal tide has magnitudes comparable to those from the diurnal tide, with amplitudes of ~2→4 K at ~80 km and ~0→2 K at ~63 km. In the extra-tropics the zonal acceleration due to breaking tides has largest impact in the lower thermosphere, with a relatively insignificant impact on the zonal winds of the upper mesosphere as noted by Miyahara and Wu [1989].

2.3.3 The Transformed Eulerian-mean Equations

The overall time mean transport of individual fluid parcels in the middle atmosphere is known as the Lagrangian circulation. Here contributions to the circulation are primarily made from dynamical forcing and secondarily from

differential radiative heating due to the annual solar cycle, which has it's greatest influence at the stratopause where radiative heating is largest. Since calculation of this Lagrangian mean motion is very complicated and when applying it's conventional Eulerian mean counterpart large differences in the mean circulation may occur due to the associated Stokes drift [Dunkerton, 1978, Stokes, 1847], it is generally approximated by using the TEM equations introduced by Andrews and McIntyre [1976]. The TEM equations are used to describe the total atmospheric state as a super-position of zonal mean and "eddy" parts. Here the eddies are wave-like zonal asymmetries which describe the local departure from the zonally averaged fields. The zonal averaging approach is justified since both absolute angular momentum and the daily incident solar flux at the top of the atmosphere are independent of longitude. Therefore the majority of atmospheric variability is captured in the zonally averaged fields. As detailed in Andrews et al. [1987], using spherical coordinates (λ, ϕ, z) and potential temperature θ as the thermodynamic variable the zonal mean meridional and vertical components $(\bar{v}^*, \bar{w}^*)$ of the TEM equations can be defined as,

$$\bar{v}^* = \bar{v} - \frac{1}{\rho_o}\frac{\partial}{\partial z}\left(\frac{\rho_o \overline{v'\theta'}}{d\bar{\theta}/dz}\right) \tag{2.4}$$

$$\bar{w}^* = \bar{w} + \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}\left(\frac{\cos\phi \overline{v'\theta'}}{d\bar{\theta}/dz}\right). \tag{2.5}$$

Here over-bars represent zonal (Eulerian) means and primes denote the eddy fields, all other variables have their usual meanings. This zonally averaged approximation to the true Lagrangian motion is often referred to as the "residual mean meridional" circulation. Taking the zonal mean form of the primitive equations [Holton, 2004] and substituting in the above expressions for the residual velocities, the TEM zonal momentum balance and thermodynamic equations can be written as,

$$\frac{\partial \bar{u}}{\partial t} + \bar{v}^* \left[\frac{1}{a\cos\phi} \frac{\partial}{\partial \phi}(\bar{u}\cos\phi) - f \right] + \bar{w}^* \frac{\partial \bar{u}}{\partial z} = \frac{1}{\rho_o a \cos\phi} \nabla \cdot \bar{F} + \bar{X}, \quad (2.6)$$

$$\frac{\partial \bar{\theta}}{\partial t} + \frac{\bar{v}^*}{a} \frac{\partial \bar{\theta}}{\partial \phi} + \bar{w}^* \frac{\partial \bar{\theta}}{\partial z} = \bar{Q} - \frac{1}{\rho} \frac{\partial}{\partial z} \left[\rho \left(\overline{v'\theta'} \frac{1}{a} \frac{\partial \bar{\theta}/\partial \phi}{\partial \bar{\theta}/\partial z} + \overline{w'\theta'} \right) \right]. \quad (2.7)$$

Here $\bar{Q}$ is the net diabatic heating which in the middle atmosphere is due to the balance between radiative heating and cooling. The mechanical forcing due to the momentum deposition from breaking waves is separated into resolved and unresolved terms on the r.h.s. of equation (2.6). Where the contribution from unresolved forcing, in the middle atmosphere usually attributed to momentum deposition from small scale gravity waves, is denoted by $\bar{X}$ and the contribution from resolved motion, usually attributed to Rossby waves, is obtained from the divergence of the two-dimensional Eliassen-Palm (EP) flux vector $\bar{F}$ [Eliassen and Palm, 1961] which has the following meridional and vertical components,

$$\bar{F}^{(\phi)} = \rho_o a \cos\phi \left(\frac{\partial \bar{u}}{\partial z} \frac{\overline{v'\theta'}}{d\bar{\theta}/dz} - \overline{u'v'} \right) \quad (2.8)$$

$$\bar{F}^{(z)} = \rho_o a \cos\phi \left[\left(f - \frac{1}{a\cos\phi} \frac{\partial(\bar{u}\cos\phi)}{\partial \phi} \right) \frac{\overline{u'v'}}{d\bar{\theta}/dz} - \overline{u'w'} \right] \quad (2.9)$$

.

From equation (2.6) it is clear that the TEM zonal acceleration is composed of terms describing the Coriolis acceleration, horizontal and vertical advection of momentum by the residual circulation and wave forcing. The second term on

the r.h.s of Equation 2.7 represents the eddy heat flux convergence which has a relatively small magnitude under quasi-geostropic scaling [Shine, 1989]. If the eddy heat flux convergence and theta tendency terms in Equation 2.7 are ignored the resulting circulation is known as the "diabatic circulation" [WMO, 1986]. The diabatic circulation is known to be good approximation to the true Lagrangian circulation provided that eddy transience (variation with time) and dissipation can be ignored [Dunkerton, 1978]. However in the total absence of eddy forcing the mean meridional circulation must vanishes implying a state of radiative equilibrium, which as discussed above is not apparent throughout the middle atmosphere. Thus, implicit in the diabatic circulation approach is the assumption that while eddy transience and dissipation can be ignored in the thermodynamic equation it is still of primary importance in the mean momentum budget of Equation 2.6.

The TEM formulation provides valuable insight into the wave-mean flow interaction of the atmosphere, where the transformed eddy forcing terms on the r.h.s of equation (2.6) now depend on basic physical properties of the wave disturbances. Such insights include the non-interaction theorem of steady, linear, frictionless and adiabatic waves on a conservative zonal mean flow as noted in Charney and Drazin [1961], where the divergence of the EP-flux vanishes and under such conditions eddy motions produce no mean-flow interaction. The TEM formulation also highlights the concept that eddy heat flux ($\overline{v'\theta'}$) and eddy momentum flux ($\overline{v'u'}$) do not act separately on the mean field tendencies and residual circulation, however do so in a combined manner as expressed by the divergence of $\bar{F}$. Under quasi-geostrophic scaling, which generally holds for the extra-tropical middle atmosphere, it can be shown that the divergence of EP-flux is equal to the meridional eddy flux of potential vorticity (q) times the basic state density ρ_o [Edmon Jr et al., 1995].

$$\nabla \cdot \bar{F} = \rho_o \overline{v'q'} \tag{2.10}$$

Thus eddy driven mean-flow accelerations require non-zero potential vorticity fluxes. These diagnostic relations are not clearly apparent when studying the untransformed zonally averaged primitive equations. It should be noted that the TEM equations are a good approximation to the true Lagrangian circulation, particularly when eddy amplitudes are small and for timescales longer than the radiative damping timescale where the zonal wind acceleration term in equation (2.6) becomes small [Shine, 1989]. However in general it is not identical to the true circulation which is divergent due to the fact that perturbations of the background flow, produced by dynamical motions, vary with time [Andrews et al., 1987].

2.3.4 Stratospheric Forcing and Circulation

Dynamical forcing of the winter stratosphere is primarily supplied by vertically propagating Rossby waves [Shepherd, 2000]. For such waves, predominantly stationary w.r.t the ground, critical layers occur when the zonal wind is zero, occurring in the sub-tropics of the winter stratosphere. Therefore Rossby wave breaking primarily occurs between ~20-60^o latitude in a region commonly know as the "stratospheric surf zone" [Juckes and McIntyre, 1987, McIntyre and Palmer, 1983]. The breaking of Rossby waves results in the deposition of angular momentum. Since Rossby waves have westward phase speeds they deposit westward angular momentum which tends to decelerate the seasonal winter eastward zonal winds [Holton et al., 1995]. Such westward forcing has an associated Coriolis torque which induces a poleward meridional circulation for both the NH and SH winter hemispheres. By mass conservation this induced meridional mass flux produces upward vertical motion at tropical latitudes and downward vertical motion in the polar regions, known at the Brewer-Dobson circulation [Brewer, 1949, Dobson, 1956]. A schematic representation of the Brewer-Dobson circulation during the NH winter season is shown in Figure 2.3. Thus through adiabatic cooling/heating accompanying this rising/sinking

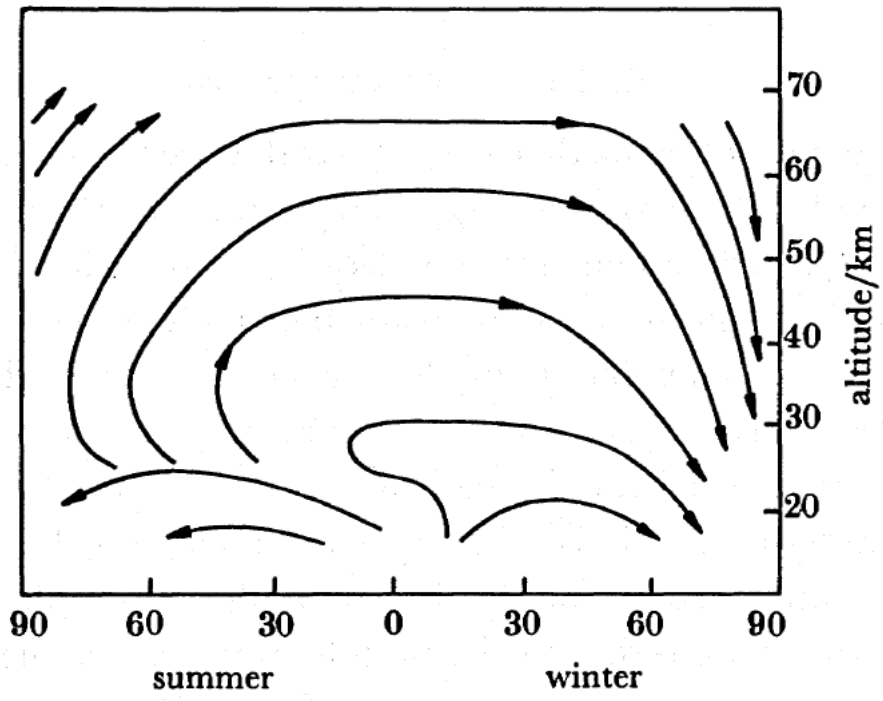

Figure 2.3: Schematic representation of the Lagrangian mean meridional circulation of the middle atmosphere under solstice conditions (January). Reproduced from Dunkerton [1978].

circulation tropical temperatures are maintained below radiative equilibrium and polar temperatures above [Andrews et al., 1987]. Hence the winter season stratospheric circulation is primarily thermally indirect as reported by Adler [1975], Hartmann [1976] and confirmed in the calculations of Haynes et al. [1991].

Numerical studies of the mean residual circulation by Eluszkiewicz et al. [1996], Gille et al. [1987], Rosenlof [1995] show that under solstice conditions the circulation is strongest and extends to higher altitudes for the NH winter compared to the SH, with maximum residual vertical velocities in the polar regions. This inter-hemispheric difference arises from the fact that Rossby wave generation is greater in the NH leading to larger momentum deposition and subsequently a stronger poleward branch of the Brewer-Dobson circulation and hence greater downwelling at polar regions [Shepherd, 2000], resulting in observed NH mid-

dle stratospheric polar temperatures being perturbed further from T_r values than in the SH, as described above in Section 2.2.1.

The Brewer-Dobson circulation during summer months is primarily restricted to the lower stratosphere as detailed in Figure 2.3, where Rossby waves reach critical layers as the zonal wind changes to a westward direction. Evidence presented by Rosenlof [1996] has shown that both planetary and synoptic scale (wavenumbers 1-8) Rossby waves provide the majority of the dynamical forcing in the NH summer lower stratosphere. For the SH summer lower stratosphere the mean meridional circulation is stronger compared to the NH, as shown by high latitude temperatures being perturbed further from T_r values due to increased downwelling and adiabatic warming. Since Rossby wave generation is smaller in the SH additional forcing is required to produce this inter-hemispheric difference. As noted by Rosenlof [1996] this additional forcing is most probably attributed to the dissipation of small scale gravity waves (see below).

The separated winter polar stratopause feature seen in the winter hemispheres is often misrepresented in modern global circulation models (GCM's), however it has been well produced in the studies of Beagley et al. [1997], Tomikawa et al. [2008]. This feature is primarily due to the influence of vertically propagating small scale gravity waves (see below) affecting the mean meridional circulation [Hitchman et al., 1989]. This reasoning has been verified by the model study of Watanabe et al. [2008], which showed that under solstice conditions planetary wave activity near and above stratopause heights is mainly restricted to the equatorward side of the polar night jet, while in the polar regions small scale gravity waves near and above the stratopause are the dominant source of dynamical forcing. As described below the gravity wave forcing of the winter hemisphere approaching and above the stratospause has a westward direction, which drives a poleward meridional circulation with descent at polar regions. This gravity wave driven descent produces adiabatic warming and hence is the

primary cause of the separated stratopause feature.

2.3.5 Mesospheric Forcing and Circulation

Departures of the mesospheric temperature and zonal wind distributions from their radiatively determined states are believed to be predominantly maintained by the breaking of small scale gravity waves [Andrews et al., 1987], with addition contributions from atmospheric tides [Miyahara and Wu, 1989]. Unlike Rossby waves, gravity waves which are generated in both the troposphere and stratosphere, see review of Fritts and Alexander [2003], can propagate vertically with eastward and westward phase speeds carrying eastward and westward angular momentum. The seasonal zonal winds of the stratosphere act as filters where only waves with large Doppler shifted phase speeds, occurring when the background flow is opposed to the phase speed direction, can propagated into the lower and upper mesosphere [Holton and Alexander, 2000]. Therefore due to the reversal of the seasonal solstice zonal mean winds in the stratosphere, breaking gravity waves in the mesosphere predominantly deposit eastward momentum in the summer hemisphere and westward momentum in the winter hemisphere. The resulting force decelerates the zonal mean winds in each hemisphere, which leads to the closing of the middle atmospheric jets described in Section 2.2.2.

For solstice months the above deceleration produces an equatorward meridional mass flux in the summer hemisphere and a poleward mass flux in the winter hemisphere, where the Coriolis torque associated with such a drift is in approximate balance with the wave induced zonal forcing of each hemisphere. The resulting solstice circulation in the mesosphere is a single cell, where as detailed in Figure 2.3 there is meridional drift from the summer to winter hemispheres [Holton and Alexander, 2000]. By mass conservation this single cell circulation results in rising motion throughout the summer hemisphere and sinking motion throughout the winter hemisphere. The associated adiabatic heating/cooling associated with this vertical motion maintains the summer mesosphere below

and the winter mesosphere above their respective radiatively determined values.

As stated above the mesospheric forcing is dependent on the gravity wave filtering by the stratospheric winds below, which as described above are influenced by Rossby wave dissipation. Hence the strength of gravity wave forcing in each hemisphere is coupled to the Rossby wave forcing below. The large Rossby wave forcing of the NH winter season results in strongly reduced eastward winds, and often to local westward winds in the longitudinal direction. Such flow will freely transmit a larger spectrum of eastward propagating gravity waves into the mesosphere, thus reducing the westward forcing. Conversely, the small Rossby wave forcing of the SH winter season results in strong eastward flow and a stable SH winter polar vortex. Under such conditions there is a freer transmission of gravity waves with westward phase speeds [Whiteway et al., 1997]. Hence the westward mesopheric gravity wave forcing of the SH winter is stronger than that seen in the corresponding NH winter season [Shine, 1989]. This inter-hemispheric differences results, through increased meridional circulation and subsidence at high latitudes, to warmer temperatures in the polar regions of the SH winter upper mesosphere compared to the NH.

For equinox months the filtering of the stratospheric zonal mean winds results in gravity wave breaking predominantly producing a weak westward force in both hemispheres. This results in a relatively weaker two cell circulation, with rising vertical motion at tropical latitudes and sinking vertical motion at both polar regions throughout the mesosphere [Andrews et al., 1987].

Using the TEM equations Haynes et al. [1991] have postulated the so called "downward control" theory. They showed that for extra-tropical latitudes, which have approximately constant values of absolute angular momentum per unit mass with height due to the dominance of the Earth's rotation, the residual mean meridional circulation at a certain altitude is directly proportional to the

magnitude and latitudinal gradient of the eddy forcing above (on the same angular momentum contour) and is unrelated to the value of eddy forcing below. Furthermore in the time mean the circulation driven by the wave forcing at a particular altitude always extends downward. Under downward control the mean residual vertical velocity $\bar{w}^*$ at some altitude z_o is approximated via,

$$\bar{w}^*(z_o) = -\frac{1}{\rho_s(z_o)\cos\theta}\frac{\partial}{\partial y}\left[\frac{\cos\theta}{f}\int_{z_o}^{\infty}\rho_s(z)\,F(z)\,dz\right], \tag{2.11}$$

where f is the Coriolis parameter, ρ_s is the standard density, θ is latitude and F is the body force per unit mass due vertically propagating atmospheric waves. Using simple theories of saturation and observational data Haynes et al. [1991] showed that the forcing from breaking waves is constant with increasing height above the saturation level, with major contributions to the residual vertical velocity at such a level occurring within two or three scale heights (~6→7 km in the mesosphere) above. While this simple argument has been shown to reasonably represent the "true" convergence of the residual vertical velocity at a particular level [Haynes et al., 1991] for mid-latitudes there are noticeable exceptions, discussed below. As noted in Haynes et al. [1991] the above argument has important implications for numerical models of the middle atmosphere, where accurate reproduction of the circulation at a particular altitude requires sufficient knowledge of the forcing above.

Through downward control it is possible for gravity wave breaking in the mesosphere to influence the mean meridional circulation of the atmosphere below. The influence of forcing in the mesosphere on the stratospheric circulation is stronger and extends to lower altitudes when the Rossby wave forcing is weak, as confirmed by the modelling study of Garcia and Boville [1994] which showed that in the SH winter pole the majority of the vertical residual velocity at 30 km is due to forcing from ~5-6 scale heights above, while in the NH winter pole it is contained within ~3 scale heights. Thus for the SH winter polar regions the

mesospheric gravity wave forcing has a significant impact on temperatures in the lower and middle stratosphere, compared to only the upper stratosphere for the NH winter season. As noted by Garcia and Boville [1994] the Rossby wave forcing has a strong impact for two reasons. Firstly, the residual vertical velocity given by equation (2.11) is density weighted, hence for the stratospheric circulation if Rossby wave forcing is present it will dominate any contributions made from forcing which occurs at larger altitudes above. Secondly, as noted above strong Rossby wave forcing in the winter stratosphere reduces the magnitude of mesospheric gravity wave forcing.

While the mesospheric circulation is dominated by gravity wave dissipation, planetary scale waves can also have a significant impact. Evidence that such waves are present throughout the mesosphere, both summer and winter hemispheres, has been given by numerous observational studies including Garcia et al. [2005]. As noted in Smith [2003] the origin of quasi-stationary planetary waves in the mesosphere is likely due to two mechanisms: 1) in the winter hemisphere vertical propagation of forced Rossby waves from the stratosphere upwards through the mesosphere, or 2) in-situ generation by dissipating gravity waves which have substantial longitudinal asymmetries of momentum flux due to the strong longitudinal variation of the stratospheric flow. In addition to quasi-stationary waves, westward propagating waves of planetary scale have also been observed in the upper mesosphere, i.e. the 2-day and 4-day waves. Such travelling waves have been attributed to both manifestations of Rossby-gravity normal modes and baroclinic/barotropic instability associated with the upper mesospheric jets [Tunbridge and Mitchell, 2009].

During winter months waves which have propagated from the stratosphere dominate the planetary scale forcing of the lower mesosphere [Smith, 2003], and under wave breaking deposit westward (negative) angular momentum, a convergence of EP-flux, which acts to enhance the residual circulation. Here

magnitudes of forcing are similar to those seen in the stratosphere, where NH values are considerably stronger than those seen in the SH [Marks, 1989]. Previous studies have shown that in the winter upper mesosphere planetary scale EP-flux values have a dipole structure, with regions of convergence in the tropics and mid-latitudes and divergence pole-ward of ~50^o latitude [Marks, 1989, McLandress et al., 2006]. As shown by Plumb [1983] regions of EP-flux divergence can occur when the meridional gradient of potential vorticity is negative, such regions occur above the PNJ primarily due to the negative vertical wind shear [Fleming et al., 1988]. As such negative meridional gradients of potential vorticity are an indication of wave generation through baroclinic or barotropic instability [Andrews et al., 1987], the EP-flux convergence/divergence dipole structure in the winter upper mesosphere suggests that planetary wave forcing in this region is dominated by waves which are generated in-situ [Smith, 1997, 2003]. In the summer upper mesosphere forced Rossby waves from the lower mesosphere are not present, and regions of divergent EP-flux at mid-latitudes are again associated with in-situ generated waves above the summer mesosphere jets due to baroclinic and barotropic instability. The magnitude of planetary wave generation in the upper mesosphere is strongly coupled to the strength of the gravity wave forcing of each season [Norton and Thuburn, 1999], since it is this forcing that is primarily responsible for weakening the jets and producing the regions of instability primarily responsible for wave generation.

The observational evidence of Karlsson et al. [2007] and the numerical studies of Karlsson et al. [2009] suggests that there is an inter-hemispheric coupling between the strength of the winter Brewer-Dobson circulation and the residual circulation in the opposing summer mesosphere. The mechanism for this inter-hemispheric coupling when considering a strengthening of the Brewer-Dobson circulation, equally valid for a weakening, is as follows. As described above, a strengthening of the Brewer-Dobson circulation results in weaker eastward winds which through altering the gravity wave filtering reduces the westward

drag of the winter mesosphere. Furthermore, the westward gravity waves break at lower altitudes due to weaker Doppler shifting of their vertical wavelengths from the weaker eastward zonal flow. Hence the residual circulation of the winter mesosphere is reduced, which through mass conservation induces colder temperatures in the polar winter mesosphere and warmer temperatures in the tropical mesosphere [Labitzke, 1972]. Through thermal wind balance this warmer tropical mesosphere, with an associated increase of the meridional temperature gradient, results in stronger eastward winds in the summer subtropical mesosphere. The stronger eastward winds result in eastward (westward) gravity waves having smaller (larger) intrinsic phase speeds and hence they break at lower (higher) altitudes, which results in increased (decreased) eastward forcing at lower (higher) latitudes [Shepherd and Shaw, 2004]. This vertical dipole structure of anomalous forcing results in a reduction and downward shift of the seasonal eastward gravity wave forcing per unit mass, and subsequently through a weaker upward residual circulation induces warmer temperatures on the poleward side of the reduced forcing region. Again the balanced response to this warming is increased eastward winds and the mechanism initiates a feedback between the zonal flow and forcing which propagates warmer temperatures to the polar regions [Becker and Fritts, 2006]. As such temperature responses in the summer polar mesosphere have been shown to be reversed when a weakening of the winter hemisphere Brewer-Dobson circulation occurs, and similar responses are found when considering both northern and southern hemispheres, suggests that the coupling mechanism is linear. Temperature responses in the polar summer mesosphere typically occur ~15-20 days after the change in forcing of the winter stratosphere [Kornich and Becker, 2010].

2.4 The Semi-Annual Oscillation

At the equatorial stratopause there is a semi-annual oscillation (SAO) of the zonal mean wind from eastward to westward flow, as first noted by Reed [1966]. The westward phase of the stratopause SAO occurs during solstice months with

typical maximum magnitudes of ~25→35 ms^{-1}, while the eastward phase occurs during equinox seasons with comparable magnitudes. During equinox months radiative heating at the stratopause is symmetrical about the equator and the eastward zonal flow is believed to be driven by dissipating fast (short period) Kelvin waves and internal gravity waves. Below the stratopause the forcing is believed to be dominated by the planetary scale Kelvin waves, while near the stratopause Kelvin waves account for ~30→70% of the forcing (depending on location and time) while the remaining contribution comes from small scale gravity waves [Hitchman and Leovy, 1988].

For solstice seasons the radiative heating near the stratopause is located off the equator in the summer hemisphere. Such conditions result in a T_r distribution with a non-zero meridional gradient across the equator. In-order to maintain thermal balance at the equator this T_r gradient is damped through inertial adjustment, which involves the meridional advection of heat and momentum [Dunkerton, 1989]. In order to conserve absolute angular momentum air parcels approaching the equator in the summer hemisphere must have decreasing zonal velocity resulting in westward flow [Shepherd, 2000]. In the winter hemisphere westward acceleration of the mean flow has been associated with the meridional transport of momentum by extra-tropical Rossby wave breaking. Analysis of numerical studies performed by Hamilton and Mahlman [1988] concluded that meridional advection of heat and momentum dominates in the summer hemisphere, while contributions from Rossby wave breaking are more significant in the winter hemisphere.

2.5 Summary

In this chapter we have discussed the seasonal evolution of both the temperature and wind fields of the middle atmosphere. While such distributions are primarily determined from radiative transfer, large perturbations from this state are caused by the dynamical forcing associated with the breaking of various at-

mospheric waves. Regions which have the largest deviations from a radiatively determined state include the winter polar stratosphere and both the winter and summer polar mesosphere. These deviations are primarily a consequence of adiabatic heating (cooling) due to the descent (ascent) of the meridional circulation at polar regions, produced by zonal forcing from breaking wave motions.

In the stratosphere the dynamical forcing is dominated by the breaking of vertically propagating Rossby waves. The resulting meridional circulation extends to higher altitudes in the winter hemisphere and is stronger for the NH compared to the SH, due to larger Rossby wave generation. Consequently NH polar winter temperatures are warmer (further from radiatively determined values) than those seen in the SH winter polar regions.

In the mesosphere the dynamical forcing is dominated by the breaking of small scale vertically propagating gravity waves. Under solstice conditions the gravity wave breaking opposes the seasonal zonal mean winds of each hemisphere, causing the zonal mean jets of each season to decrease with height and inducing a summer to winter hemisphere single cell meridional circulation. Adiabatic heating (cooling) at the winter (summer) polar regions results in a meridional temperature gradient which has a smaller and even reversed sign compared to that seen under radiative forcing alone. The extent to which the gravity wave forcing of the mesosphere impacts the stratospheric circulation and temperature below is determined through the downward control principle. Such impact has been shown to be greater in the SH winter compared to the NH.

Chapter 3

Data Sets

This chapter details the four data sets used in this thesis, namely the Met Office stratospheric assimilated data set, the Upper Atmosphere Research Satellite (UARS) Microwave Limb Sounder (MLS), the Earth Observing System (EOS) MLS and the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) experiments.

For the Met Office analyses details are given of the dynamical model, assimilation scheme and observations used to populate the data set. This includes the description of major changes made to any aspect of the system. For the UARS MLS, EOS MLS and SABER experiments details of measurement technique, orbit and coverage, spatial resolution, sources of error and results from validation with correlative data sets are presented.

3.1 Met Office Stratospheric Assimilated Analyses

3.1.1 Background

The analyses that will be validated in this project are obtained from the Met Office stratospheric assimilated dataset. The dataset spans the period from the 17th October 1991 to the present. Each analysis produced by the Met Office comprises 3 dimensional global gridded fields of temperature, geopotential height

and wind components. In essence each analysis can be considered as the best guess of the geophysical state of the atmosphere for a particular time. Since 1995 the Met Office have produced 4 analyses per day. However only analyses produced at 1200 UTC are archived. The Met Office produce this dataset using a stratospheric configuration of their numerical weather prediction (NWP) Unified Model (UM), with the aid of a stratosphere-troposphere data assimilation (DA) system.

Since the 17th October 1991 the Met Office has used three different methods of DA to produce stratospheric analyses; Analysis Correction [Lorenc et al., 1991], 3D-Var [Lorenc et al., 2000] and 4D-Var [Rawlins et al., 2007]. However as the majority of this project focuses on the quality of mesospheric analyses, where limited data is assimilated in the Met Office operational model, full details of the above DA schemes are not discussed further. However a time line of when each scheme was implemented is given in Table I.

3.1.2 Unified Model Description and Changes

At the heart of the NWP scheme that produces the analyses is a global model of the troposphere and stratosphere. This model is known as the Met Office's Unified Model (UM) which is detailed in Cullen [1993]. The UM became operational in 1991 and is a combination of the global NWP and climate models previously used at the Met Office. The UM is a grid point model where the dynamical equations that describe the evolution of the atmosphere are numerically integrated forward in time. The UM also includes a library of physical parameterizations to describe processes which either have smaller length scales than the UM's grid resolution or are not represented in the dynamical equations of the model. Examples of these parameterizations include the effects of convection, radiation and gravity wave drag (GWD).

As for the DA schemes described above the UM has had several changes since

1991 when the stratospheric assimilated dataset was first produced.

Initially the UM was based on the traditional hydrostatic primitive equations [White and Bromley, 1988]. The vertical coordinate of the UM was a "hybrid" function of pressure described in Simmons and Burridge [1981]. In the horizontal the dynamical variables were stored on a regular latitude-longitude Arakawa B grid [Arakawa and Lamb, 1977]. The vertical distribution of variables was the same as in the previous 15 level NWP model [Gadd, 1985]. There were 42 pressure levels from the ground to the model top at 0.28 hPa. This resulted in a resolution of $\sim$ 1.6 km in the stratosphere. The UM had a split explicit, conservative, finite difference integration scheme as discussed in Cullen and Davies [1991]. For this scheme the solution procedure was split into "adjustment" and "advection" parts, which were solved using different time steps. The set of physical parameterizations used in the initial UM configuration were in line with the Hadley Centre Atmospheric Model version 2b (HadAM2b) detailed in Hall et al. [1995]. This includes the radiation scheme by Ingram [1990] and the GWD scheme by Wilson and Swinbank [1990]. A full list of all parameterizations with references can be seen in the appendix of Swinbank et al. [1998].

On the 10^{th} October 2003 the dynamical core of the UM was changed and became referred to as the "New Dynamics". Some of the major features of this change, detailed in Davies et al. [2005], are listed below.

- The dynamical equations were changed to the non-hydrostatic fully compressible deep atmosphere formulation. The rationale behind using this equation set is to avoid unnecessary approximations, such as the shallow atmosphere approximation while retaining a complete representation of the Coriolis force. Also relaxing the hydrostatic assumption has been shown to improve the non-linear behaviour of the equations [Lions et al., 1992], allowing the use of horizontal grid resolutions of $\lesssim$ 10 km. A full comparison of this equation set with the previously used hydrostatic primitive

equations is given in White et al. [2005].

- The vertical terrain-following co-ordinate is now height based, making it easier to maintain global conservation principles with the new dynamical equation set.
- The dynamical variables are still stored horizontally on a regular latitude-longitude grid, except they are now staggered on an Arakawa C-grid [Arakawa and Lamb, 1977]. This grid staggering has been proven to be the most accurate for the process of geostrophic adjustment.
- Variables are stored using the Charney-Phillips vertically staggered grid.
- A Semi-Lagrangian (SL) advection scheme is used in finding solutions of the finite difference equations for all prognostic variables except density, thus allowing longer time steps which were previously constrained by the maximum velocity in any part of the computational domain using the Eulerian advection scheme.
- There are now 50 model levels from the ground to the model top which was extended to ~63 km.

Until March 2006 the Met Office were producing analyses using two different global configurations of the UM. A tropospheric configuration and a stratospheric configuration used to create the stratospheric assimilated dataset.

The tropospheric configuration was able to resolve 216 waves round a latitude circle (N216), implying that the model has 432 grid points in the longitudinal direction. The vertical range of this tropospheric model was from the ground to ~39 km.

The stratospheric configuration was only able to resolve 48 waves around a latitude circle (N48), obviously implying a lower resolution in the horizontal with only 96 grid points in the longitudinal direction. However compared to the

tropospheric model the stratospheric model had a larger vertical range extending from the ground up to ~63 km.

Post March 2006 the stratospheric and tropospheric configurations were merged to create one model. This new stratospheric model maintains the vertical range of the old stratospheric model with 50 levels from the ground to ~63 km. However the horizontal resolution was refined to 0.375^o (480 grid points) in latitude and 0.5625^o (640 grid points) in longitude (N320).

Post November 2009 the stratospheric model lid was extended to include further representation of the upper mesosphere. This new configuration now includes 70 model levels ranging from the ground to the model lid at ~80 km. The horizontal resolution was initially identical to the previous 50 level configuration, however in March 2010 was further refined to 0.2343^o (768 grid points) in latitude and 0.3515^o (1024 grid points) in longitude (N512).

Analyses produced by different configurations of the UM shall be referred to according to the number of vertical model levels used in the model. Hence analyses from November 1991 to April 2000 shall be referred to as L42, from October 2003 to October 2009 as L50 and from November 2009 to the present day as L70. Below is a list of all major changes to the UM from 1991 to the present day. Note changes made at a particular date apply to all subsequent UM configurations unless stated.

29th October 1991: The initial UM configuration was based on the hydrostatic primitive equation model described above and in Cullen [1993]. There are 42 pressure levels with the model lid at 0.28 hPa.

25th February 1992: The orographic GWD scheme produced large amounts of noise in the stratosphere. Therefore the well tested Palmer et al. [1986] scheme was introduced up to 20 hPa. Above 20 hPa a simple Rayleigh friction scheme was introduced.

5th January 1994: The long-wave radiation scheme was altered to give smoother and more accurate heating rates in the stratosphere [Morcrette et al., 1986].

28th January 1998: Operations were switched to the new Cray T3E super computer. The T3E ozone climatology was later found to be incorrect leading to increased temperatures in the mesosphere and decreased temperatures at the stratopause for the next 2 years.

16th May 2000: The sub-grid physical parameterizations were changed to bring them in line with the Hadley Centre Atmospheric Model version 3 (HadCM3) [Pope et al., 2000]. The two changes that made possibly the largest impact on the stratospheric analyses are listed below.

- The orographic GWD scheme was changed to more sophisticated Gregory et al. [1998] scheme that includes the effects of anisotropic orography. Above 20 hPa a simple Rayleigh friction scheme was still applied.

- The introduction of the Edwards and Slingo [1996] radiation scheme. The new scheme includes the heating effects of $H_2O, O_3, O_2, N_2O, CH_4$, CFC11 and CFC12.

14th November 2000: Total of model levels changed to 40 with the resolution of the stratosphere being ~2.5 km. Model lid is changed to 0.1 hPa. Problems with the ozone climatology data were corrected.

10th October 2003: The dynamical core of the UM was changed to the "New Dynamics" described above. Following Scaife et al. [2002] the GWD scheme was replaced with an Ultra Simple Spectral Parameterisation (USSP) described in Warner and McIntyre [2001]. Number of model levels increased to 50.

14th March 2006: Tropospheric and stratospheric models merged to form one stratospheric model. In addition to these changes the USSP scheme was turned off.

5th December 2006: The USSP scheme was added back in and the upper level horizontal diffusion was removed.

10th November 2009: Introduction of the extended 70 level model with model lid at ~85 km.

3.1.3 Model Levels

Post introduction of the "New Dynamics", model levels in the UM use a normalised height based vertical coordinate η with a staggered Charney-Phillips grid. For such a grid the prognostic variables θ and vertical velocity w are stored on levels which are commonly known as "theta levels". While the variables ρ, pressure and horizontal components of velocity u and v are stored on intermediate levels between adjacent theta levels, commonly known as "rho levels".

The η vertical coordinate has a value of zero at the lower surface boundary and a value of one at the model lid. Theta and rho η levels near the surface

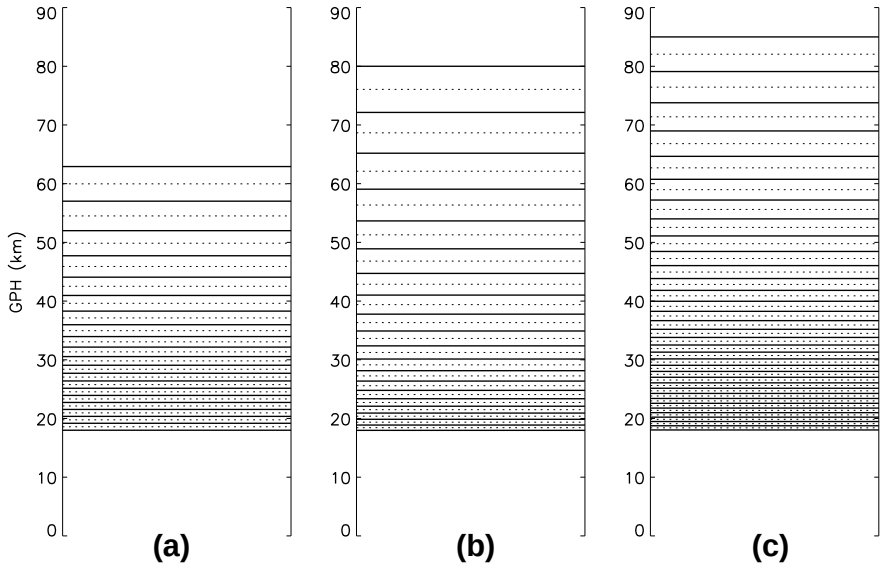

Figure 3.1: Vertical grid levels for (a) 50 level model; (b) 70 level model; (c) 85 level model. Solid lines correspond to theta levels while dashed lines correspond to rho levels.

boundary are purely terrain following, with a gradual transition to "flat" levels approaching the model lid. Full details of the vertical coordinate system and spatial discretisation of the model prognostic variables can be found in Staniforth et al. [2004]. The vertical model grid configurations for the L50 and L70 model configurations can be seen in Figure 3.1. Note only the flat theta and rho model levels of the middle atmosphere are shown in Figure 3.1, below which the model levels are terrain following. For the L50 grid configuration there are 30 terrain following levels in the troposphere and lower stratosphere, while the L70 configuration has an increased vertical resolution in this region with 50 terrain following levels.

Figure 3.1 also details the vertical resolution for the 85 level configuration of the UM. The L85 model configuration is currently used operationally for seasonal forecasting models at the Met Office and is detailed here since this configuration is applied in the UM trials discussed in Section 8. As for the L70 level, the L85 configuration has 50 terrain following levels in the troposphere and lower stratosphere, however compared to the L70 model it has increased resolution in

the upper middle atmosphere and extends to ~85 km.

3.1.4 Observations and Changes

Each day the Met Office receives hundreds of thousands of atmospheric observations from various sources. Each type of observation has it's own characteristics including resolution, systematic and random errors and the location and time of observation. Obviously the choice and availability of the observations used in the DA process will have an impact on the accuracy of the analyses produced.

Temperature observations are available from numerous sources including radiosondes, aircraft observations and satellite soundings. Below is a list of the sources of temperature observations used in creation of the stratospheric assimilated dataset from 1991 to the present.

- Aircraft observations which have vertical positions converted to pressure levels.
- "SATEM"reports, which are layer mean temperature soundings from the TIROS Operational Vertical Sounder (TOVS) [Smith et al., 1979] series of polar orbiting satellites operated by the National Oceanic and Atmospheric Administration (NOAA). These soundings have a horizontal resolution of ~500 km and a vertical range up to 1 hPa. They are produced by the National Environmental Satellite Data and Information Service (NESDIS).
- Soundings from the High Resolution Infrared Radiation Sounder (HIRS) [Koenig, 1980], also from the TOVS series of satellites. These soundings have a horizontal resolution of ~120 km and a vertical range up to 0.4 hPa.
- Radiosonde soundings available up to and between 50-10 hPa.
- The Advanced TOVS (ATOVS) [Reale et al., 2008] series of satellites NOAA-15, NOAA-16, NOAA-17 and NOAA-19 containing the following instruments.

  - The Advanced Very High Resolution Radiometer (AVHRR) [Cracknell, 1997].
  - An improved version of the HIRS.
  - The Advanced Microwave Sounding Unit (AMSU) [Aumann et al., 2003], which consists of three separate modules, AMSU-A1, AMSU-A2, and AMSU-B.

- The Atmospheric Infrared Sounder (AIRS) [Aumann et al., 2003]

- Global Positioning System Radio Occulation (GPSRO) temperature profiles from the CHAllenging Mini-Satellite Payload (CHAMP) [Wickert et al., 2004], Global navigation satellite systems Radio Occulation Receiver (GRAS) [Loiselet et al., 2000], Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) [Rocken et al., 2000] and the Gravity Recovery and Climate Experiment (GRACE) [Tapley and Reigber, 2002] missions.

- The Special Sensor Microwave Imager/Sounder (SSMIS) [Kunkee et al., 2008] instrument on board the Defence Meteorological Satellite Program (DMSP) satellite DMSP-16.

- The Infrared Atmospheric Sounding Interferometer (IASI) [Siméoni et al., 1997] on board the European Space Agency (ESA) polar orbiting satellite METOP.

Below is a list of all the observational changes in creating the stratospheric assimilated dataset from 1991 to the present. Note changes made at a particular date apply to all subsequent dates unless stated.

29th October 1991: Aircraft, SATEM reports, HIRS and radiosonde data profiles used.

18th April 2001: HIRS and SATEM reports no longer used. 3D-Var DA allows the direct assimilation of measured radiances. The Met Office calculates radiances using the fast Radiative Transfer for TOVS (RTTOV) model. Radiances from the following instruments on board NOAA-15 and NOAA-16 satellites are assimilated. AVHRR, improved HIRS, AMSU-A1, AMSU-A2, and AMSU-B.

10th October 2002: Radiances from NOAA-17 added to the system.

22th February 2005: AIRS radiances added to the system. Unfortunately there was a bug in the AIRS assimilation code. For the upper most pressure levels model radiance values were incorrectly calculated using background surface temperatures values. This lead to negative analyses increments and a cold temperature bias for the upper most levels, especially at low latitudes.

14th March 2006: Bugs in the AIRS code was fixed. RTTOV-v5 is upgraded to version 7, RTTOV-v7 [Saunders et al., 2002b].

20th September 2006: GPSRO CHAMP and GRACE data added to the system. These profiles have a vertical range from 4-27 km. Also added are SSMIS radiances from DMSP-16.

5th December 2006: GPSRO CHAMP data withdrawn from the system.

6th March 2007: Additional channels from HIRS on NOAA-17 and NOAA-16 added to the system.

15th May 2007: GPSRO COSMIC data added into the system.

7th November 2007: Radiances from IASI added to the system.

1st April 2008: The range of GPSRO data which is assimilated is extended to 0-40 km.

22nd July 2008:	GPSRO GRAS data added to the system. AIRS cloudy radiances are now assimilated.
13th July 2009:	Introduction of NOAA-19 data to the system.
2nd November 2010:	SSMIS data added to constrain the mesosphere.

3.1.5 Summary of Changes

A summary of all changes made in creation of the stratospheric assimilated dataset from 1991 to the present can be seen in Table I.

3.2 The Upper Atmosphere Research Satellite Microwave Limb Sounder

3.2.1 Background

The UARS MLS experiment was the first of the two MLS experiments. It is one of nine instruments on board the UARS satellite launched on the 12^{th} September 1991. Full scientific observations were obtained daily since the 29^{th} September 1991 to 15^{th} March 1994. After this date measurements became increasingly sparse to conserve the lifetime of the instrument. The final observations were obtained on 25^{th} August 2001. Temperature data products were retrieved from O_2 63 GHz radiance measurements using the technique documented in Waters et al. [1975]. All UARS temperature profiles in this project are version 5 of the data set.

Year	Day / Month	Model levels / Model lid	Dynamical Core	DA Scheme	GWD Scheme	Radiation Scheme	Observations
1991	29/11	42/0.28 hPa	Eulerian Dynamics	AC	Wilson and Swinbank	Ingram	Aircraft, SATEM, HIRS, Radiosonde
1992	25/02				>20 hPa Palmer <20 hPa Raleigh		
2000	16/05				>20 hPa Gregory <20 hPa Raleigh	Edwards-Slingo	
2000	14/11	40/0.1 hPA		Incremental 3D-Var			
2001	18/04						Aircraft, Radiosonde, AVHRR, HIRS, AMSU-A1, AMSU-A2
2003	10/10	50/0.1 hPa	Semi Lagragian Dynamics		>20 hPa Gregory USSP Added In		AMSU-B Data Now Used
2005	22/02						Bugged Airs Radiance Data Added
2006	14/03			Incremental 4D-Var	>20 hPa Gregory USSP turned off		Airs Radiance Bugged Fixed
2006	20/09						GPSRO CHAMP and GRACE Data Added
2006	05/12				>20 hPa Gregory USSP Added Back In		GPSRO CHAMP Withdrawn
2007	15/05						GPSRO COSMIC Data Added
2007	07/11						IASI Radiance Data Added
2008	22/11						GPRSO GRAS Data Added
2009	13/07						NOAA-19 Data Added
2009	10/11	70/0.01 hPa					
2010	2/11						SSMIS data added

Table I: Summary of changes made to Unified Model, Data Assimilation Scheme and Observations from 1991 to present.

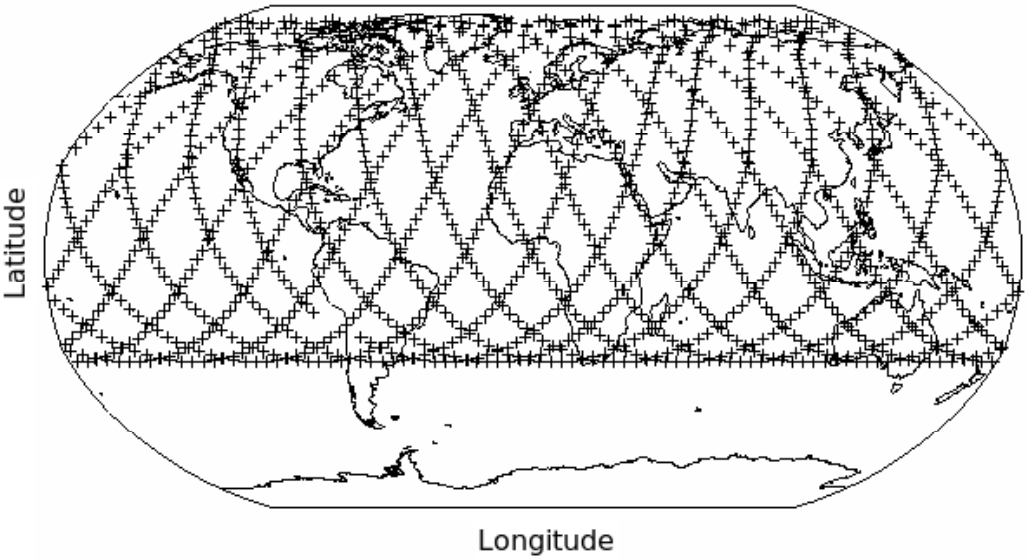

Figure 3.2: UARS MLS tangent point locations (crosses) for a 24-h period of coverage in the north viewing mode (80^oN to 34^oS).

3.2.2 Orbit and Coverage

The UARS satellite has an inclined 57^o orbit with a height of ~585 km. As the MLS instrument measures radiances at 90^o to the motion of the UARS satellite, the latitude coverage is from 34^o in one hemisphere to 80^o in the opposite hemisphere. Every 36 days the UARS satellite performs a 180^o yaw manoeuvre where the latitude coverage of opposite hemispheres is interchanged. Each limb scan takes ~65.5 s with profile tangent point loci separated by 4^o latitude. There are a total of 16 orbits per day, each taking ~90 minutes, resulting in a total of ~1300 vertical scans per day.

The UARS MLS vertical coverage of temperature is scientifically useful from ~32 hPa to ~0.46 hPa ranging from the middle stratosphere to the lower mesosphere. Temperatures are reported at an interval of six surfaces per decade change in pressure. Figure 3.2.2 shows the tangent point locations of the vertical limb scans for one 24-h period of UARS coverage.

3.2.3 Spatial Resolution

As for the EOS MLS instrument the vertical resolution of the UARS MLS can be estimated from the horizontal integration of 2D averaging kernels, see equation (9) of [Livesey et al., 2003]. The following averaged kernel widths are taken from a typical profile of the instrument obtained on September 17^{th} 1992. This is 6.5 km at 32 hPa, 6.0 km at 15 hPa, 6.5 km at 1 hPa and 5 km at 0.46 hPa. These values have been scaled from log pressure into kilometres using a scale height of 16 km per decade change in pressure. The horizontal along track resolution, distance between each successive vertical profile, is ~ 495 Km. The across track horizontal resolution is controlled by the antenna's FOV, and is ~9.5 km for the 63 GHz O_2 measurements.

3.2.4 Sources of Error

Following Fishbien et al. [1996] the sources of error contributing to all UARS data can be decomposed into random and systematic parts depending on the timescale of their variability. Error sources which vary on the smallest timescales are random and contribute only to the precision of the data. Examples of these sources include measurement noise, radiometric calibration and forward model errors. A detailed account of all random errors contributing to precision of UARS data can be seen in Fishbien et al. [1996]. Systematic error sources are defined as those which may vary over any time scale. These sources include smoothing errors, which occur from the inaccuracies of the averaging kernels in reproducing the true atmospheric state, and model parameter errors occurring from inaccuracies of instrument calibration, spectroscopic parameters and approximations used in the forward model. The primary sources of systematic error are due to the inaccuracies of the O_2 line broadening parameter and forward model errors. A detailed account of all systematic errors contributing to the accuracy of UARS temperature data can be seen in Fishbien et al. [1996].

3.2.5 Validation with Correlative Data Sets

Previously Livesey et al. [2003] validated the version 5 temperature profiles against the National Centers for Environmental Prediction (NCEP) Global Data Assimilation System (GDAS) stratospheric analysis data set [McPherson et al., 1979]. Here comparisons were made between vertical profiles of both data sets for the first year of UARS data. In addition Wu et al. [2003] validated UARS version 5 retrieved temperature profiles against the COSPAR International Reference Atmosphere 1986 (CIRA'86) data [Fleming et al., 1990] and the Improved Stratospheric and Mesospheric Sounder (ISAMS) [Rodgers et al., 1996], High Resolution Doppler Imager (HRDI) [Hays et al., 1993], Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere-1 (CRISTA-1) [Grossmann et al., 2004] satellite temperature profiles and the Rayleigh lidar measurements from seven Network for the Detection of Atmosphere Composition Change (NDACC) observational sites. The biases observed in the above studies are discussed below.

In the lower stratosphere UARS MLS data has a cold bias with magnitudes generally below ~4 K. In the upper stratosphere UARS MLS data has a warm bias with magnitudes generally below ~2 K, except with respect to NCEP analyses where warm bias magnitudes are between ~4→9 K. In the lower mesosphere UARS has both warm and cold biases with magnitudes less than ~6 K.

3.3 The Earth Observing System Microwave Limb Sounder

3.3.1 Background

The EOS MLS experiment is the second of the two MLS experiments. It is one of four instruments on board the National Aeronautics and Space Administration (NASA) Aura satellite. Launched on the 12^{th} of July 2004 EOS MLS has been obtaining full scientific observations from the 13^{th} August 2004 to date. The

EOS MLS data products used in this project are global profiles of atmospheric temperature produced at a one day granularity, from midnight to midnight universal time. Temperature data products are obtained from observed thermal emissions near the 118 GHz O_2 and 240 GHz $O^{18}O$ spectral lines, using the retrieval algorithm described in Livesey et al. [2006].

3.3.2 Orbit and Coverage

The Aura satellite is part of the NASA "A-Train" group and has an ascending-node 705 km sun-synchronous polar orbit with a 98^0 inclination. The MLS instrument field of view (FOV) is in the direction of the orbital motion producing vertical scans of the Earth's limb in the orbit plane. The latitude coverage is from 82^0S to 82^0N with each limb scan synchronised to the orbit track. In total there are ~14.5 complete orbits for each day with a total of 240 scans per orbit. This results in ~3500 different observational site vertical profile scans in 24 h, one granule. All vertical scans in each orbit are phased so that at each equator crossing the limb scan tangent point loci occur directly on the equator. This results in vertical profiles which are one orbit apart having approximately the same latitude and being separated by ~100 minutes.

The MLS vertical coverage for temperature is from ~5 km to ~90 km ranging from the upper troposphere into the upper mesosphere. Figure 3.3 shows the shows the locations of the vertical limb scans for one 24-h period of EOSMLS coverage.

3.3.3 Spatial Resolution

The EOS MLS retrieval algorithms operates in a 2-D manner [Livesey and Read, 2000], with dimensions in the along-track orbit and vertical directions. This approach allows the spatial resolution of the MLS instrument to be estimated from 2-D averaging kernels [Rodgers, 1990, 2000]. Horizontal along track integration of the 2-D kernels results in vertical averaging kernels which describe the vertical

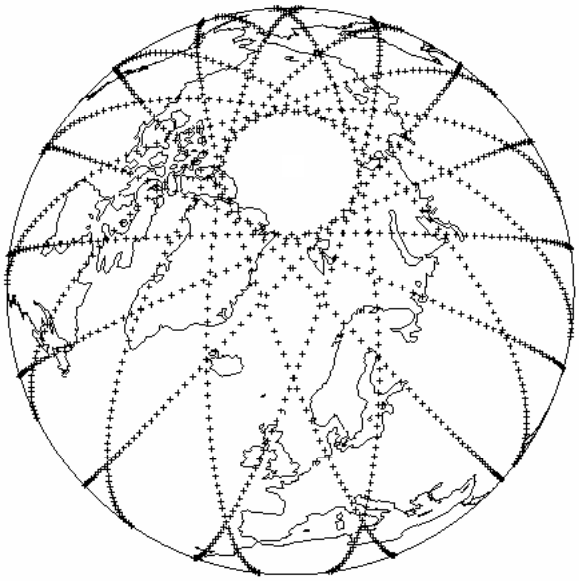

Figure 3.3: EOS MLS measurement locations (crosses) for a 24-h period of coverage

resolution of the MLS instrument at each retrieval level. The vertical resolution is taken to be the thickness of the atmospheric layer contributing to the retrieved value at half the maximum kernel value for that retrieval, otherwise know as full width half maximum (FWHM). This is 5.3 km at 316 hPa, 5.2 km at 100 hPa, 3.5 km at 32 hPa, 4 km at 10 hPa, 8 km at 1 hPa, 9 km at 0.1 hPa and 14 km at 0.01 hPa. Vertical integration of the 2-D kernels results in horizontal kernels which describe the horizontal along-track resolution of the MLS instrument. The FWHM value of the horizontal kernel estimates the horizontal along-track resolution as ~165 km. The cross-track horizontal resolution is controlled by the horizontal width of the MLS FOV. For the 240 GHz radiometer this is ~6 km and for the 118 GHz radiometer it is ~12 km.

3.3.4 Sources of Error

Systematic uncertainties arise in MLS data products from instrumental issues such as radiometric uncertainty, FOV calibration and approximations in the re-

trieval implementation.

The largest single contribution to MLS systematic errors is caused by Gain Compression (GC). This error is induced by observed spectral signals being compressed when viewed against the atmospheric background temperature scene of ~300 K compared to 0 K. The magnitude of this compression is ~1.5%. Although errors caused by GC were only recognised late in the development of the v2.2 algorithm their magnitude and sign can still be estimated through a model based upon laboratory measurements of MLS flight-hardware-spare intermediate amplifiers. GC error causes vertical oscillations for MLS retrieved temperatures between 316 hPa and 10 hPa and a 1-3 K high bias for levels above 10 hPa [Schwartz et al., 2007].

Apart from GC the largest contributions of systematic error come from the radiometric measurement system and the forward model used for retrieval. While the magnitude of each source of systematic error can be calculated, unlike GC the sign of these biases is unknown. The overall bias uncertainty caused by all systematic errors, apart from GC, is 5 K at 100 hPa, 2-2.5 K from 100 hPa to 0.01 hPa and 3 K at 0.001 hPa.

3.3.5 Validation with Correlative Data Sets

Validation of EOS MLS temperature data was performed in Schwartz et al. [2007]. This study computes the zonal mean difference, and standard deviation about this mean, between selected days of v2.2 EOS MLS temperature data and the following independent correlative data sets, assimilated temperature analyses from both the Goddard Earth Observing System (GEOS-5) [Reinecker et al., 2007] and the European Centre for Medium-Range Weather Forecasts (ECMWF) [Simmons et al., 2005] global models, temperature observations from various instruments including the CHAllenging Minisatellite Payload (CHAMP) [Wickert et al., 2004], Atmospheric Infrared Sounder (AIRS) [Aumann et al., 2003],

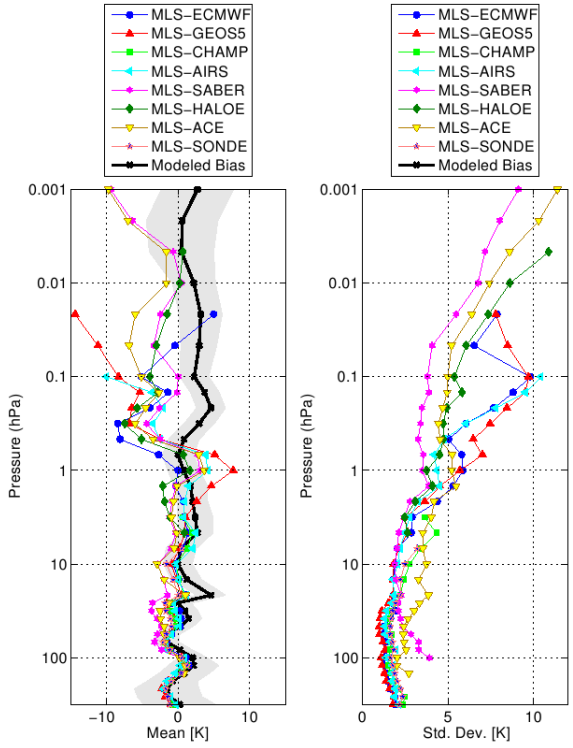

Figure 3.4: Summary of temperature biases and scatter between EOS MLS and eight correlative data sets. Mean differences are shown in the left panel and scatter about these differences in the right. The modelled systematic uncertainties caused by gain compression is shown in black. The contribution of remaining biases of unknown sign are shown by the grey envelope.

Sounding of the Atmosphere using Broadband Radiometry (SABER) [Russell III et al., 1999], Halogen Occulation Experiment (HALOE) [Russell et al., 1993], Atmospheric Chemistry Experiment (ACE) [Bernath et al., 2004] satellites and the global radiosonde network [Durre et al., 2006]. Figure 26 of Schwartz et al. [2007], shown as Figure 3.4, details a summary of the global mean EOS MLS temperature bias relative to the above eight correlative data sets.

From Figure 3.4 we find that in the lower stratosphere EOS MLS data generally has a cold bias with magnitudes no greater than ~3 K. Here biases have an oscillating vertical structure which is explained and well predicted by the GC model, except at 21.6 hPa where a predicted warm bias of ~4 K is not observed.

In the upper stratosphere EOS MLS biases are less than ~2 K, except when compared to GEOS-5 where warm bias of up to ~8 K occur near 1 hPa. In the lower mesosphere all biases are cold with magnitudes of ~0→8 K. The magnitude of this general cold bias in the lower mesosphere can not be explained by known sources of systematic uncertainty. In the upper mesosphere the majority of biases are again cold with magnitudes of ~0→8 K, except a cold bias of ~13 K w.r.t to GEOS-5 data and a warm bias of ~5 K w.r.t to ECMWF analyses at ~0.03 hPa.

3.4 Sounding of the Atmosphere using Broadband Emission Radiometry

3.4.1 Background

The SABER experiment is one of four instruments on board NASA's Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics (TIMED) satellite. Launched on the 7^{th} of December 2001 SABER has been obtaining full scientific observations from the 22^{nd} of January 2002 to date. The main SABER data products used in this project are global profiles of atmospheric temperature produced on a one day granularity, from midnight to midnight universal time. SABER temperature data are obtained from thermal emissions near the 15 μm CO_2 spectral channel, using retrieval concepts similar to those developed for the Nimbus 7 Limb Infrared Monitor of the Stratosphere (LIMS) experiment [Gille and Russell, 1984]. Below ~65 km the SABER retrieval algorithm uses a forward radiance model based on local thermodynamic equilibrium (LTE) assumptions [Remsberg et al., 2004]. Above ~65 km a second algorithm is used which includes additional effects of radiation and chemical reactions, i.e based on non-LTE assumptions [Remsberg et al., 2008]. Full technical details of the SABER instrument can be seen in Russell III et al. [1999]. All SABER temperature data used in this project are version 1.07.

3.4.2 Orbit and Coverage

The TIMED satellite has a near sun-synchronous orbit with 71.4^o inclination and a height of ~625 km. The SABER instrument FOV is 90^o to the velocity vector of the TIMED satellite producing vertical scans. The SABER instrument has two modes of viewing, north and south. In north viewing mode the tangent point locations of the vertical scans vary from 83^{o}N to 52^{o}S. This mode occurs for 60-63 days after which a yaw manoeuvre turns the instrument to it's south viewing mode, where the tangent point locations of the vertical scans vary from 83^{o}S to 52^{o}N. The south viewing mode again occurs for 60-63 days after which the viewing sequence is repeated. This results in only latitudes from 52^{o}N to 52^{o}S having continuous coverage. Vertical profiles are taken every ~58 seconds with an along track profile spacing of ~ 3^o. There are a total of 15 orbits each day, each ~97 minutes long, resulting in a total of ~1500 vertical scans per day. The positions of observational profile locations are distributed uniformly along longitude circles and yaw manoeuvre's occur on nearly the same dates for succeeding years.

The SABER vertical coverage for temperature is from ~10-110 km ranging from the upper troposphere to the lower thermosphere. Figure 3.4.2 shows the tangent point locations of the vertical limb scans for one 24-h period of SABER coverage.

3.4.3 Spatial Resolution

The SABER instrument has a single axis scan mirror that measures radiances emitted by the atmosphere along a ray path. Each ray path is identified by it's tangent height, defined as the altitude of the point of closest approach to the surface of the Earth. The SABER instrument has a high vertical resolution of ~2 km throughout it's vertical range. The along track horizontal resolution, here defined as the spacing between each successive profile, is ~400 km. The cross-track resolution is controlled by the width of the SABER FOV. For the infrared broadband radiometer measuring the narrow and wide CO_2 15 μm channels this is ~1.68 and ~1.49 km respectively.

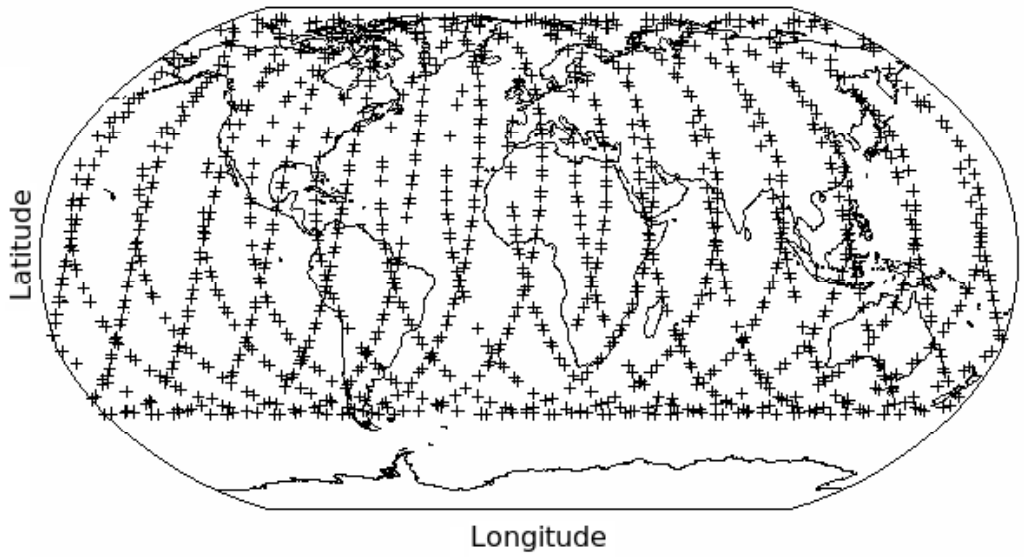

Figure 3.5: SABER tangent point locations (crosses) for a 24-h period of coverage in the north viewing mode (83^oN to 53^oS).

3.4.4 Sources of Error

The primary source of random error in the LTE algorithm is caused by high frequency spacecraft vibrations, "jitter", that nominally leads to profile registration errors of ±20 m. The secondary source of random error is noise from the SABER radiance detector. Values of jitter error are of the order 0.3 K in the lower stratosphere and increase to 0.6 K in the lower mesosphere. In the mid-mesosphere the values of jitter and detector noise error are of the same order. The primary source of systematic error in the LTE algorithm is from errors in the measured radiance values. These value are of the order 1.0 K in the lower stratosphere and positive. In the upper stratosphere and lower mesosphere these errors can increase to 1.6 K and are negative. The secondary source of systematic error is in the determination of the reference pressure used in pressure registration of calculated radiance profiles.

Additional uncertainties dominate above ~65 km when using retrievals based on non-LTE assumptions. Since CO_2 is known not to be well mixed above ~75 km

[López-Puertas et al., 1998], temperature retrievals using the non-LTE algorithm will depend on prior information about the decrease of CO_2 with altitude and it's uncertainty [Mertens, 2001]. Uncertainties in parameters used for the computation of vibrational temperatures in the non-LTE algorithm also contribute to uncertainties in retrieved temperatures. The most important of these parameters is the rate that CO_2 vibrations are quenched by collisions with atomic oxygen, which is determined by the rate coefficient, k_o, in equation (1) of Mertens [2001].

3.4.5 Validation with Correlative Data Sets

The work of Remsberg et al. [2008] provides a detailed account of the validation of SABER v1.07 temperature retrievals. This account includes the results from comparing SABER profiles with the following data sets, Rayleigh lidar measurements from two Network for the Detection of Atmosphere Composition Change (NDACC) observational sites, temperature profiles from the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) [Fischer et al., 2007], Spectral Airglow Temperature Imager (SATI) [Sargoytchev et al., 2004], EOS MLS, AIRS and HALOE experiments and climatological temperature fields from falling spheres measurements detailed in Lübken [1999]. Discussion of the above validation follows.

In the lower stratosphere SABER data has a warm bias with magnitudes generally smaller than ~4 K. In the upper stratosphere SABER data has a cold bias with magnitudes ranging from ~1→5 K, except w.r.t to MIPAS data where warm biases of ~2 K are observed. In the lower mesosphere SABER data has a cold bias with magnitudes again ranging from ~1→5 K, except w.r.t EOS MLS profiles where a warm bias of ~3→5 K from ~1.0-0.3 hPa is observed. In the upper mesosphere SABER data has both cold and warm biases with magnitudes ranging from ~5 K and ~9 K respectively.

3.5 Summary

Analyses Summary

This chapter has described in detail the changes made to the dynamical model, assimilation scheme and observations used to create the stratospheric assimilated data set. Basic formulations of the AC, 3D-Var and 4D-Var assimilation schemes were presented as-well as a basic overview of the Met Office assimilation cycle. The changes made to the system which have possibly the largest impact on the analyses include.

- Radiation scheme changed to the Edwards-Slingo scheme in May 2000.
- Introduction of 4D-Var assimilation in March 2003.
- Horizontal grid refinement when switched to the operational stratospheric model in March 2006.
- Dynamical model changed to the new dynamics in October 2003.
- Correction of the bugged AIRS data assimilation in March 2006.
- Introduction of the USSP scheme in October 2003, subsequently removed in March 2006 and re-introduced back into the system in December 2006.

Satellite Data Summary

For the three independent satellite data sets details of the vertical resolution, measurements technique, coverage and spatial resolution are described. Also discussed are the main sources of random and systematic errors which describe the precision and accuracy respectively for each instrument. Also discussed is a separate second estimation of each instruments accuracy, calculated from the observed temperature biases between each instrument and profiles from correlative data sets. A comparison of EOS MLS, UARS MLS and SABER coverage, vertical resolution and observed temperature biases throughout the middle atmosphere is given in Table II.

	EOS MLS	UARS MLS	SABER
Coverage	Global	34^oN(S) to 80^oS(N)	34^oN(S) to 80^oS(N)
Vert. Resolution Mesosphere	~8-9	~ 5	~2
Vert. Resolution Stratosphere	~3.5-5.3	~ 6-6.5	~2
	Temperature Biases		
Upper Mesosphere ~0.01 hPa	-2...0	N/A	-5...+9
Middle Mesosphere ~0.1 hPa	-8...0	N/A	-5...+5
Lower Mesosphere ~0.5 hPa	-7...+4	-4...+6	-5...+5
Stratopause ~ 1.0 hPa	0...+5	-4...+8	-5...+5
Middle stratosphere ~10 hPa	-1...0	0...+8	-5...+3
Lower stratosphere ~100 hPa	0...+1	N/A	+1...+4

Table II: Comparison of EOS MLS, UARS MLS and SABER coverage, vertical resolution (km) and observed temperature biases (K) in the stratosphere and mesophere.

Chapter 4

Zonal Mean Temperature

This chapter discusses the zonal mean temperature fields of the stratospheric assimilated data set. Climatologies of the analyses are constructed for selected periods and then compared to observed fields. For L50 and L70 analyses the zonal mean temperature biases with both EOS MLS and SABER data are presented. For L42 analyses the zonal mean temperature biases with UARS MLS data are also presented. It should be noted that UARS MLS, EOS MLS and SABER data are not assimilated into the analyses fields, and as such they are independent observations. Systematic biases are detailed for all analyses and probable causes are diagnosed.

4.1 L50 Analyses Climatology

Figure 4.1 shows the monthly zonal mean temperature fields of the L50 analyses, averaged over 5 years 2005-2009 for January to October and 4 years 2005-2008 for November to December. For qualitative comparison the zonal mean temperature fields of the EOS MLS and SABER instruments averaged over the same period and the SPARC climatological fields are also shown.

The EOS MLS data are reported on a pressure grid with 35 pressure levels which has 12 intervals per decade change in pressure from 1000-100 hPa, 9 intervals per decade change in pressure from 100-10 hPa, 6 intervals per decade change

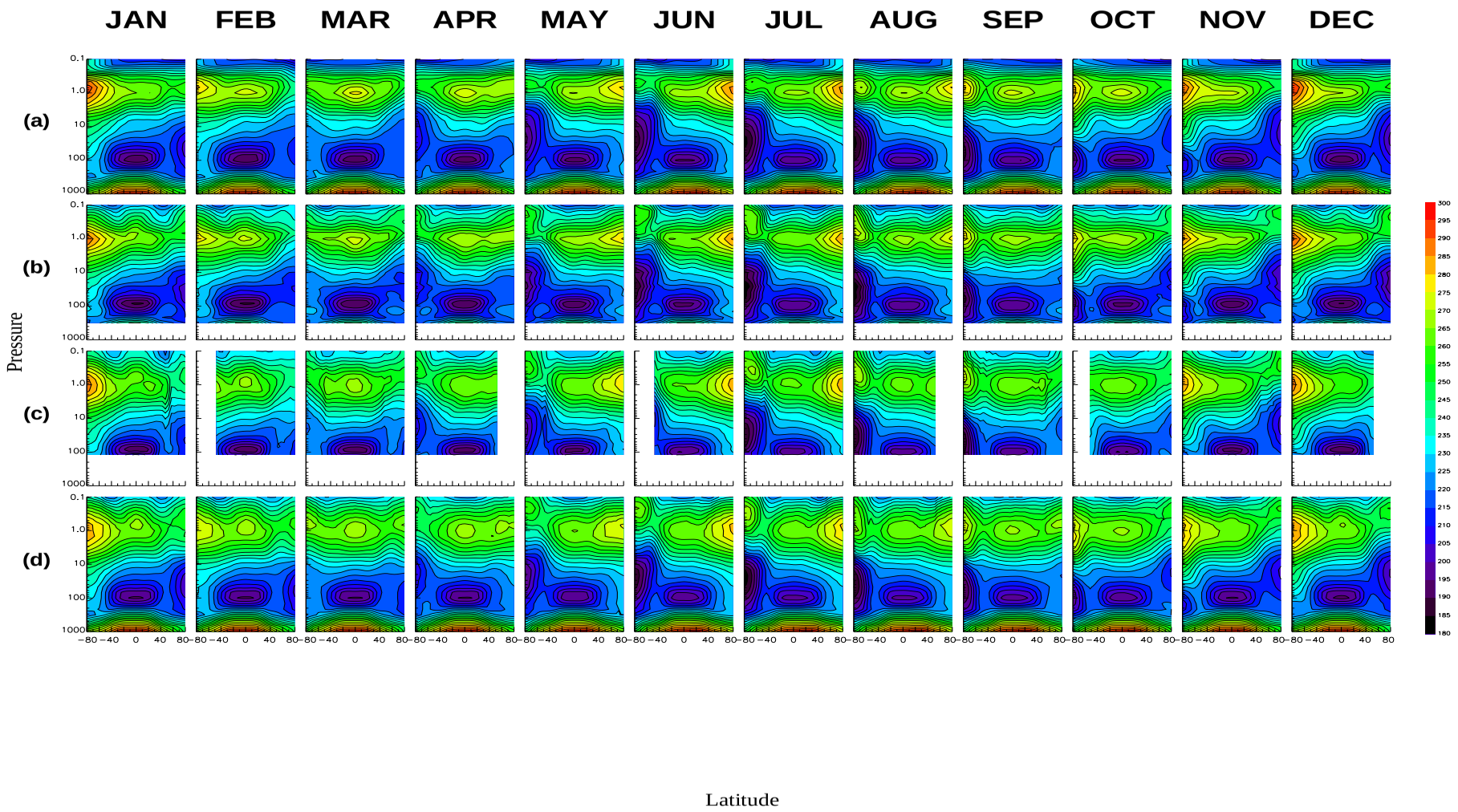

Figure 4.1: Climatological monthly zonal mean temperature fields from 2005-2009 for (a) Met Office L50 analyses; (b) EOS MLS v2.2 dataset; (c) SABER v1.07 dataset; (d) SPARC 2004 climatological dataset. Contour interval is 10 K.

in pressure from 10-0.1 hPa and 3 intervals per decade change in pressure from 0.1-0.01 hPa. The EOS MLS data has been placed into 5^o latitudinal bins. The SABER data was placed into identical latitudinal bins as the EOS MLS data and vertically into bins which have 33 pressure levels with 12 intervals per decade change in pressure from 1000-100 hPa and 6 intervals per decade from 100-0.1. The SPARC zonal mean temperatures are reported on a grid which has 33 pressure levels from 1000-0.0046 hPa and a latitudinal resolution of 4^o [Randel et al., 2004].

Figure 4.1 (b) shows that the EOS MLS data has a steep vertical temperature gradient between ~1-0.46 hPa when compared to the SABER data. This persistent feature is due to the consistent EOS MLS warm bias at 1 hPa rapidly changing sign to a relatively strong cold bias at 0.46 hPa, as detailed in Figure 3.4.

Figure 4.1 (c) shows that for SABER months which have near global latitudinal coverage, due to the presence of a yaw manoeuvre, discontinuities in the temperature distribution occur at ~53^o degrees of each hemisphere. These discontinuities arise due to certain latitude regions having temperature values averaged over different time periods. Between ~53^o south and ~53^o north of each hemisphere the temperatures are averaged over the complete month, while poleward of ~53^o temperatures are averaged over a smaller subsection of days in the month determined by the date of the yaw manoeuvre. This explanation is verified in Figure 4.2, which shows smooth latitudinal temperature gradients when the SABER results for January in Figure 4.1 (c) are decomposed into separate south and north facing modes.

Stratosphere

Qualitatively the L50 analyses results from Figure 4.1 show that throughout

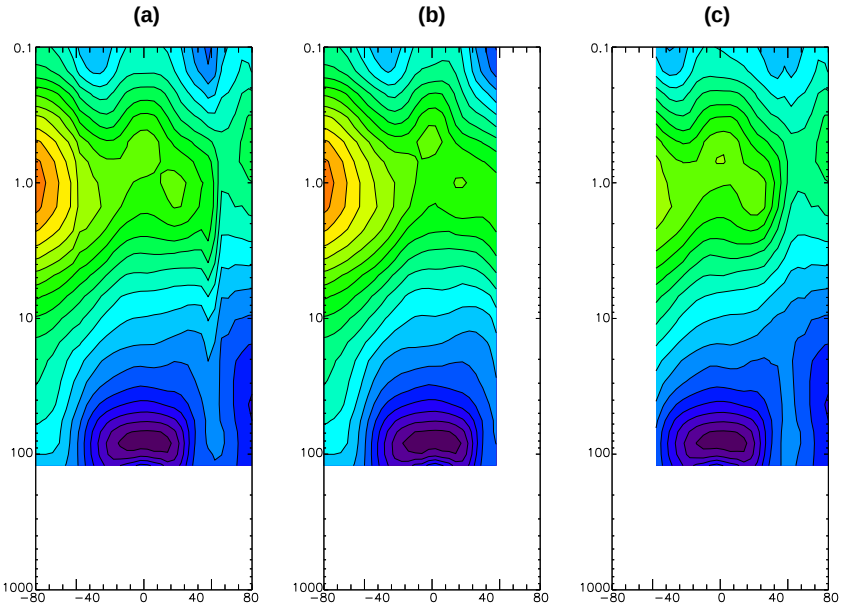

Figure 4.2: Decomposition of SABER monthly zonal mean temperature for January averaged from 2004-2008. (a) Zonal average over full month; (b) Zonal average for south facing mode; (c) Zonal average for north facing mode. Contour interval same as in Figure 4.1.

the stratosphere the Met Office analyses accurately reproduce the main seasonal features of the atmosphere temperature distribution discussed in Section 2.2.1. Both equinox and solstice months of each hemisphere have latitudinal temperature gradients that are in qualitative agreement with EOS MLS and SABER values.

A poor aspect of the analyses reproduction of stratospheric temperatures is the location of the separated polar winter stratopause discussed in Section 2.2. For the NH winter months the polar stratopause is located at ~0.9→0.8 hPa, compared to the observed and climatological values of ~0.3→0.2 hPa. In the SH winter months the analyses representation is slightly better, with the polar stratopause located at ~0.7→0.5 hPa compared to the observed and climatological values of 0.3→0.15 hPa.

Lower Mesosphere

Qualitatively speaking the representation of the lower mesosphere is considerably worse than that of the stratosphere. The L50 analyses of Figure 4.1 (a) generally have the correct characteristic decrease in temperature with altitude for the mesosphere above ~1.0 hPa. However compared to Figures 4.1 (b) and (c) there are excessively high vertical temperature gradients from ~0.5→0.4 hPa, which result in analyses temperatures generally being ~20 K colder than observed values from ~0.3→0.1 hPa. These excessive vertical temperature gradients are present across all latitudes, however are greatest above the summer middle to high latitude stratopause. Another aspect of the poor mesospheric representation is the small latitudinal variation of temperature from ~0.5→0.1 hPa in the analyses compared to observed and climatological values, most prominent at high latitudes in both winter hemispheres from ~0.5→0.1 hPa. Such deficiencies result from the poor representation of the separated winter stratopause described above.

4.2 Validation Procedure

The range of vertical temperature profiles used in the validation was restricted to a 3-hour window either side of the daily analysis time of 1200 UTC. For each observational profile in this temporal range, temperature values from the corresponding analysis grid region were horizontally averaged to the profile location. These horizontally averaged analysis values are then vertically interpolated to the nearest observation pressure level, with the differences between analyses and observed values binned according to latitude and pressure values and monthly averaged.

Profiles of monthly mean zonal mean temperature differences were obtained with the profile time window extended to 6-hours either side of the daily anal-

ysis time for selected months, results not shown. The differences between the 3-hour and 6-hour results were found to be negligible. We therefore assume that all information on the differences in the zonal temperature structure is captured by the 3-hour window of the MLS and SABER profiles.

4.3 L50 Monthly Zonal Mean Temperature Error

Figures 4.3 and 4.4 show the monthly mean zonal mean temperature differences between L50 analyses values and respectively the EOS MLS and SABER data for months between January 2005 and October 2009. Since the results are calculated as analyses values minus MLS or SABER data, blue contour levels correspond to the analyses having a cold temperature bias while red contour levels correspond to the analyses having a warm temperature bias.

For a direct comparison of EOS MLS and SABER validation results Figure 4.5 shows the monthly mean zonal mean temperature errors of Figure 4.3 minus those from Figure 4.4; this can be interpreted as SABER minus EOS MLS temperatures. Here blue contour levels correspond to SABER data being colder than EOS MLS and red contour levels correspond to SABER data being warmer than EOS MLS. Note that above ~1.0 hPa the values of SABER minus EOS MLS profiles and the bias estimates of both data sets discussed in Sections 3.3.5 and 3.4.5 are typically smaller than the analyses biases, described in detail below. Hence we can confidently state that for the majority of months the bias signal above ~1.0 hPa discussed below are predominantly due to the inaccuracies of the analyses fields, not biases of the independent observational data. The noticeable exception to this is for January 2008 where from Figure 4.5 we find that the differences between SABER and EOS MLS validation results from ~1.0→0.4 hPa are noticeably larger than all for all other months, reducing confidence in the fact that the biases present in the validation results are primarily due to inaccuracies of the analyses fields for this month.

When detailing biases of the L50 analyses from January 2005 to October 2009 a comparison is made between analyses pre and post March 2006. This date represents when the stratospheric and tropospheric global models were merged as detailed in Section 3.1.2. This comparison is made due to the large number of changes made to the system at this date including an increase of horizontal resolution, introduction of 4D-var assimilation, correction to the bugged AIRS assimilation code and the upgrading of RTTOV-v5 to RTTOV-v7.

4.3.1 Lower Mesosphere Cold Bias

Figures 4.3 and 4.4 both show that the largest zonal mean temperature error in the analyses is a cold bias that occurs in the uppermost model levels across all latitude values. This cold bias is predominantly confined between ~0.3→0.1 hPa, however it does extend down to ~0.8 hPa for specific months and latitudes discussed below.

Tropical Latitudes

Pre March 2006 From ~0.2→0.1 hPa maximum cold bias values of 22→30 K for the months of December, January and February (DJF) are larger than the 14→22 K values seen for all other months. For all months the above maximum cold bias values decrease in magnitude down to ~0.8→0.9 hPa.

Post March 2006 From ~0.2→0.1 hPa maximum cold bias values are generally lower than corresponding months pre March 2006. Maximum values of 14→18 K for DJF are again larger than the maximum values of 10→14 K seen for all other months. As for months pre March 2006 the maximum values decrease in magnitude for lower pressure levels. However now cold bias regions only extend down to ~0.4→0.3 hPa. The restriction of this cold bias region to above ~0.4→0.3 hPa for months post March 2006 is assumed to be a direct effect of the corrections made in the AIRS data assimilation code discussed in Section 3.1.4.

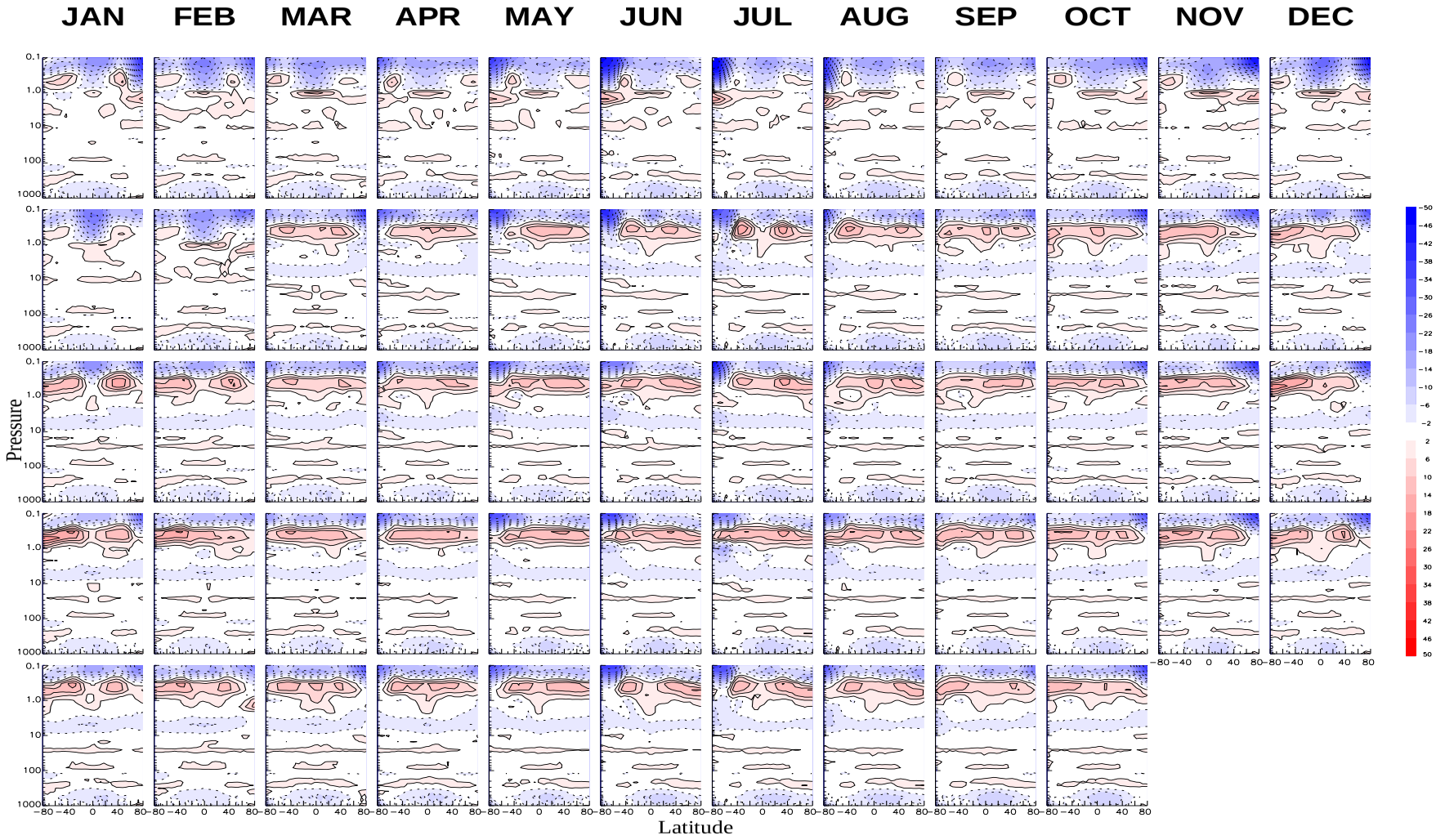

Figure 4.3: Monthly zonal average differences between L50 analyses temperatures and vertical temperature retrievals from the EOS MLS instrument from January 2005 to October 2009. Contour interval is 4 K. Negative biases are indicated by dashed contours.

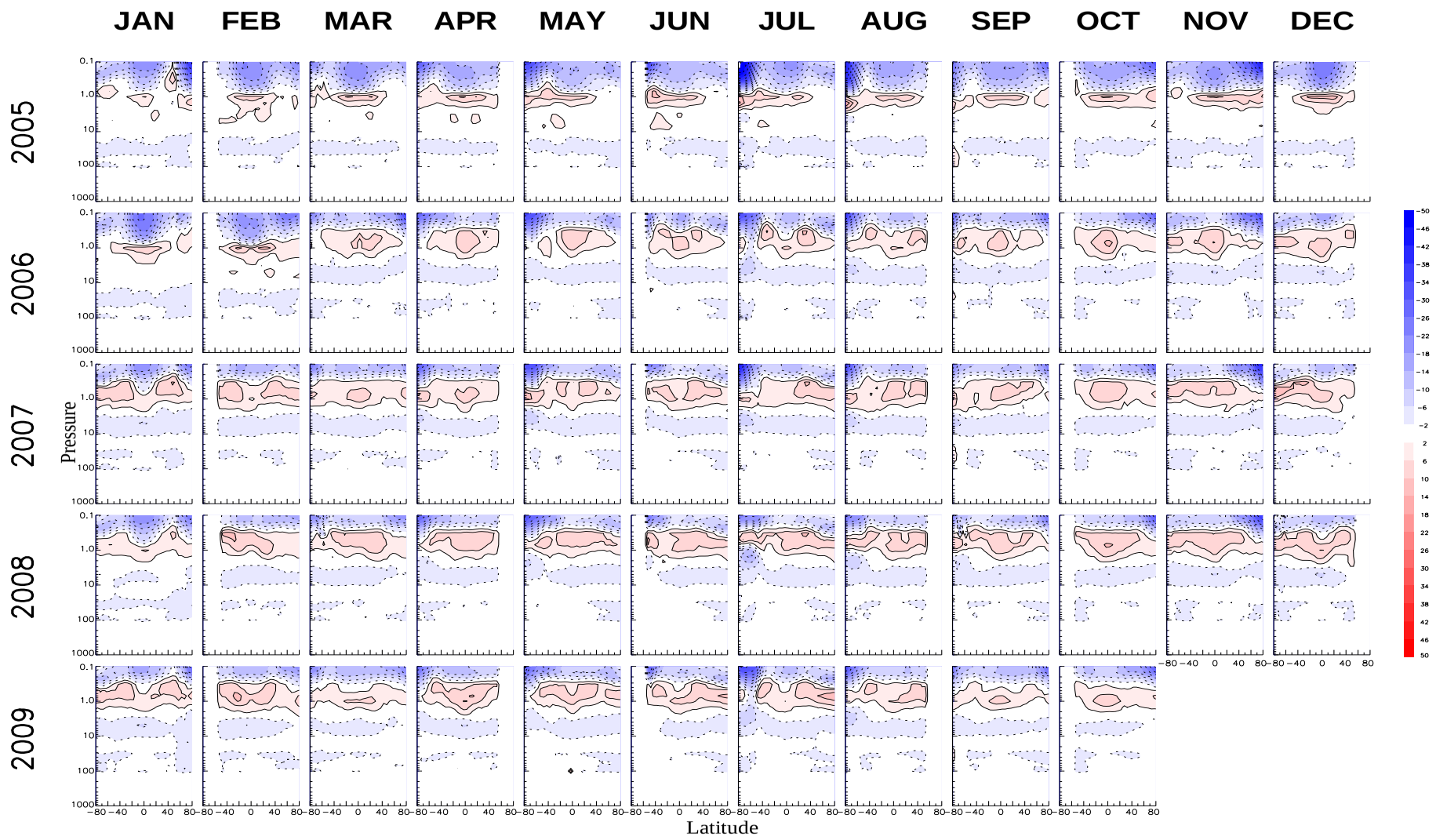

Figure 4.4: Monthly zonal average differences between L50 analyses temperatures and vertical temperature retrievals from the SABER instrument from January 2005 to October 2009. Contour interval same as Figure 4.3.

Figure 4.5: Results of Figure 4.4 minus results of Figure 4.3. Contour interval same as Figure 4.3.

For tropical latitudes from ~0.2→0.1 hPa validation results using EOS MLS and SABER data are qualitatively similar, consistent with the previous work of Schwartz et al. [2007] which reports a bias between EOS MLS and SABER of less than 2 K for this latitude and pressure range.

Equinox Polar Latitudes

Unlike tropical latitudes, there is relatively little difference between the cold polar biases of both hemispheres for equinox months pre and post March 2006. This smaller difference is expected as the impact of the bug in the AIRS assimilation code is reduced for polar latitudes, where the difference between surface and upper model level temperatures are smaller when compared to those for tropical latitudes. For all months bias maxima occur at ~0.1 hPa and typically decrease with altitude until ~0.3→0.4 hPa, except in the SH of September 2005 which extends down to ~1.0 hPa.

Validation results using EOS MLS data show that for both March and September NH biases values are primarily between ~18→22 K, and typically ~4→8 K larger than SH values of ~14→18 K. The most noticeable exception occurs in March 2006, where NH bias values of ~38 K are ~20 K larger than those seen in the SH.

Validation results using SABER data are noticeably different than those described above using EOS MLS data. With respect to SABER bias values for both hemispheres in March and September are now qualitatively similar, with typical magnitudes of ~22→26 K. The difference between the EOS MLS and SABER results is most noticeable in the NH of March, where SABER temperatures are typically ~4 K colder than EOS MLS, and in the SH of March and September, where SABER temperatures are typically ~4 K warmer than EOS MLS, see Figure 4.5. This difference in polar temperatures is qualitatively consistent with the biases between EOS MLS and SABER data reported in Figure 22 of Schwartz et al. [2007]. However magnitudes of the polar differences reported in Figure 4.5

for March and September are ~2→4 K larger than those of Schwartz et al. [2007], which use profiles averaged over 3 months.

Winter Polar Latitudes

Figures 4.3 and 4.4 show that there is a clear seasonal cycle for cold biases seen at high latitudes in the lower mesosphere. Here larger cold biases are seen in the winter polar regions of each hemisphere when compared to those from equinox and summer seasons.

Pre March 2006 Cold bias regions of both hemispheres have maximum values from ~0.2→0.1 hPa which typically extend and decrease in value until ~1.0→0.9 hPa.Validation results using EOS MLS show there is typically a cold bias of ~38→42 K for NH winter months, except for January 2006 where bias values are ~18 K. Biases with respect to EOS MLS data for SH winter months are typically ~4→8 K larger than those seen in the NH winter, with magnitudes of ~42→50 K. Validation results using SABER data are consistent with those from EOS MLS for SH winter months, however are typically ~4 K smaller for NH months. This difference between NH biases is again consistent with the biases between EOS MLS and SABER data for this latitude and pressure range reported in Schwartz et al. [2006].

Post March 2006 Maximum cold biases from ~0.2→0.1 for winter polar regions follow an identical pattern to those seen in months pre March 2006. However the values of these cold biases are lower than those seen for months pre March 2006 and generally only extend down to ~0.5 hPa. Biases for NH winter months using EOS MLS data are now ~30→38 K, except for January 2007/2009 where values ~22→26 are reported. Biases with respect to EOS MLS data for the SH winter months are typically ~4→12 K smaller than months pre March 2006, with magnitudes of ~30→42 K similar to those seen in the NH winter. Validation results using SABER data are qualitatively similar to those from EOS MLS for both hemispheres.

Summer Polar Latitudes

Cold bias regions for polar latitudes of both northern and southern summer months are contained between ~0.3→0.1 hPa. Unlike the biases seen for the winter polar latitudes the size and magnitude of these cold bias regions is very similar for months both pre and post March 2006. Here the maximum values tend to occur below the model lid at pressure values of ~0.15 hPa and decrease in magnitude both above and below. Validation results using EOS MLS and SABER data are qualitatively similar. Bias magnitudes for both NH and SH summer months have similar magnitudes of ~10→14 K.

From ~0.3→0.1 hPa both EOS MLS and SABER data have consistent cold biases with the correlative data discussed in Sections 3.3.5 and 3.4.5. Excluding the EOS MLS comparison with GEOS-5 and ECMWF analyses, magnitudes of both global cold biases are ~1→5 K. Therefore it is possible that the cold biases seen at the uppermost levels could be larger than that suggested by Figures 4.3 and 4.4. While global biases between EOS MLS and SABER data in this pressure range are near zero, the superior vertical resolution of the SABER instrument, over four times that of EOS MLS, would place more confidence in the SABER validation results where noticeable differences are seen compared to results using EOS MLS data, i.e. polar latitudes.

4.3.2 Impact of Dynamical Forcing

The cold biases at winter polar latitudes suggests that there is insufficient adiabatic warming, due to the vertical descent of the mesospheric circulation in these regions. By the reasoning explained in Section 2.3.5 this would suggest that wave forcing in the extra-tropical winter hemisphere is too weak, and the modelled mesospheric circulation produces temperatures which are too close to their radiative equilibrium values at polar latitudes. Since the biases are present at mesospheric levels and latitudes poleward of the polar night jet, this strongly suggests that the weak forcing is due to a misrepresentation of small scale grav-

ity waves within the UM. With direct observations only assimilated below 1.0 hPa [Saunders et al., 2002a], the data assimilation process in unable to correct for such large cold bias values in the lower mesosphere. This large cold bias is typical of modern GCM's [Beagley et al., 1997, Garcia and Boville, 1994, Pawson et al., 2000] and is commonly known as the "cold pole problem".

As stated above from ~0.3→0.1 hPa the L50 analyses have cold biases of varying magnitude across all latitudes. Largest bias values have a strong seasonal signal, maximum bias values are seen in the winter polar regions, strongly suggesting that they are primarily caused by a misrepresentation of the lower mesospheric circulation. However, theses cold biases are consistently accompanied with cold biases of ~6→14 K in the tropics and opposing summer hemispheres. If biases seen in the analyses from ~0.3→0.1 hPa were completely related to an inaccurate circulation, then one would expect to see a warm bias in the summer hemisphere polar regions of solstices months. Therefore it is implausible that an inaccurate representation of the lower mesospheric circulation alone, with unrealistic adiabatic warming/cooling at winter/summer polar latitudes, could result in bias values with identical signs across an isobaric surface as reported in Figures 4.3 and 4.4. This strongly suggests that the cold biases from ~0.3→0.1 hPa are caused by a combination of misrepresenting the lower mesospheric circulation and other possible sources of error such as inaccuracies within the radiation or data assimilation schemes, discussed below.

The argument that small scale gravity wave forcing in the lower mesosphere is underestimated in the L50 analyses is consistent with the results discussed in 7.3.1. Here off-line calculations of the gravity wave forcing in the lower mesosphere of the L50 analyses have values which are lower than previous estimates for this region. The difference between the off-line calculations and previous estimates is most noticeable in the winter lower mesosphere, where as described above we find the largest biases in the L50 analyses.

4.3.3 Impact of the Ozone Climatology

The consistent cold bias signal of ~2→10 K across all latitudes from ~0.3→0.1 hPa, most prominently shown in the validation results with respect to SABER, would strongly suggest that the radiative heating in this region of the atmosphere is underestimated. As previously noted in Jackson et al. [2008], this inaccurate radiative heating maybe caused by the monthly prescribed ozone climatology of Li and Shine [1995], used in the calculation of L50 short wavelength radiative heating rates, underestimating ozone mixing ratios in the lower mesosphere [Mathison et al., 2007]. The Li and Shine ozone climatology is a monthly averaged data set primarily based on satellite data from the Solar Backscatter Ultraviolet radiometer (SBUV), Stratospheric Aerosol and Gas Experiment II (SAGE II), Solar Mesosphere Explorer (SME) and the Total Ozone Mapping Spectrometer (TOMS) over the period 1985 to 1989. Previously Mathison et al. [2007] conducted experiments detailing the impact in the L50 model when the Li and Shine ozone climatology is replaced by; (a) SPARC ozone climatology based on similar satellite data to Li and Shine over the period 1979-1997 [Randel and Wu, 1999]; (b) ozone fields from ECMWF assimilated analyses; (c) assimilation of ozone [Jackson, 2007] using EOS MLS and SBUV data; (d) assimilation of SBUV ozone data only. These experimental ozone fields, derived using more recent data for the SPARC climatology and up to date fields for the ECMWF and assimilation experiments, showed increased values of monthly averaged ozone mixing ratios compared to the Li and Shine climatology throughout the lower mesosphere for both January and February 2006. Both the SPARC and the two assimilated ozone fields have mixing ratios which are ~2.8 times greater than Li and Shine values approaching ~0.1 hPa, see Figure 6 of Mathison et al. [2007]. ECMWF fields show smaller differences and are ~2.2 times greater than Li and Shine values approaching ~0.1 hPa. The consistency of these differences highlights the deficiency of the Li and Shine climatology approaching the L50 model lid, however there is still considerable uncertainty in ozone mixing ratios of all four experimental fields in the mesosphere.

Forecast model runs were performed by Mathison et al. [2007] over January and February 2006 using the four experimental ozone fields described above and then compared to control results using the Li and Shine climatology. The resulting change in average 24 hour forecast temperatures between the control and experiment fields over the first 15 days of February are shown in Figure 4.6, reproduced from Figure 8 of Mathison et al. [2007].

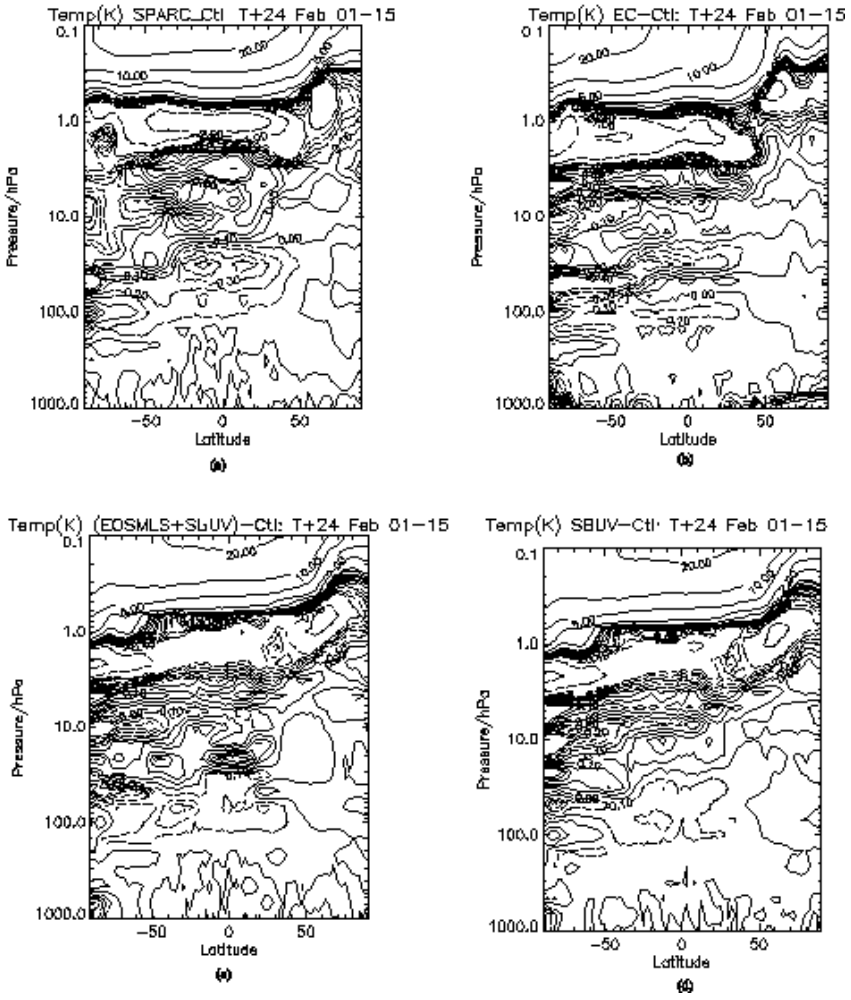

Figure 4.6: Temperatures differences between experimental and control run ozone experiments for the 1^{st} to the 15^{th} February 2006, see text for full details. (a) SPACR climatology, (b) ECMWF assimilated fields, (c) EOSMLS + SBUV data and (d) SBUV data only. Negative contours are dashed. Units (K). Reproduced from Mathison et al. [2007].

Largest temperature differences are seen between ~0.3 →0.1 hPa, where all experimental fields show an increase of forecast temperatures with maximum magnitudes of ~20 K seen between the SH pole and mid-latitudes of the NH. For all experimental fields the magnitude of temperature differences decreases with decreasing altitude to ~10 K at ~0.3→0.2 hPa. From ~0.6→0.3 hPa the temperature differences for all experimental ozone fields are consistent with each other, with differences continuing to decrease with decreasing altitude from ~10 K at ~0.3 hPa to ~2 K at ~0.6 hPa. The evidence presented in Mathison et al. [2007] clearly shows that the Li and Shine ozone climatology is inconsistent with more recent observational evidence of ozone concentrations in the lower mesosphere, i.e. mixing ratios are too small. As reported above, the magnitude of temperature differences obtained when replacing the Li and Shine climatology with more recent representations of ozone concentration are of the same order as the latitudinally independent cold bias signal seen in the lower mesosphere of the L50 analyses. This would strongly suggest that the deficiencies in the Li and Shine ozone climatology are likely responsible for the latitude independent cold bias signal seen between ~0.3→0.1 hPa in the L50 analyses. Upgrading the representation of ozone concentration in the lower mesosphere, by any of the four alternatives discussed in Mathison et al. [2007], could result in the following impact on biases between ~0.3→0.1 hPa for L50 analyses post March 2006.

For all seasons the cold biases of ~2→10 K seen between the mid-latitudes of each hemisphere would be reversed to warm biases of ~4→8 K. The strength of this reversal would be greatest when using the SPARC climatology. Cold biases in both the NH and SH polar regions of the equinox months would be reduced to ~4→8 K. For solstice months cold biases in both the NH and SH winter polar regions would be reduced by less than ~5 K, while cold biases in the summer polar regions of both hemispheres would be reversed to a warm bias of ~8→20 K. The magnitude of this reversal in the summer polar regions would be largest when using the ECMWF ozone fields.

Based on the above hypothesis, it is apparent that the replacement of the Li and Shine ozone climatology could still result in degraded analyses temperatures between ~0.3→0.1 hPa [Mathison et al., 2007], i.e. warm biases between mid-latitudes of all months and in the summer polar regions. It is possible that such degrading is a consequence of inaccuracies present in the experimental ozone fields, which as noted above contain considerable uncertainty. However, previously Sassi et al. [2004] suggested that ozone mixing ratios follow the diurnal cycle [Chapman et al., 1970], having reduced concentrations during daylight hours when solar radiation is absorbed, resulting in a significant impact on the short wavelength heating rates in the lower mesosphere. Hence models which apply an ozone climatology, where concentrations of ozone do not follow a diurnal cycle and are constant in time, will tend to overestimate mixing ratios and therefore radiative heating due to ozone. The results of Sassi et al. [2004] show differences between temperatures obtained using a fully interactive ozone chemistry model [Park et al., 2004] and those obtained when using an ozone climatology compiled from the interactive experiment. Temperatures using the ozone climatology are ~2→6 K warmer from ~0.3→0.1 hPa between mid-latitudes of both hemispheres over all seasons, and ~1→2 K colder for winter polar regions. These results are qualitatively consistent with the temperature differences reported between mid-latitudes and from ~0.3→0.1 hPa in Figure 8 of Mathison et al. [2007], where despite SPARC and both assimilated experimental ozone fields showing qualitatively similar increases of mixing ratios compared to the Li and Shine climatology, SPARC temperatures are noticeably warmer than those obtained using either of the two assimilated field experiments. Here the use of assimilated observational ozone profiles, which are used in the radiative heating calculations [Jackson, 2007], at appropriate local times are more likely to better represent the diurnal change of ozone concentration in the lower mesosphere, hence producing lower temperatures than the SPARC climatology which uses fixed ozone mixing ratio values.

It is therefore possible that the latitudinal independent cold bias signal of ~2→10 K seen between ~0.3→0.1 hPa for all months of L50 analyses could possibly be amended by upgrading the Li and Shine ozone climatology, however this hypothetical situation would result in warm biases between mid-latitudes of all months and summer polar regions of solstice months. This over compensation of the observed cold bias, to a hypothetical warm bias between mid-latitude values is most likely to be less extreme when using assimilated ozone fields to replace the Li and Shine climatology, as it is plausible that excessive radiative heating due to fixed ozone mixing ratios would be reduced. This would suggest that probably the most effective course of action to reduce systematic biases due to the deficiencies of the Li and Shine ozone climatology would be the assimilation of real time ozone data. However this choice of action does significantly affect the operational running time of analyses output, where as noted in Mathison et al. [2007] the assimilation of real time ozone profiles could add an additional ~20-25% to computational expense. As an alternative approach, it would be possible to upgrading the Li and Shine ozone mixing ratios with a climatology that includes a diurnal cycle.

As noted above, the hypothetical impact of upgrading the Li and Shine ozone climatology could result in warm biases of ~8→20 in the summer hemisphere polar regions of solstice months. Justification of this hypothesis is shown in Figure 13 of Mathison et al. [2007], where forecast cold biases of ~8→12 K with respect to EOS MLS temperature profiles in the polar regions of the southern summer hemisphere are replaced by warm biases of ~8→16 K when assimilated EOS MLS and SBUV ozone profiles replace the Li and Shine climatology. While such hypothetical warm polar biases of the summer hemisphere are an obvious deficiency, they are theoretically consistent with the cold biases seen in the opposing winter polar regions described above, placing additional confidence in the statement that gravity wave forcing of the lower mesosphere is underesti-

mated in the L50 analyses.

Similar arguments to those presented above for solstice months also apply to equinox seasons. Here the hypothetical impact of upgrading the Li and Shine climatology could possibly result in a reversal of cold to warm biases between mid-latitudes of each hemisphere from ~0.3→0.1 hPa, accompanied by a small reduction of the cold biases seen in polar regions of each hemisphere. As for the solstice seasons this is theoretically consistent with the assumption that gravity wave forcing of the lower mesosphere is too weak, where the weaker equinox two cell circulation in both hemispheres produces insufficient adiabatic cooling/warming at tropical/polar latitudes.

4.3.4 Stratopause Warm Bias

Figures 4.3 and 4.4 both show that the cold bias regions from ~0.3→0.1 hPa discussed above are predominantly accompanied by a warm bias region directly below. These warm bias regions are generally found from ~2.0→0.3 hPa over various latitudes.

Tropical Latitudes

Pre March 2006 EOS MLS validation results show there is a consistent warm biases for all months occurring between ~1.5→0.9 hPa with maximum values of 6→10 K, except January 2005 which has maximum values of 2→6 K. Similar magnitudes are seen for SABER validation results, however bias regions have larger vertical range from ~2.0→0.8 hPa.

Post March 2006 Magnitudes of the warm bias with respect to EOS MLS data have increased by ~4 K and now extends from ~1.5→0.3 hPa. SABER validation results have identical magnitudes to months pre March 2006 and consistent with EOS MLS results now extend upwards to ~0.3 hPa. The extension of this warm bias to higher altitudes is assumed to be a direct effect of the corrections made in

the AIRS data assimilation code, where the cold biases pre March 2006, discussed in Section 3.1.4, are assumed to have reduced the consistent warm bias signal evident in months post March 2006.

Extra-tropical Latitudes

Pre March 2006 EOS MLS validation results show warm biases occur from ~3.0→0.9 hPa directly below the strong polar winter cold bias regions of both hemispheres. Here larger bias values of 14→18 K are seen in the SH winter, compared with 10→14 K for the NH winter months. EOS MLS validation also reports warm biases between ~0.8→0.3 hPa. Under solstice conditions for both hemispheres these warm bias regions are centred at polar latitudes in the summer hemisphere and mid-latitudes in the winter hemisphere, with magnitudes between ~2→10 K. SABER validation results are qualitatively similar to those of EOS MLS from ~3.0→1.0 hPa. However the warm bias regions reported in EOS MLS validation results from ~0.8→0.3 hPa are not seen in Figure 4.4, as SABER data is consistently ~2→6 K warmer than EOS MLS data in this latitude and pressure range. This temperature bias between EOS MLS and SABER data is consistent with magnitudes presented in Figure 22 of [Schwartz et al., 2007] for extra-tropical latitudes.

Post March 2006 As for tropical latitudes warm biases w.r.t to EOS MLS data extend over a larger vertical range, extending upwards to ~0.3 hPa. EOS MLS validation results show warm bias regions at extra-tropical latitudes consistently occur at identical pressure levels to warm bias seen in the tropics. Consequently the majority of months have latitude independent warm biases from ~0.9→0.3 hPa with magnitudes consistently ~10→14 K. The exception to this is at polar latitudes of the winter months for both hemispheres, where warm bias values are significantly reduced. This reduction is most evident for polar regions of SH winter months in 2006, where there is actually a cold bias of ~2→10 K between ~0.9→0.3 hPa. SABER validation results typically have magnitudes ~2→6 K

smaller than EOS MLS results from ~0.9→0.3 hPa, consistent with the differences noted in months pre March 2006 discussed above.

From 0.8→0.3 hPa EOS MLS data, excluding comparison with SABER, has an observed cold bias of 0→7 K with the correlative data discussed in Section 3.3.5. From 0.8→0.3 hPa SABER data, excluding comparison with EOS MLS, has an observed cold bias of 1→5 K with the correlative data discussed in Section 3.4.5. Therefore it is possible that the warm biases in this region could be smaller than that suggested in Figures 4.3 and 4.4. Evidence from Remsberg et al. [2008], Schwartz et al. [2007] suggests that SABER temperature profiles have predominantly closer values to correlative data than EOS MLS data in this pressure range. The vertical resolution of SABER data is also approximately four times that of EOS MLS data for this pressure range. Qualitatively this would tend to place a higher level of confidence in the SABER validation results from ~0.8→0.3 hPa.

4.3.5 Impact of the Radiation Scheme

The latitudinally independent nature of the above warm biases, most prominent in months post March 2006, would strongly suggest they are caused by inaccuracies of the radiative heating calculations at these levels. The currently operational Edwards-Slingo radiation scheme covers both short and long wavelength spectral bands, where the primary variables used in the code are the upward and downward (diffuse plus direct) radiative fluxes and the direct incident solar flux at the model lid boundary; full details can be found in Edwards et al. [2004]. At the model lid the total downward long wavelength radiative flux and downward short wavelength diffuse radiative flux are set equal to zero. Therefore the boundary condition for the total downward flux at the model lid is equal to the direct incident solar flux, which depends directly on the solar irradiance in the spectral band at the model lid, obtained from a predefined solar spectrum, and the solar zenith angle. One of the major features of the Edwards-Slingo scheme is the fact that the resolution of the incorporated spectral bands is not fixed. This

flexibility is achieved by the use of dependent pre-processed spectral files which can be used in subsequent versions of the UM. The spectral files contain specific physical information on the radiative active gases and atmospheric aerosols used in the radiation scheme. Such physical information includes the number of species used in the radiation scheme, spectral band widths for each gaseous species, the fraction of the input solar spectrum in each band, scattering coefficients for water droplets and ice crystals, absorption coefficients and information on the *k*-fit for gaseous transmissions [Edwards et al., 2004]. Full details of the format and information stored in these spectral files can be found in Edwards and Thelen [2003].

When the Edwards-Slingo radiation scheme was applied to the operational middle atmospheric configuration of the UM in April 2000, producing L42 analyses, the spectral files used were based on those from the Hadley Centre Coupled Climate Model version 3 HadCM3 model detailed in Gordon et al. [2000]. Both the HadCM3 spectral files and the Edwards-Slingo radiation scheme were initially developed to accurately represent tropospheric radiative heating, and were not optimised for stratospheric and lower mesospheric conditions. The solar spectrum used in the HadCM3 spectral files is that of Labs and Neckel [1970], compiled from observational data sets obtained in the 1950-1960's with a spectral resolution of 0.01 μm. Inaccuracies of the upper stratospheric and lower mesospheric heating rates obtained when using the HadCM3 spectral files was demonstrated in Jackson et al. [2001] who, using a 49 level configuration of the UM, found that longwave heating rates from the Edwards-Slingo scheme were only consistent with those from a line-by-line model up to ~2 hPa. Around and above the stratopause region long wavelength heating rates from the Edwards-Slingo scheme were found to be too large when compared to the line by line model, i.e. cooling in this region is too small. Largest differences were found between ~2.0→0.3 hPa, with magnitudes of the order 1-1.5 K day^{-1}.

With the introduction of the semi-Lagragian dynamical core in October 2003, the spectral files used by the Edwards-Slingo scheme were upgraded to those from the Hadley Center Global Environment Model (HadGEM) [Davies et al., 2005]. One of the major changes in this upgrade was the introduction of the Kurucz and Bell [1995] solar spectrum, which is based on a theoretical model [Zhong et al., 2008] and has a spectral resolution of 1 cm^{-1}. Noticeable differences between the above solar spectra primarily occur for small wavelengths $\lambda < 1\mu$m, where the newer Kurucz spectrum has larger solar irradiance values [Thelen, 2006]. In the upper stratosphere ozone dominates the radiative transfer in the short wave with a peak at ~0.25 μm (Hartley Band). Therefore one would expect the HadGEM spectral files, with the Kurucz spectrum, to produce larger ozone heating rates than when using the HadCM3 files. This is indeed the case, as reported by [Thelen, 2006] who found ozone heating rates ~2.0→0.4 hPa were ~1.5→2.0 K day^{-1} warmer for HadGEM files compared to HadCM3. From the same study it was noted that radiative heating rates in the longwave region produced by the HadCM3 and HadGEM spectral files are qualitatively similar. Here both files have heating rates due to CO_2 and O_3 which are qualitatively similar to those from the GFDL line by line model, however they both underestimate the cooling due to H_2O by ~0.8 K day^{-1} at ~1.0 hPa, consistent with the excessive long wavelength heating reported in Jackson et al. [2001].

The spectral files used in the operational L50 model are essentially identical to those in the HadGEM models, with minor changes to the representation of atmospheric aerosols. It is therefore reasonable to expect that the use of HadGEM spectral files in the L50 analyses would lead to similar heating rate biases as those discussed above. Here the combined inaccuracies of both the short and long wavelength regions could likely lead to possible excessive net heating rates of ~1.0→2.0 K day−1 from ~2.0→0.4 hPa in the L50 analyses. With the radiative time scale of the atmosphere in this region ~5→7 days [Gille and Lyjak, 1986, Haynes, 2005], it is plausible that the above hypothetical heating rate errors

could result in warm temperature biases of the order ~7-14 K from ~2.0-0.3 hPa. These hypothetical warm biases have similar magnitudes to the observed warm biases detailed in Figures 4.3 and 4.4 however with ~4→6 K larger values for pressure levels closer to 0.3 hPa, particularly for the SABER validation results which has a higher level of confidence in this pressure range. However as described above, smaller ozone mixing ratios in the currently operational Li and Shine ozone climatology produce temperature values between ~0.6→0.3 hPa which are too cold when compared to temperatures produced by more recent and accurate representations of ozone concentrations. Accounting for these cold temperatures could adjust the above hypothetical warm biases to values which are qualitatively similar to those observed. It is therefore likely, and most probable, that the combined deficiencies in the Li and Shine ozone climatology and the heating rate calculations of the Edwards-Slingo radiation scheme are the cause of the warm biases seen in upper stratosphere and lower mesosphere regions of the L50 analyses. Therefore as stated in Mathison et al. [2007] any upgrades to the radiation scheme should be tested alongside upgrades to the representation of ozone in the NWP system.

The above radiative deficiencies are known to the Met Office and their effects are well documented in Bushell [2005], where climate model runs using the HadGEM spectral files produce stratopause temperatures which are consistently warmer than various other GCM's and earlier versions of the UM. Experimental modifications to the HadGEM spectral files to reduce the above warm bias have made by Wenyi Zhong of the Imperial College London and are detailed in Thelen [2006]. Results from these experiments proved positive, warm biases from 6 year climate runs were reduced by ~7 K between ~1.0→0.3 hPa as documented in Figure 3 of Thelen [2006], resulting in incremental changes to the spectral files used in recent versions of HadGEM models. While initial tests of the Zhong modifications in the NWP global model also proved positive in the stratosphere, there was a detrimental impact on tropospheric temperatures significant enough

to reject it's incorporation into the operational system [Bushell, 2010]. Subsequently the use of the modified Sun-Edwards-Slingo radiation scheme [Sun and Rikus, 1999] is currently being considered to address the above deficiencies.

It should be noted that the Met Office are currently working to address the above deficiencies of radiative heating throughout the lower mesosphere, which we have attributed to be most likely caused by a combination of inaccuracies from both the representation of ozone mixing ratios and heating rates produced by the Edwards-Slingo radiation scheme. Therefore the remaining body of work in this thesis does not address these issues, but concentrates on characterisation and reduction of biases caused by the inaccurate representation of the mesospheric circulation which is primarily driven by small scale gravity wave dissipation.

4.4 Climatological Comparison of L42 and L50 Analyses

From 1991 to the present day several changes have been made to all aspects of the NWP model that creates the stratosphere assimilated dataset. In order to understand the overall impact these changes have made in the middle atmosphere we shall compare the validation of L42 analyses, from 1992-1994 using UARS MLS data, with those from the validation of L50 analyses discussed above.

Figure 4.7 shows validation results of monthly zonal mean temperature errors between L42 analyses and UARS MLS values averaged over the years 1992-1994. For comparison the monthly zonal mean errors of Figure 4.3 averaged over the years 2005-2009 are also shown. Note only temperature differences in the UARS latitude and pressure range are shown for the EOS results. Latitudinal and vertical bins used in the calculation of the UARS MLS results were identical to those used for EOS MLS.

From Figure 4.7 (a) we find that overall the L42 analyses from 1992-1994 have a cold bias across all latitudes and heights with respect to the UARS MLS temperatures. This cold bias tends to between -2→-6 K through the middle and upper stratosphere, increasing to -10→-18 K above the stratospause at polar latitudes of both winter hemispheres. From Section 3.2.5 we find that UARS MLS data, excluding comparison with CIRA'86, generally has a warm bias between 2→6 K with correlative data above the stratopause. This would suggest that the cold biases discussed above could possibly have magnitudes less than those presented in Figure 4.7.

Generally the validation results of L50 analyses from 2005-2009, detailed above and shown in Figure 4.7 (b), show that analyses temperatures are typically warmer than observations compared to the cold biases seen for L42 analyses from 1992-1994. Here biases magnitudes are between 2→14 K at and above the stratopause level.

Figure 4.7 (c) shows the difference between UARS MLS monthly zonal mean temperature profiles, averaged from 1992-1994, and those from EOS MLS, averaged from 2005-2009. Generally UARS MLS temperatures are 2→6 K warmer than EOS MLS values throughout the stratosphere and lower mesosphere. Larger warm biases of 10→14 K are seen at polar latitudes, primarily in the winter seasons of both hemispheres in the lower mesosphere, and for various months in the middle stratosphere. The magnitude of this temperature bias is significantly larger than climate cooling trends [Ramaswamy et al., 2001]. Thus the above bias could possible be related to systematic biases between each instrument, however a large contribution to this bias could arise from the inter-annual variability of the atmosphere which is known to be larger for the winter season. With only three years of UARS MLS data compare to five for EOS MLS, it is impossible to state confidently on the magnitude of the systematic biases between the two observational data sets.

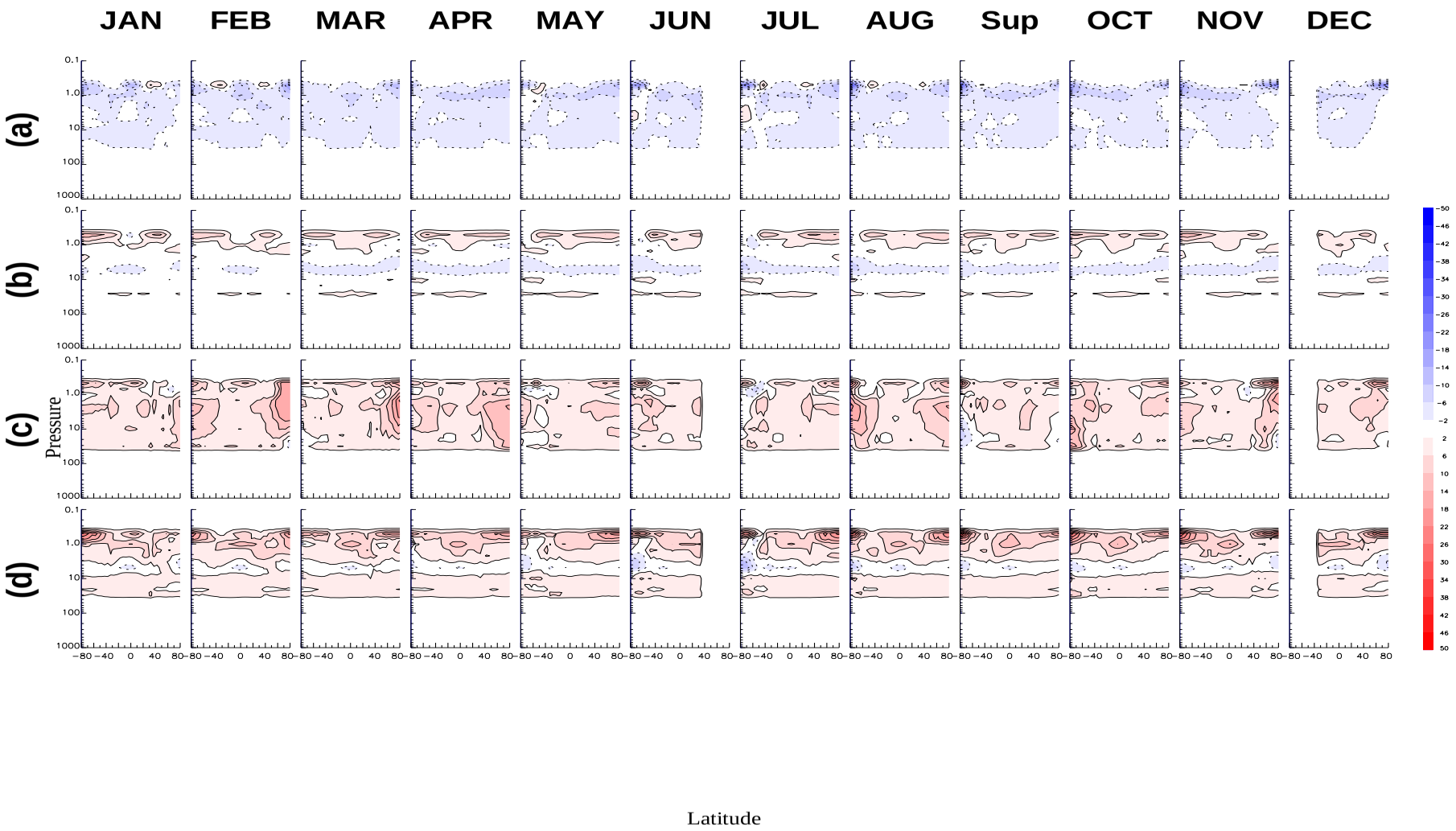

Figure 4.7: Monthly zonal average temperature differences between (a) L42 Analyses and UARS MLS data, averaged over 1992-1994; (b) L50 Analyses and EOS MLS data, averaged over 2005-2009; (c) UARS MLS data averaged over 1992-1994 and EOS MLS data averaged over 2005-2009; (d) Validation results of (b) minus (a). Contour level 4 K.

Figure 4.7 (d) shows the L50 validation results minus those of the L42, essentially Figure 4.7 (b)-(a). These results show, (L50 analyses - EOS MLS) - (L42 analyses - UARS MLS), which can be interpreted as (L50 analyses - L42 analyses) + (UARS MLS - EOS MLS). Allowing for the warm bias of UARS MLS profiles compared to EOS MLS, we find that in the upper stratosphere and lower mesosphere the L50 analyses errors from 2005-2009 are generally ~2→10 K warmer across all latitudes compared to L42 analyses errors from 1992-1994. The magnitude of this temperature difference increases approaching ~0.3 hPa and is again significantly larger than climate cooling trends, suggesting that the above difference between L42 and L50 analyses is due to changes in the model configuration.

The latitudinal independence of the above warm bias would strongly suggests that the difference in analyses errors from the UARS to EOS temporal ranges is primarily due to the change of radiation scheme used. The radiation scheme used in the L42 analyses from 1992-1994 is detailed in Ingram [1990] and is based on the scheme described in Stratton [1999] with modifications to the long wavelength heating rates following the work of Morcrette et al. [1986]. This scheme was changed in 2000 to the now presently used [Edwards and Slingo, 1996] scheme. Previously Jackson et al. [2001] noted that from ~2.0→0.3 hPa the long wavelength heating rates from the Edwards-Slingo scheme are larger than those from the scheme described by Stratton [1999], with magnitudes ranging from ~0.5-1.0 K day^{-1}. As noted above the L50 analyses from 2005-2009 also use HadGEM based spectral files which have been shown to produce excessive heating rates around and above the stratopause region. With radiative time scales in this region of the atmosphere being ~5→7 days, the increased heating rates of the Edwards-Slingo scheme compared to the scheme described by Stratton [1999] is the most probable cause of the ~2→10 K difference in temperature biases seen between the L42 and L50 analyses in the upper stratosphere and lower mesosphere.

4.5 L70 Analyses

As noted in Section 3.1.2 the model lid of the operational model was extended to 0.01 hPa during November 2009. This section details the L70 zonal mean temperature fields and their zonal mean temperature biases with both EOS MLS and SABER data. The impact of increasing the model domain on observed biases is discussed through comparison with the previous results of the L50 model discussed above. It should be noted that the most significant change, apart from the increased domain, is the use of the vertical levels detailed in Figure 3.1 (b) which has a degraded resolution throughout the upper stratosphere and lower mesosphere compared to the previous L50 configuration. The spectral files (different vertical resolution based on the vertical level configuration), ozone climatology and parameters of the USSP gravity wave scheme used in the L70 model are identical to those used in previous L50 configuration.

Figure 4.8 shows the monthly zonal mean temperature values for the L70 analyses from November 2009 to September 2010. The temperatures are reported on grids which have 31 standard UARS pressure levels and a latitudinal resolution of 2.5^{o}. For comparison monthly zonal mean temperatures from the EOS MLS instrument, including the upper mesosphere up to 0.01 hPa, are shown in Figure 4.9.

Lower Mesosphere

Compared to L50 analyses, the L70 representation of the lower mesosphere is qualitatively superior. Here the excessive vertical temperature gradients seen in the L50 analyses at ~0.5-0.4 are not present. Furthermore the L70 analyses show an improvement in representing the latitudinal variation of temperature between ~0.5-0.1 hPa, with meridional temperature gradients having qualitatively the correct sign and magnitude for all months. The exception to this is for November, December, Feburary and March where in the NH meridional tem-

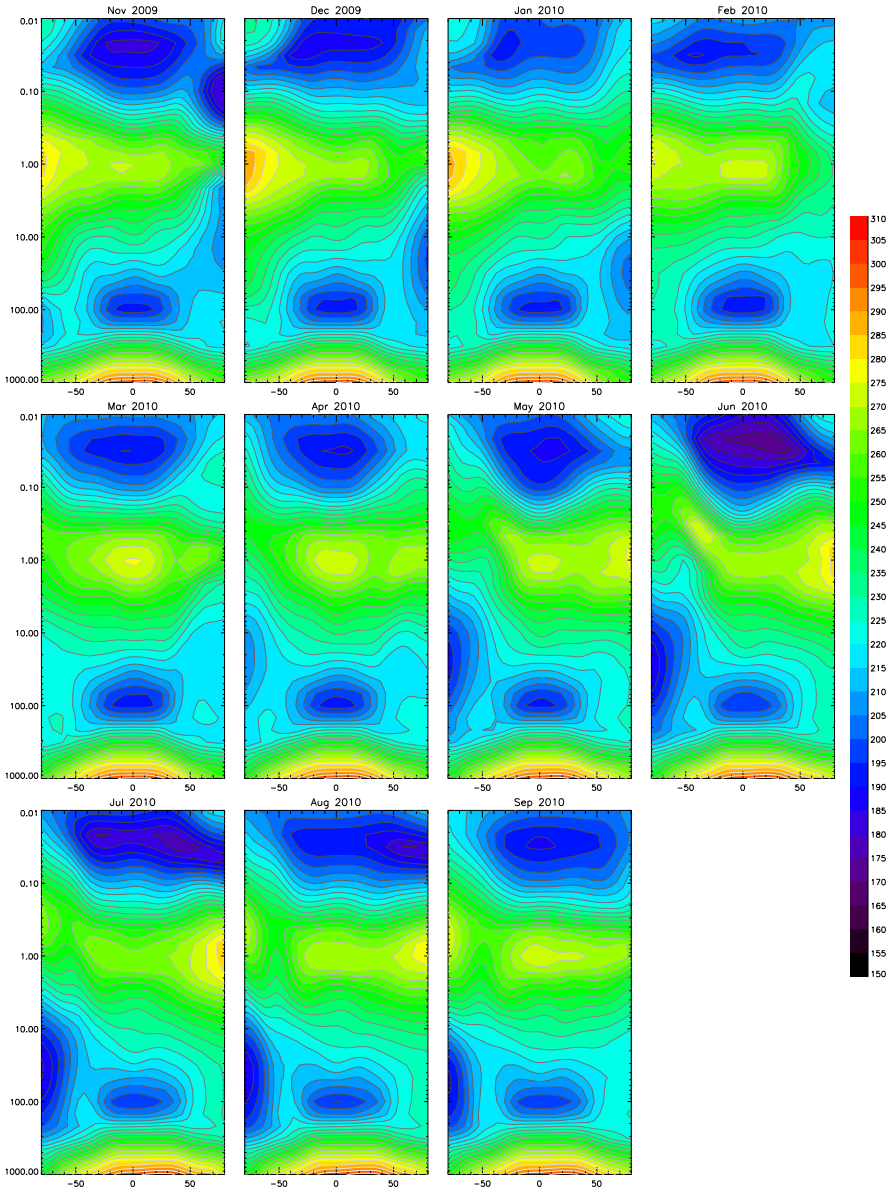

Figure 4.8: Monthly zonal mean temperatures of the L70 analyses from November 2009 to September 2010. Contour interval is 5 K.

perature gradients poleward of ~40^{o} result in decreasing temperatures towards polar latitudes, instead of increasing as observed by EOS MLS values. Another noticeable deficiency is the poor representation of the separate and elevated winter polar stratopause feature discussed in Section 2.2, which is absent for November and located too low at ~0.8 hPa for December and January.

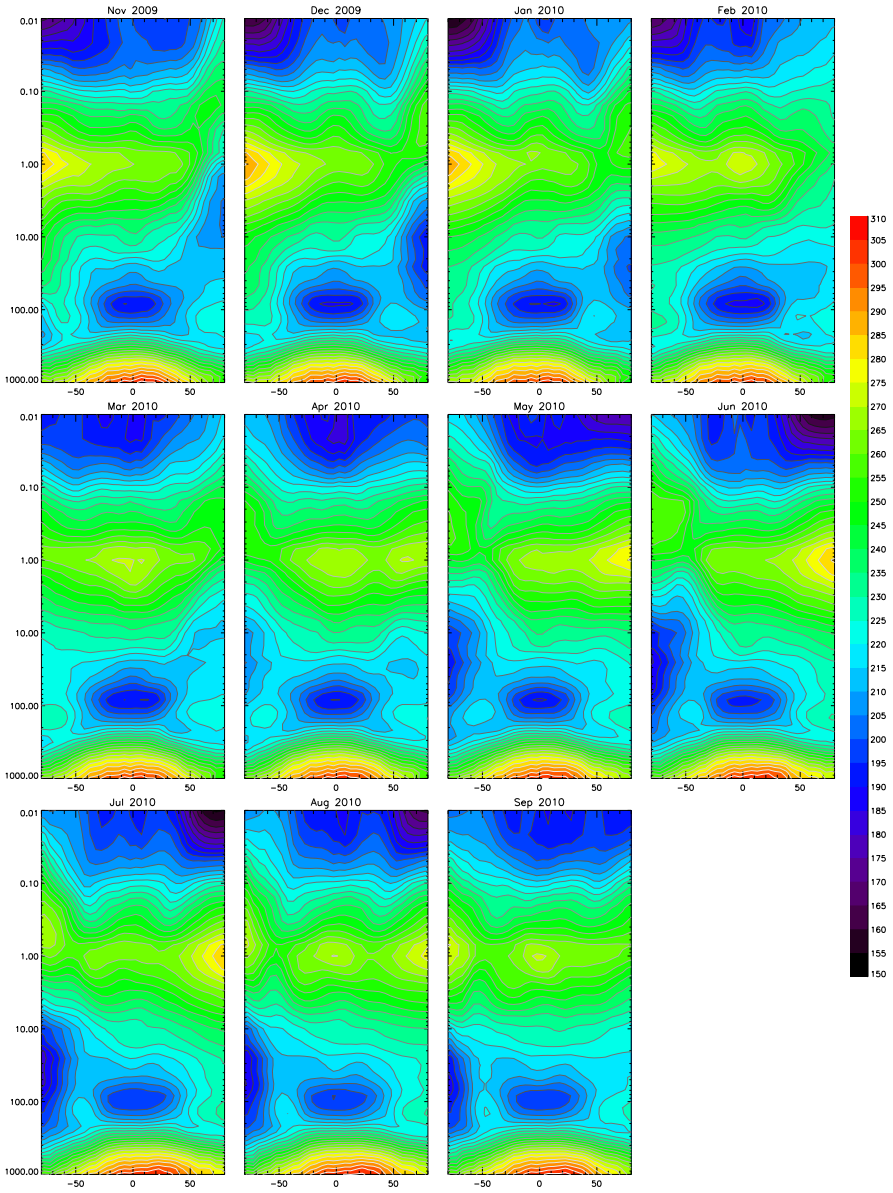

Figure 4.9: Monthly zonal mean temperatures from the EOS MLS experiment. Contour interval is 5 K.

Upper Mesosphere

In the upper mesosphere the L70 representation is relatively poor when compared to the EOS MLS values. For solstice months both summer hemispheres have minimum temperatures located at tropical latitudes instead of in the polar

regions at 0.01 hPa as observed. This results in meridional temperature gradients of both the summer hemispheres having the incorrect sign above ~0.05 hPa, where temperatures decrease towards tropical latitudes. Such gradients are inconsistent with the characteristic monotonic increase of temperature from the summer to winter hemispheres of the upper mesosphere. For February, March and April minimum analyses temperature values are also seen at tropical latitudes from ~0.6-0.3 hPa. Above ~0.3 hPa the analyses temperatures primarily increase towards 0.01 hPa, a feature which is not present in the EOS MLS data.

4.6 L70 Monthly Zonal Mean Temperature Error

Figure 4.10 shows the monthly mean zonal mean temperature difference between L70 analyses values and the EOS MLS dataset, from November 2009 to May 2010. Figure 4.11 shows the monthly mean zonal mean temperature difference between L70 analyses values and the SABER dataset, from November 2009 to May 2010. As for the L50 analyses, blue contour levels correspond to analyses having a cold temperature bias and red contours a warm bias.

4.6.1 Lower Mesosphere

Tropical Latitudes

For tropical latitudes of all months there is a consistent cold bias in the lower mesosphere. Bias magnitudes with respect to EOS MLS typically have maximum values of ~18→22 K at ~0.1 hPa, with values decreasing with an increase of pressure and latitude values. Noticeably the above cold biases extend to lower altitudes when compared to typical L50 analyses locations, this feature is most prominent in the NH winter months where tropical cold biases extend down to 1.0~ hPa. The magnitudes of the above tropical cold bias from ~0.4→0.1 hPa are generally consistent with those seen at similar altitudes and latitudes in the L50 analyses, described above. The validation results with respect to SABER show cold biases of similar magnitude however with marginally smaller vertical

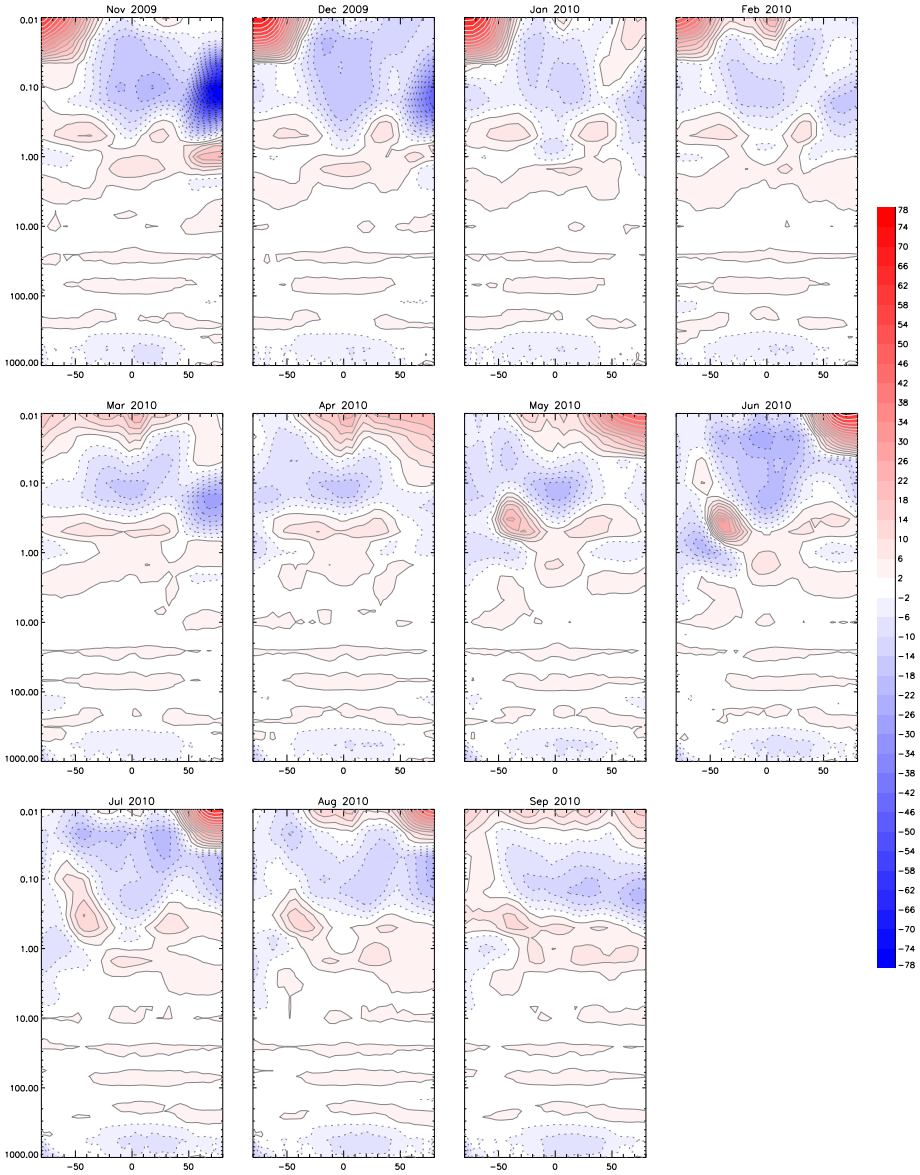

Figure 4.10: Monthly zonal average differences between L70 analyses temperatures and vertical temperature retrievals from the EOS MLS instrument from November 2009 to May 2010. Contour interval is 4 K. Negative biases are indicated by dashed contours.

range, most noticeable in January and February 2010. As noted above the SABER instrument has a higher vertical resolution that EOS MLS at this pressure range, placing more confidence in the SABER results.

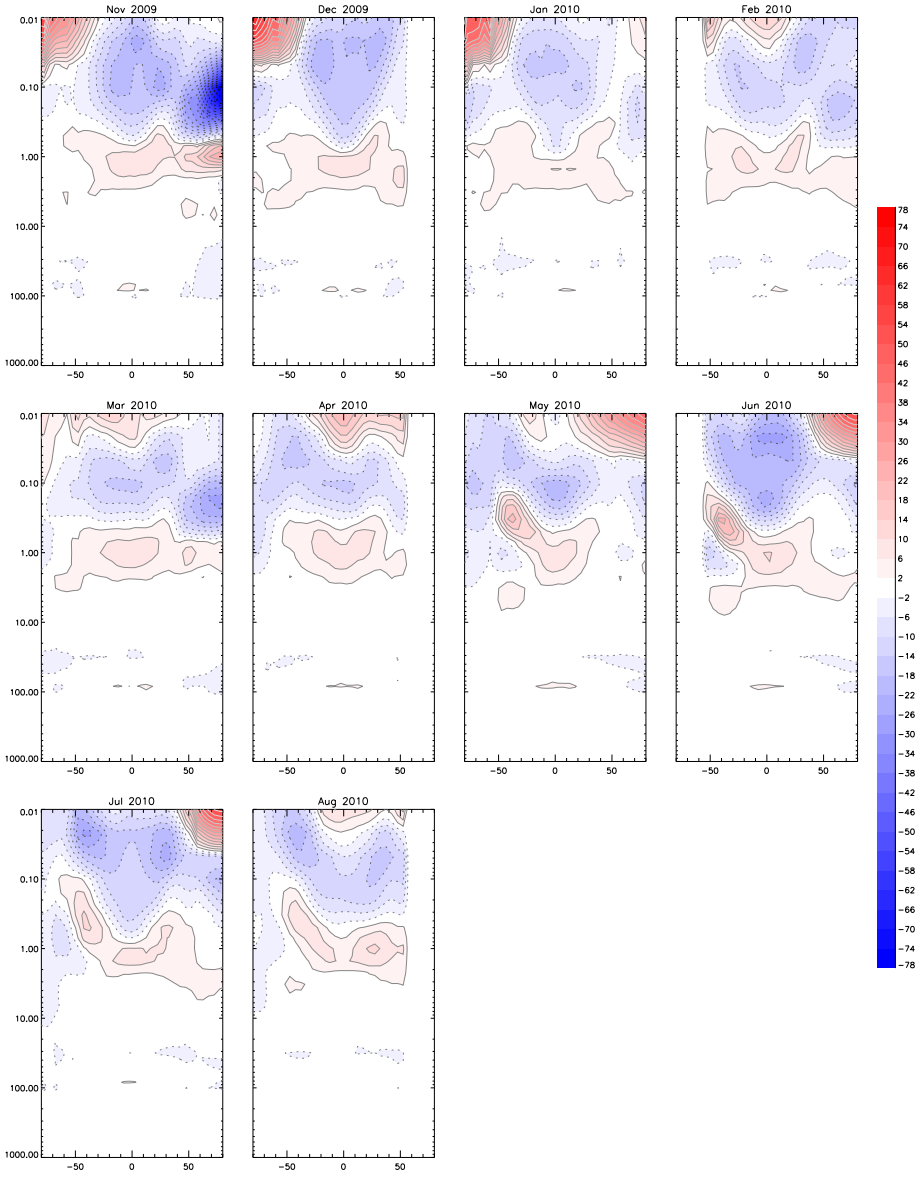

Figure 4.11: Monthly zonal average differences between L70 analyses temperatures and vertical temperature retrievals from the SABER instrument from November 2009 to May 2010. Contour interval is 4 K. Negative biases are indicated by dashed contours.

Spring/Autumn Polar Latitudes

For February, March and April the L70 analyses have cold biases in polar regions of both hemispheres, with maximum values at ~0.2→0.1 hPa that decrease with

altitude. Here NH bias of the L70 analyses tend to extend to ~1.0 hPa, unlike corresponding L50 cold biases which only extend to ~0.3→0.4 hPa. The validation results using EOS MLS and SABER data are qualitatively consistent with each other. Here polar cold bias magnitudes seen in both hemispheres of the L70 analyses are typically smaller than corresponding L50 values. The exception to this is the NH of March 2010, where cold biases of ~26→30 K with respect to EOS MLS and SABER data are ~2→6 K larger than corresponding L50 values. The improved L70 representation is prominent in the SH polar regions and most noticeable in April 2010, where cold biases have magnitudes of ~6→10 K compared to corresponding L50 values of ~26→30 K. The NH polar regions of April 2010 also show a distinct improvement compared to L50 values, with cold bias magnitudes ~6→10 K smaller. Confidence in this improved representation is however limited by only one year of L70 analyses available for validation, compared to 5 years for the L50 analyses.

In September the L70 analyses have cold biases of ~14→22 K in the NH polar regions and a small magnitude warm bias of ~2→6 K in the SH polar regions w.r.t to EOS MLS data. As for February, March and April the L70 representation is an improvement on the L50 biases for this region, most noticeable in the SH where L50 cold biases of ~20 K have been replaced by a small warm bias of ~4 K. However, as stated above confidence in this improved representation is limited.

Winter Polar Latitudes

For NH winter months the lower mesosphere is dominated by a large cold bias which occurs at polar latitudes centred between ~0.3-0.1 hPa. Here cold biases decrease with altitude and typically extend down to between ~0.8→1.0 hPa. For November and December this polar cold bias has respective magnitudes approaching ~70 K and ~44 K when validated against both EOS MLS and SABER data.

The large ~70 and ~44 K cold biases seen respectively in the lower mesosphere of

November and December 2009, are a result of a continuous decrease in analyses minimum temperatures poleward of 60^{o}N at 0.1 hPa that was noted shortly after the L70 analyses were introduced [Jackson, 2010]. This decrease in temperature resulted in the development of a spuriously cold polar vortex at 0.1 hPa. The problem was related to the assimilation process, as a free running version of the model was reported to warm by 10 K day^{-1}, where unrealistic vertical correlations in the background covariance matrix resulted in analysis increments of the stratosphere producing spurious (negative) temperature increments throughout the mesosphere. These unrealistic mesospheric increments were also assumed to be amplified by the increased amount of stratospheric data being assimilated, due to the introduction of NOAA-19 satellite information, and the relatively coarse vertical resolution of L70 analyses, ~5 km in the lower mesosphere. This problem was subsequently addressed by the Met Office throughout December 2009 using a variety of emergency changes, including a linear reduction of analyses increments for the top 8 model levels [Jackson, 2010]. Due to the existence of the spurious decrease in model temperatures at 0.1 hPa, the cold biases seen in the lower mesosphere of November and December 2009 are likely to be uncharacteristic of the L70 analyses. Furthermore, post December 2009 the analyses increments for the mesospshere were no longer added to the analyses fields.

With the uncharacteristic winter polar vortex described above, and the limited number of L70 analyses available for validation, comparisons between L70 and L50 analyses for the NH winter season are obviously unrepresentative of typical differences in biases magnitudes seen in this region. However January 2010 winter polar cold bias maximums of ~14→18 K with respect to EOS MLS data, and ~10→14 K with respect to SABER data, are ~8→12 K smaller than typical L50 analyses values of ~22→30 in the same region, as would be expected from the reasoning described below.

The improved representation of polar winter temperatures in the L70 analy-

ses, compared to those of the L50 model, is also evident in the SH. Here polar cold biases in May, June and July 2010 from ~0.3→0.1 hPa have magnitudes of ~2→10 K w.r.t EOS MLS data, noticeably smaller than the ~26→38 K values typical of the L50 analyses. Validation results using SABER data show similar bias magnitudes. However for the above months L70 analyses have cold polar biases from ~3.0→0.7 hPa in the SH, a feature which is not prominent in the L50 analyses except in the SH winter months of 2006. Magnitudes of these biases are ~2→10 K, ~10→22 K and ~2→10 K for May, June and July respectively and are accompanied by a warm bias of ~14→18 K located at mid-latitudes from ~0.7→0.2 hPa for each month. This cold and warm dipole bias feature of the SH winter lower mesosphere is a direct consequence of the inaccurate representation of the separate winter stratopause discussed above.

Summer Polar Latitudes

Cold bias magnitudes of the SH summer polar regions are located between ~0.3→0.1 hPa, and typically have magnitudes of ~2→10 K with respect to EOS MLS data. The validation results using SABER data are qualitatively consistent with those from EOS MLS. These values are typically ~4 K smaller than corresponding L50 analyses values. Cold biases of the NH summer months are qualitatively identical to those of the SH season, with values again ~4 K less than corresponding L50 analyses values.

4.6.2 Impact of Increased L70 Model Lid

The reduction of bias values seen at polar latitudes between ~0.3→0.1 hPa in the majority of L70 analyses compared to L50 values is both expected and explained by the downward control arguments of Haynes et al. [1991]. Following the arguments presented in Section 2.3.5, it is reasonable to expect that the increased vertical domain of the L70 analyses, approximately three scale heights above the model lid of the L50 model, would better represent the eddy forcing above the lower mesosphere and hence produce a more accurate representation of the

circulation and hence polar temperature via the theory discussed in Section 2.3.5. This is indeed the case, where as described above the cold polar biases seen in the L50 model are generally reduced for the L70 analyses implying a more accurate representation of the lower mesospheric circulation. Additionally to the above downward control argument, one would expect further inaccuracies of the L50 lower mesospheric circulation to be compounded by dynamical and physical boundary conditions at the model lid. In order to conserve mass throughout the model domain the prognostic vertical velocity w is set to zero at the model lid, implying no dynamical heating/cooling at the top theta level of the L50 model. This non-physical feature due the limited vertical domain is not apparent at ~0.1 hPa in the L70 analyses, hence one would expect a more realistic representation of the circulation and temperature at this altitude. Obviously identical boundary conditions for w and associated inaccuracies are present at the L70 model lid.

The gravity wave forcing, which has been shown to dominate the momentum budget of the mesosphere, is supplied in both the L50 and L70 models by the USSP scheme fully detailed in Section 6. As explained in Section 6.1.7 the boundary condition of this physical parameterisation allows remaining momentum flux from the scheme to propagate out of the model domain, known as a "transparent" lid, for both the L50 and L70 analyses. If, for the sake of argument the "opaque" lid condition was employed in the L50 analyses then all remaining momentum flux would be deposited at the model lid. For this hypothetical situation the L50 analyses would then contain a measure of the total gravity wave forcing in the atmosphere above the vertical domain at the model lid, thus by downward control producing a more realistic circulation in the lower mesosphere in the L50 analyses. In some sense applying a transparent lid in the L50 analyses doesn't allow the USSP to use all available information to accurately reproduce the lower mesospheric circulation; however extending the domain in the L70 model reduces this deficiency for the lower mesosphere as can be seen by the reduction of observed polar temperature biases. As stated above for the

vertical velocity boundary condition, similar arguments and associated deficiencies when using a transparent lid condition will also occur at the model lid of the L70 model, resulting in an unrealistic representation of the upper mesospheric circulation. Discussion of the possible impact the transparent lid condition has on the upper mesospheric circulation is detailed below.

Consistent with L50 analyses, there is a latitude independent cold bias signal from ~0.3→0.1 hPa for L70 analyses. The independent nature of this cold bias is most prominent in the validation results with respect to SABER data, which as stated above we have higher confidence in for this pressure range. Since this feature is prominent in both L70 and L50 analyses at qualitatively similar altitudes, additional confidence can be placed in the assumption that the cause of this bias is due to deficiencies in the Li and Shine ozone climatology which remains in operational use for the L70 model. However there are minor discrepancies between the magnitudes of the latitudinally independent cold bias signals seen in the L50 and L70 analyses. Here the latitude independent cold bias magnitudes of the L70 analyses are smaller than those seen in the L50 model by ~4 K, most noticeable at mid-latitudes of each hemisphere. Again due to the limited temporal range of L70 analyses available for validation it is impossible to state confidently the reason for this reduction in the latitude independent cold bias signal from ~0.3→0.1 hPa. One possible explanation is that the L70 analyses contain additional information about radiative processes in the atmosphere above, such as radiative heating due to molecular oxygen, which is treated in the Edwards-Slingo scheme, and is known to be non-trivial above ~0.1 hPa [Andrews et al., 1987]. It is possible that such additional heating information is effectively propagated downwards by radiative diffusion in the L70 analyses, while in the L50 analyses the total downward long wavelength radiative flux and downward short wavelength diffuse radiative flux in the Edwards-Slingo scheme are set to zero at ~0.1 hPa. It is therefore possible that this additional diffuse heating in the L70 analyses could reduce the magnitude of the latitude

independent cold bias when compared to L50 values.

Again consistent with the L50 analyses, warm bias regions occur in the L70 analyses below the cold biases discussed above. However the latitude independent nature of this warm bias in the L70 analyses from ~0.8→0.3 hPa is not as prominent as that displayed in the L50 analyses. For this pressure range validation results using both EOS MLS and SABER data show that warm biases are generally restricted to mid-latitudes, with cold biases occurring at tropical and polar latitudes of certain hemispheres. From ~3.0→0.8 hPa the majority of L70 analyses do display a latitude independent warm bias, most noticeable in the validation results using SABER data, with magnitudes similar to those seen in the L50 analyses from ~1.0→0.3 hPa. The exceptions to this are the months of May, June and July, where as noted above the inaccurate representation of the separate stratopause in the SH winter polar regions result's in a cold bias for both months. The spectral files used in the L70 analyses are identical to those used in the L50 model except for the vertical resolution of the monochromatic layers used, which are based on the vertical grid employed in each model. From Figure 3.1 we find that the L70 analyses have coarser vertical resolution and model levels which occur at different altitudes than the L50 model in the upper stratosphere and lower mesosphere. It is plausible that these differences could possibly result in the warm biases caused by deficiencies in the analyses spectral files to occur at different altitudes in the L70 analyses compared to the L50 model. Since the latitude independent warm biases of both models have qualitatively similar magnitudes, additional confidence can be placed in the assumption that the cause of this bias is due to inaccuracies of the spectral files used in the two different model configurations.

4.6.3 Upper Mesosphere

The upper mesosphere is dominated by strong warm bias regions which have a clear seasonal signal. Warm biases are seen confined to summer polar latitudes

of each hemisphere with maximum magnitudes occurring at the model lid and decreasing in value as they extend downward to typically ~0.07 hPa. A detailed description of all biases in the upper mesosphere is given below.

Tropical Latitudes

For NH winter months the tropical cold biases seen at ~0.1 hPa decrease towards the model lid, where bias magnitudes are typically between ~2→6 K with respect to EOS MLS data. For these months, approaching the model lid the cold biases with respect to SABER data are typically ~4→8 K larger. For February→May the cold biases seen at ~0.1 hPa decrease towards ~0.04 hPa where they reverse sign to a warm bias which increases towards the model lid. Here typical warm bias magnitudes at the model lid are ~14→22 K with respect to EOS MLS data. For June and July 2010 the cold biases with respect to EOSMLS and SABER seen at 0.1 hPa extend throughout the entire upper mesosphere to the model lid, where magnitudes have again decreased to ~2→6 K. The warm biases seen between ~0.03→0.01 hPa for February→May have noticeably smaller magnitudes when validated using the SABER data, here values are typically between ~2→14 K.

Spring/Autumn Polar Latitudes

For February there is a strong warm bias in the SH of ~46→50 K with respect to EOS MLS data and a cold bias of ~6→10 in the NH between ~0.04→0.01 hPa. This cold NH bias is also evident, with similar magnitudes, in the validation results using SABER data. For March 2010 there are warm biases in polar regions of both hemispheres between ~0.04→0.01 hPa, where magnitudes in the SH of ~14→18 K with respect to EOS MLS data are larger than the ~2→6 K values of the NH. Here the validation results using SABER data show a ~6→10 K decrease in the SH warm polar bias and a cold bias of ~2→6 K in the NH. For April 2010 the dominant warm bias has migrated to the NH polar region where it has a magnitude of ~18→22 K with respect to EOS MLS data near the model lid. The SH polar region of this month now has a cold bias of ~2→10 K with respect to both EOS MLS and SABER data between ~0.04→0.01 hPa.

For August there is a warm bias seen at approaching the model lid in the NH polar regions, while in the SH there is a cold bias at high latitudes. Here both bias magnitudes and locations are a mirror image about the equator of those seen in February. For September the polar regions of both hemispheres have a warm bias with similar magnitudes of ~6→14 K.

Winter Polar Latitudes

For NH winter polar regions the magnitudes of biases approaching the model lid are considerably less than those seen in the opposing summer hemisphere, described below. However for November and December 2009 there are strong cold bias signals directly above ~0.1 hPa in the polar NH. These cold biases decrease rapidly with height and are associated with the spurious cold polar vortex at 0.1 hPa detailed above. Above ~0.04 hPa both November and December 2009 have cold biases with magnitudes less than ~10→14 K with respect to EOS MLS data. For these months similar biases are seen in the SABER validation results. For January 2010 above ~0.04 hPa there is a warm bias region at polar latitudes of ~6→10 K. The magnitude and region of this warm biases is smaller for the SABER results, where warm biases are between ~2→6 K.

Consistent with results from NH winter months, the biases approaching the model lid in the SH winter polar regions have considerably smaller magnitudes than those seen in opposing summer hemisphere. In May, June and July 2010 above ~0.04 hPa there are cold biases of ~2→14 K with respect to EOS MLS data, with similar biases also reported using the SABER data.

Summer Polar Latitudes

For the SH summer season the strong warm bias has largest values for the solstice month of December 2009, where magnitudes at the model lid approach ~70→74 K with respect to EOS MLS data. Both November 2009 and January 2010 have summer polar warm biases with smaller magnitudes of ~58→62 K with respect

to EOS MLS data. For the above months the validation results using SABER data are consistently ~6→10 K smaller than those obtained using EOS MLS data.

Consistent with the SH, the warm biases of the NH summer season have maximum values located at polar latitudes near the model lid. For May 2010 the warm polar bias has a maximum magnitude of ~46→50 K with respect to EOS MLS data. In June 2010 this warm bias has increased to ~62→66 K and then decreases to ~54→58 K for July 2010. Again similar biases are reported in the SABER validation results, however with magnitudes generally smaller by ~4→8 K.

In the upper mesosphere EOS MLS data has a consistent cold bias with the correlative data discussed in Section 3.3.5. Excluding comparison with GEOS-5 and ECMWF analyses, from ~0.08→0.03 this cold bias is between ~3→8 K and from ~0.03→0.01 this cold bias is between ~0→3 K. Therefore it is possible that the warm and cold biases in this pressure range could be respectively smaller or greater than those presented in Figure 4.10. In the upper mesosphere at ~80 km the SABER data has an observed bias of -5→9 K with the correlative data discussed in Section 3.4.5. Therefore it is possible that the biases w.r.t SABER data in the upper mesosphere could either be smaller or larger than those presented in Figure 4.11. The SABER data at this altitude is also known to have a higher uncertainty in the summer polar latitudes [Remsberg et al., 2008]. However more confidence would be placed in the SABER validation results at this altitude than those obtained with EOS MLS data, which has a vertical resolution of ~14 km compared to ~2 km for the SABER instrument.

4.6.4 Impact of Radiative and Dynamical Forcing

As in the lower mesosphere from ~0.3→0.1 hPa there is a latitude independent cold bias signal for the majority of L70 analyses in the upper mesosphere, extending from ~0.1~0.07 hPa. For tropical latitudes of solstice months the cold

biases at ~0.1 hPa do extend well into the upper mesosphere and only reduce in magnitude above ~0.02 hPa. An obvious and possible cause of the cold bias signals between ~0.1~0.07 hPa is again the deficiencies in the Li and Shine ozone climatology noted above. Ozone mixing ratios are known to decrease with altitude into the upper mesosphere, where the column above ~60 km contributes less than 1% to total column ozone values, and radiative heating due to ozone decreases towards the mesopause [Groves, 1982]. It is therefore expected that cold biases caused by deficiencies in the Li and Shine ozone climatology will have a smaller impact for higher altitudes in the upper mesosphere. Therefore it is unlikely that the tropical cold biases in the upper mesosphere seen under solstice conditions would be caused by deficiencies in the Li and Shine climatology.

As noted above the largest error in the upper mesosphere is the strong warm bias which occurs from ~0.04→0.01 hPa in the summer polar latitudes of each hemisphere. This warm bias would strongly suggest that there is insufficient adiabatic cooling driven by the vertical ascent of the single cell mesospheric circulation of this season. Thus by similar reasoning for the cold winter polar biases of the L50 analyses, the above seasonal warm biases would suggest that wave forcing in the extra-tropical upper mesosphere is too weak. Since the wave forcing of the upper mesosphere is known to be dominated by breaking gravity waves, especially in the summer hemisphere where Rossby wave activity is limited by the zonal wind direction, the above warm polar biases are most likely to be attributed to insufficient small scale gravity wave forcing at upper mesospheric levels in the UM. This argument is supported by evidence in Section 5.2 which shows that the mesospheric zonal winds are too strong when approaching the model lid compared to previous observations. Also consistent with this argument is the evidence presented in Section 7.3.2 which shows that off-line estimates of the gravity wave forcing in the upper mesosphere of the L70 analyses are typically smaller than previous estimates from independent modelling studies.

As noted above the vertical velocity w is set to zero at the model lid. This non-physical boundary condition is likely to have a larger impact on the circulation approaching the L70 model lid, compared to the L50, due to the coarser vertical resolution. Based on similar arguments discussed above for the L50 model the application of a transparent lid in the L70 analyses is also likely to produce an unrealistic meridional circulation in the upper mesosphere of the L70 analyses. It is known from previous studies that gravity waves carry significant values of momentum flux into the thermosphere. Off-line calculations of the momentum flux, see Figure 7.4 in Section 7.2, showed that at the L70 model lid not all of the momentum flux launched in the troposphere had been deposited into the model domain. Therefore it is likely the application of a transparent lid in the L70 analyses will contribute to the underestimated gravity wave forcing suggested by the above temperature biases, and evidence presented in Section 7.3.2.

Again following the reasoning presented above for the L50 analyses, if this bias was entirely related to the misrepresentation of the circulation alone one would expect a cold bias in the opposing winter hemisphere of comparable magnitude. From the EOS MLS and SABER validation results we find that this is not the case and cold biases of the polar winter upper mesosphere have magnitudes of ~2→10 K. This would strongly suggest that approaching the L70 model lid the biases result from a combination of misrepresenting the circulation and a possible latitude independent bias most likely attributed to inaccuracies within the radiation scheme. The magnitude of biases in the polar winter and summer upper mesosphere would strongly suggest that there is a possible latitude independent warm bias for this altitude range in the L70 analyses. This hypothesis is supported by the biases at the model lid for the equinox seasons, where the mesopheric circulation is weaker and one would therefore expect smaller biases to result from the inaccurate representation of small scale forcing, having primarily a warm bias signal with magnitudes of ~2→14 K. This latitide warm

bias results in minimum tropical temperatures of the L70 analyses to be located at ~0.05 hPa compared to ~0.01 hPa for EOS MLS observations, as detailed by Figures 4.8 and 4.9.

One plausible explanation for the above hypothetical warm bias at the L70 model lid is inaccuracies of the radiative heating rates caused by the non-physical nature of the model lid boundary. Here the model lid is assumed to be the top of the atmosphere, where the solar irradiance from the predefined solar spectrum are applied. In reality incident radiation from the sun will be affected in the atmosphere above the L70 model lid due to absorption, emission, reflection and scattering effects whose combined influences are likely to lower the energy of the spectrum before reaching 0.01 hPa. Therefore the modelled solar irradiance values, which have no knowledge of the reduction in energy due to the atmosphere above, will likely contain excessively large values at the model lid. Hence the heating rates at these altitudes could possibly be overestimated resulting in the hypothetical warm biases detailed above. Another possible and likely cause of large biases due to the radiation scheme is the inherent inaccuracy of the spectral files in the upper mesospheric region [Manners, 2010]. This inaccuracy occurs due to the fact that the spectral files were originally optimised for tropospheric conditions, hence they are particularly poor in the upper mesosphere. Any inaccuracy in these spectral files will also be compounded by relatively poor vertical resolution of the upper mesosphere. While it is known to the Met Office that the spectral files in the upper mesosphere are highly questionable, to date no definitive study has been made on their impact [Manners, 2010]. Therefore it is impossible to comment confidently on the magnitude and sign of the temperature biases that they produce. However based on the above evidence it is most likely they contribute to a warm bias in the upper mesosphere.

Approaching the model lid of the L70 analyses, ~80 km, there is a considerable decrease in the atmospheric density. Due to this decrease the mean free

path of molecules becomes increasingly longer and collisions become less frequent, consequently the energy which is absorbed through radiative transfer can be partially re-emitted before being quenched by collision, i.e becoming thermal heat. This is commonly known as classical breakdown of local thermodynamic equilibrium (LTE). In the upper mesosphere the breakdown of LTE conditions can result in a reduction of the total radiative energy absorbed by ozone and molecular oxygen [Fomichev and Shved, 1988, Mlynczak and Solomon, 1993]. As the Edwards-Slingo radiation scheme does not account for the breakdown of LTE it is possible that ozone and oxygen heating rates are over estimated in the upper mesosphere, contributing to warm biases at the L70 model lid.

In the upper mesosphere the breakdown of LTE also impacts the longwave radiative cooling, which is primarily due to carbon dioxide. Recent studies [Coy, 2010] using a high altitude version of the GEOS-5 model [Reinecker et al., 2007] show that accounting for LTE breakdown in the longwave radiation calculations had a significant impact on model temperatures at 0.01 hPa. This study compared the monthly zonal mean temperatures for January 2007 with radiative fluxes calculated from; (a) the NASA climate radiation scheme (CLIRAD) based on LTE assumptions in the shortwave [Chou and Suarez, 1999] and longwave [Chou et al., 2001] regions; (b) independently calculated longwave fluxes including the effects of LTE breakdown using the Fomichev et al. [2004] scheme blended with the CLIRAD longwave values above 75 km. Using the longwave heating rates of (b) the temperatures at 0.01 hPa were ~15→25 K smaller across all latitudes, with maximum differences seen in the tropics and minimum differences seen in the SH summer polar regions. This evidence would suggest that the breakdown of LTE in the longwave region have a non-trivial effect approaching 0.01 hPa, and therefore it is likely that the exclusion of such effects in the Edwards-Slingo radiation scheme contribute to the warm bias seen approaching the L70 model lid.

Region	Season	EOS MLS Bias (K)	SABER Bias (K)	Probable Cause
Tropical (~0.6→0.1 hPa)	NH Winter	22-30	22-30	Ozone Climatology / Bug in Assimilation of AIRS data
	NH Summer	14-22	14-22	
	Equinox	14-22	14-22	
NH Polar (~0.6→0.1 hPa)	Equinox	18-22	22-26	Ozone Climatology / Gravity Wave Forcing
	Winter	38-42	34-38	
	Summer	10-14	10-14	
SH Polar (~0.6→0.1 hPa)	Equinox	14-18	22-26	Ozone Climatology / Gravity Wave Forcing
	Winter	42-50	42-50	
	Summer	10-14	10-14	
Global (~2.0→0.8 hPa)	All Months	6-10	6-10	Spectral Files / Ozone Climatology

Table III: Summary of monthly zonal mean temperature biases for the L50 analyses pre March 2006.

Region	Season	EOS MLS Bias (K)	SABER Bias (K)	Probable Cause
Tropical (~0.4→0.1 hPa)	NH Winter	14-18	14-18	Ozone Climatology
	NH Summer	10-14	10-14	
	Equinox	10-14	10-14	
NH Polar (~0.4→0.1 hPa)	Equinox	18-22	22-26	Ozone Climatology / Gravity Wave Forcing
	Winter	30-38	30-38	
	Summer	10-14	10-14	
SH Polar (~0.4→0.1 hPa)	Equinox	14-18	22-26	Ozone Climatology / Gravity Wave Forcing
	Winter	30-42	30-42	
	Summer	10-14	10-14	
Global (~2.0→0.4 hPa)	All Months	10-14	6-10	Spectral Files / Ozone Climatology
Global (~10→3 hPa)	All Months	2-6	2-6	RTTOV-v7 / Resolution

Table IV: Summary of monthly zonal mean temperature biases for the L50 analyses post March 2006.

4.6.5 Summary

Tables III, IV and V respectively give a summary of the biases seen in the L50 (pre March 2006), L50 (post March 2006) and L70 analyses. Biases are show for both EOS MLS and SABER validation with blue values corresponding to cold biases and red values corresponding to warm biases. Also shown in each table is the most probable cause of each bias, as detailed by the reasoning above.

Region	Season	EOS MLS Bias (K)	SABER Bias (K)	Probable Cause
Tropical (~0.04→0.01 hPa)	NH Winter	2-6	6-14	Spectral Files / Solar Spectrum / Non-LTE Effects
	NH Summer	2-6	2-6	
	Equinox	14-22	2-14	
NH Polar (~0.04→0.01 hPa)	Equinox	6-18	2-14	Gravity Wave Forcing / Spectral Files
	Winter	10-14	10-14	
	Summer	50-66	46-52	
SH Polar (~0.04→0.01 hPa)	Equinox	6-14	6-14	Gravity Wave Forcing / Spectral Files
	Winter	2-14	2-14	
	Summer	70-74	60-64	
Tropical (~0.3→0.08 hPa)	NH Winter	18-22	10-14	Ozone Climatology
	NH Summer	10-14	10-14	
	Equinox	10-14	10-14	
NH Polar (~0.3→0.08 hPa)	Equinox	18-22	18-22	Ozone Climatology / In Winter Data Assimilation
	Winter	50-60	46-56	
	Summer	2-10	2-10	
SH Polar (~0.3→0.08 hPa)	Equinox	6-10	6-10	Ozone Climatology
	Winter	2-10	2-10	
	Summer	2-10	2-10	
Mid-Lat. (~1.0→0.5 hPa)	SH Winter	14-22	14-18	Gravity Wave Forcing
Global (~3.0→0.8 hPa)	All Months	2-10	2-14	Spectral Files / Ozone Climatology

Table V: Summary of monthly zonal mean temperature biases for the L70 analyses.

Chapter 5

Zonal Mean Zonal Wind

This chapter discussed the zonal mean zonal fields of the stratospheric assimilated data set. Due to an inadequate source of independent wind observations in the upper stratosphere and lower mesosphere, it was impossible to obtain a direction validation of the zonal wind analyses. However a qualitative comparison can be made with values from the stratospheric processes and their role in climate (SPARC) climatology. Climatologies of the analyses are constructed for selected periods with the differences from SPARC values discussed.

5.1 L50 Analyses Climatology

Figure 5.1 shows the monthly mean zonal mean wind fields of L50 analyses averaged over 5 years from 2005 to 2009 for January to October and 4 years from 2005 to 2008 for November to December. Also shown are the SPARC climatological values reported on a grid with 41 pressure levels from 1000-0.0046 hPa and a latitudinal resolution of 4^o. These climatological values are obtained from the UARS reference atmosphere project (URAP), detailed in Swinbank and Ortland [2003], which combines data from Met Office analyses and the UARS high resolution Doppler imager (HRDI) [Hays et al., 1993]. Here Met Office analyses are not used above 1.0 hPa, therefore in the mesosphere the SPARC climatology is independent of the L50 analyses. The difference between the averaged values of Figure 5.1 (a) and those from the SPARC climatology are shown in Figure 5.1 (c), here the SPARC values have been linearly interpolated to the analysis locations.

Discussion of the L50 zonal wind values in Figure 5.1 compared to the SPARC climatology follows.

Southern Hemisphere Winter: For June the winter eastward mesospheric jets of the L50 analyses have values between 120-140 ms^{-1} centred at ~40-50^o, extending from ~0.4 hPa to the model lid. This distribution is typically 30-40 ms^{-1} stronger than the SPARC climatology, as shown by Figure 5.1 (b), where jet values decrease in magnitude above ~0.4 hPa. As for June the July analyses jet values are again 30-40 ms^{-1} stronger than SPARC values, however distributions for 2007 and 2008 do show jets that decrease in magnitude at the appropriate altitude and latitude.

Southern Hemisphere Summer: In December the summer westward jets of the L50 analyses have typical values between 60-80 ms^{-1}, which are ~10-20 ms^{-1} stronger than SPARC values. In addition the analysed jets are centred too near the equator at ~30-40^o compared to SPARC latitude values of ~60-70^o. In January the analyses distributions are similar to December, however with magnitudes again typically 10-20 ms^{-1} stronger than SPARC values. For both months the analyses show a very poor representation of the separate jet maximum at the stratopause typical of this season.

Northern Hemisphere Winter: In December the L50 analyses eastward jets have values of ~60-90 ms^{-1} which are typically ~10-20 ms^{-1} stronger than SPARC values and again are not closed off, extending to the model lid. In January analyses jet values are typically ~10 ms^{-1} larger than SPARC values, and occur ~10-20^o too near the pole. An exception to this is the January zonal mean winds of 2006 (field not shown), which has weak westward winds of 10-20 ms^{-1} in the lower mesosphere at polar latitudes. The westward winds seen in the NH winter of January 2006 are associated with the major stratospheric sudden warming (SSW) which occurred during this month [Coy et al., 2009].

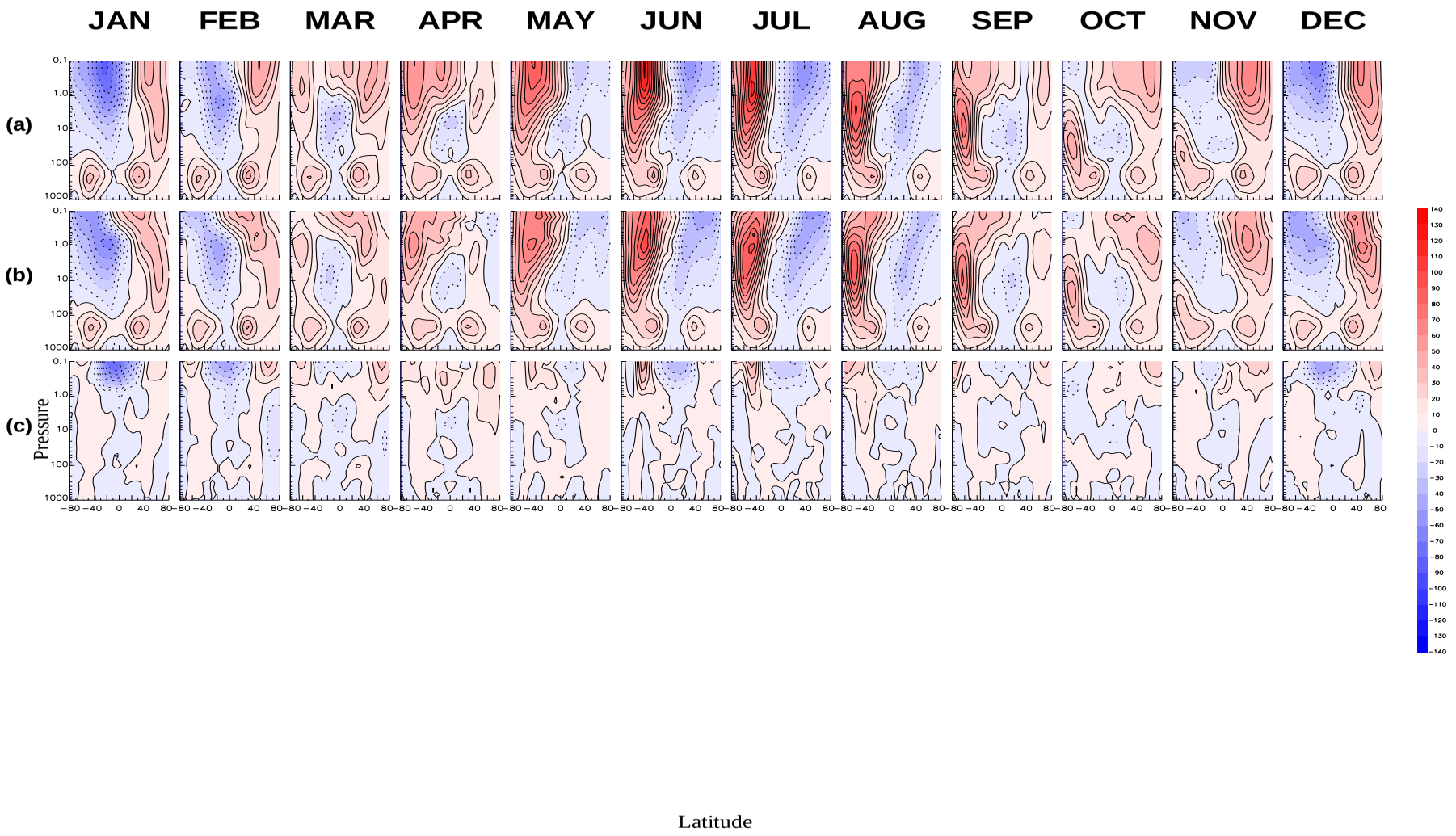

Figure 5.1: Climatological monthly zonal mean zonal wind fields from 2005-2009 for (a) Met Office L50 analyses; (b) SPARC reference atmosphere; (c) Difference between (a) and (b). Wind values have contour intervals of 10 ms^{-1} with negative values indicated by dashed contours.

Northern Hemisphere Summer: For June the westward jets of the analyses have values of 60-70 ms^{-1}, typically ~10-20 ms^{-1} stronger than SPARC values. In July the analyses representation is closer to the SPARC climatology with jet values typically ~0-10 ms^{-1} stronger. For both months the position of the jets is ~10-15^o too near the Equator when compared against SPARC locations.

Equinox Months : For both equinox seasons of each hemisphere the zonal wind distributions of the L50 analyses are qualitatively consistent with those presented in the SPARC climatology. However the eastward winds of the upper mesosphere are typically ~10 ms^{-1} stronger than SPARC values for high latitudes of both hemispheres. This discrepancy is most noticeable approaching the model lid.

For both hemispheres in the lower mesosphere the SPARC climatological eastward winds of the winter season are evident at tropical latitudes of the opposing summer hemisphere. This noticeable feature is not present in the L50 analyses fields as can be seen in Figure 5.1 (c), where the analyses winds for tropical latitudes above ~1.0 hPa during solstice months have a noticeable westward bias compared to SPARC values. To clarify this westward bias Figure 5.2 details the L50 analysis equatorial zonal mean zonal winds in the lower mesosphere averaged over 5 years from 2005 to 2009 for January to October and 4 years from 2005 to 2008 for November to December, with the SPARC climatological values shown for comparison. From Figure 5.2 we find that at from ~1.0→0.6 hPa the L50 equatorial zonal mean zonal winds have a semi-annual oscillation (SAO) signal which is comparable to the SPARC climatology. Here the L50 analyses have a stronger westward phase of the SAO during NH winter months compared to that seen during SH winter, as detailed by the SPARC climatology. However from ~1.0→0.6 hPa the L50 analyses winds during the NH winter season and both equinox seasons are ~10 ms^{-1} stronger than those of the SPARC climatology. However above ~0.6 hPa the westward winds of the solstice phase of the

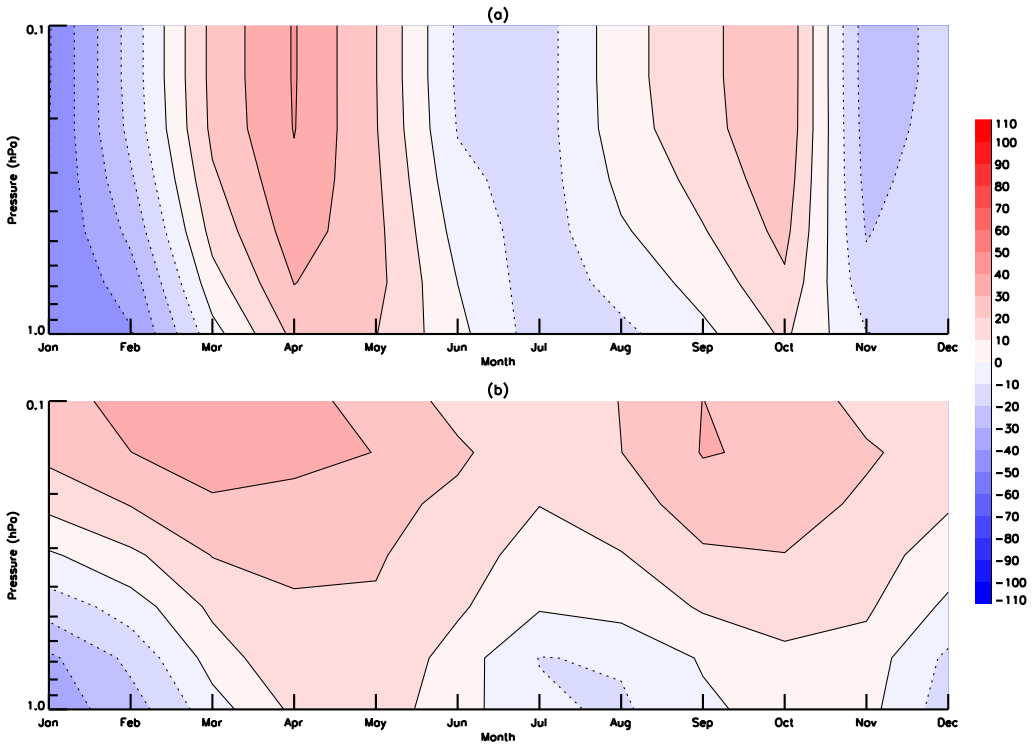

Figure 5.2: (a) Yearly averaged L50 analysis equatorial zonal mean zonal winds for the lower mesosphere; (b) Equatorial zonal mean zonal winds of the SPARC climatology. Contour intervals are 10 ms^{-1} with negative values indicated by dashed contours.

L50 SAO signal extend to 0.1 hPa, while for the SPARC climatology the zonal mean winds of this season are eastward however with magnitudes smaller than those seen for the equinox seasons. Due to uncertainty in the SPARC data and the fact that the strength of the SAO, which is known to have a large degree of inter-annual variability, is determined from numerous physical processes it is impossible to confidently identify a definitive cause of the differences in the SAO signals above ~0.6 hPa.

Additionally the L50 analyses jets show very little horizontal tilt with altitude. This is not consistent with observed jet values which have been shown to tilt towards the equator in winter hemispheres and towards the polar regions for summer hemispheres. The poor representation of this feature is apparent in

many other modern GCM's [Boville, 1995, Langematz and Pawson, 1997, Pawson et al., 2000] and is most prominent in the eastward winter winds of the L50 analyses. This poor representation was also noted in the last climatological study of the stratospheric version of the UM by Jackson et al. [2001], who used Rayleigh friction above 20 hPa. These results suggest that the introduction of the USSP gravity wave scheme has had only a small impact in correcting the orientation of the mesospheric jets, as previously detailed in Scaife et al. [2002]. Furthermore Scaife et al. [2002] suggests that improved equator-ward tilt of the winter jets is primarily due to the amount of forcing supplied and not the forcing mechanism employed.

The above results are consistent with the temperature biases of the lower mesosphere reported in Section 4.3.1. Here the excessive strength of the L50 analyses jets in the lower mesosphere compared to the SPARC climatological values, would again strongly suggest that wave forcing which acts to reduce the zonal flow at these altitudes in both the winter and summer hemispheres is too weak. The largest differences in zonal flow are located in the lower mesosphere, strongly suggesting that this weak wave forcing is primarily due to the misrepresentation of small scale gravity waves by the USSP scheme in the UM.

5.2 L70 Analyses

Figure 5.3 shows the monthly zonal mean zonal wind values for the L70 analyses from November 2009 to September 2010. These wind values are reported on grids which have 31 standard UARS pressure levels and a latitudinal resolution of 2.5^o. For comparison corresponding monthly zonal mean zonal winds from the SPARC climatology, including the upper mesosphere to 0.01 hPa, are shown in Figure 5.4.

Southern Hemisphere Summer: The westward jets for November, December and January have maximum magnitudes of $\sim 90 \rightarrow 100\,\mathrm{ms}^{-1}$ and are $\sim 20 \rightarrow 40\,\mathrm{ms}^{-1}$

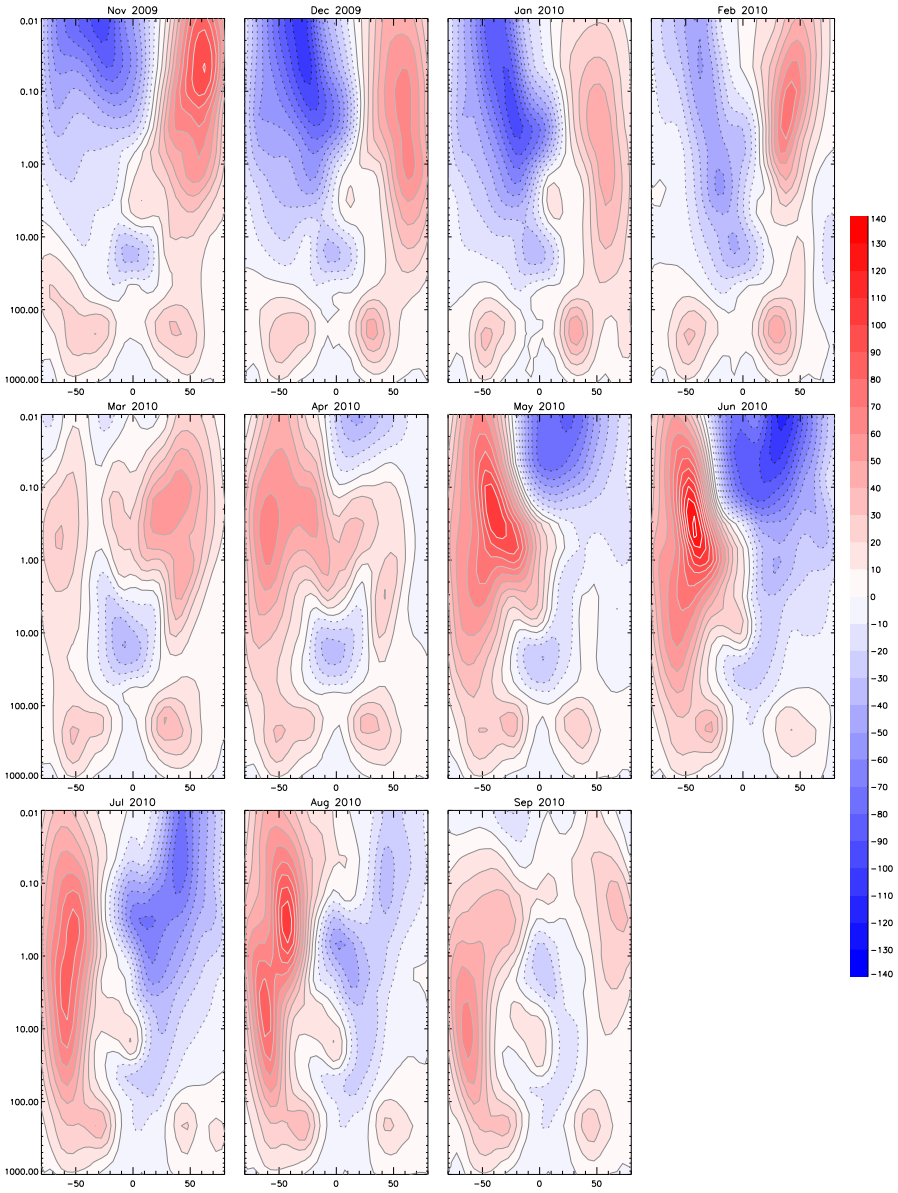

Figure 5.3: Monthly zonal mean zonal winds of the L70 model from November 2009 to September 2010. Wind values have contour intervals of 10 ms^{-1} with negative values indicated by dashed contours.

stronger than SPARC values approaching the model lid. Here the jet centre is located too near the equator and does not decrease in magnitude approaching 0.01 hPa.

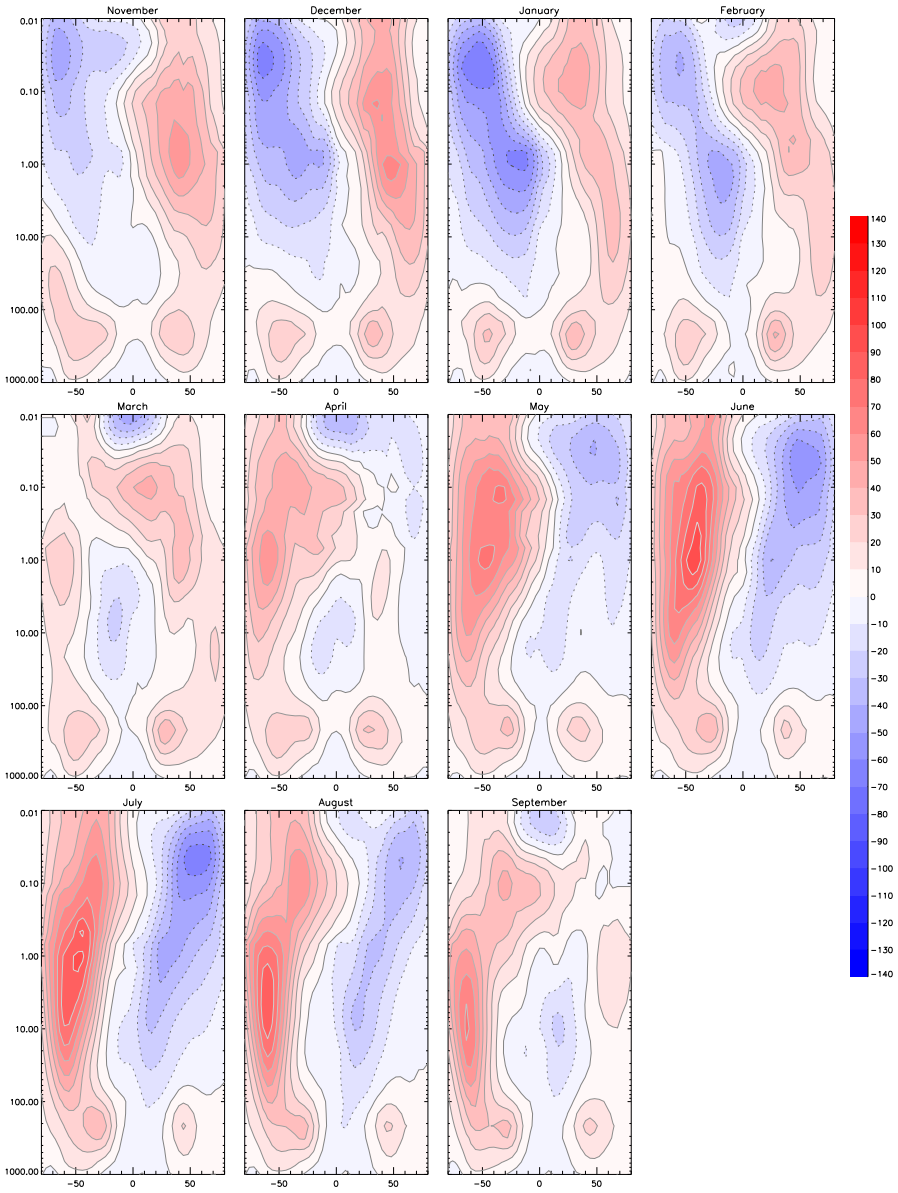

Figure 5.4: Monthly zonal mean zonal winds of the SPARC climatology from November to September. Wind values have contour intervals of 10 ms^{-1} with negative values indicated by dashed contours.

Southern Hemisphere Winter: The largest difference between the L70 eastward winds and those of the SPARC climatolgy occur in the lower mesosphere. Here the jet maximums for May and June of ~100→110 ms^{-1} and ~120→130 ms^{-1} are ~20-30 ms^{-1} stronger than SPARC values and have the opposite tilt

with height. This discrepancy between the L70 analyses and SPARC climatology is less pronounced for July 2010.

Northern Hemisphere Summer: For May and June the westward zonal winds of the L70 analyses have maximum values of ~80→90 ms^{-1} and ~100→110 ms^{-1} respectively, and are typically ~20-40 ms^{-1} stronger SPARC values in the upper mesosphere. As for the SH summer months this discrepancy is most noticeable approaching the model lid where the L70 jets have values which do not decrease in magnitude with altitude.

Northern Hemisphere Winter: For November and December the L70 zonal winds have maximum values in the upper mesosphere of ~80-100 ms^{-1}. These values are noticeably larger and occur at higher altitudes than the SPARC climatology which have maximum values of ~50-60 ms^{-1} in the lower mesosphere.

Equinox Months: For these months the zonal wind distributions of the L70 analyses are qualitatively similar to the SPARC climatology, where difference in magnitude values are generally less than ~10 ms^{-1}.

The above discrepancies of the L70 analyses zonal mean zonal winds when compared to the SPARC climatology are consistent with those expected from thermal wind balance and the temperature structure of the L70 analyses discussed in Section 4.5. As seen when comparing Figures 4.8 and 4.9, the L70 analyses have incorrect meridional temperature gradients at extra-tropical latitudes in the summer upper mesosphere of both hemispheres. Here the increase of temperature towards the polar regions results in an increased westward vertical wind shear and summer mesospheric jets which are not closed off at ~0.01 hPa and located too near the equator.

The uncharacteristic polar vortex of the NH winter hemisphere in November

results in the observed increase of temperature towards the winter pole not occurring until the upper mesosphere. Hence the winter eastward jet of the L70 analyses in November is stronger than the SPARC climaotolgy in the lower mesosphere, and only begins to close when the correct meridional gradients occur above ~0.02 hPa. Since the winter polar vortex of the L70 analyses in SH winter months are more characteristic of this season, the meridional temperature gradients have the correct sign at extra-tropical latitudes resulting in jet maximums which occur at the approximately the correct latitude. However the strength of these jets is too strong due to the excessively large meridional temperature gradients compared to observed values.

The zonal winds for all months increase when approaching the model lid in the upper mesosphere, compared to the decreasing values of the SPARC climatology. This would strongly suggest that the dynamical forcing approaching the model lid is too weak for each hemisphere. As this region of the atmosphere is primarily forced by small scale gravity wave breaking, this would suggest that the momentum deposition supplied by the USSP scheme in the L70 analyses is too weak in the upper mesosphere.

5.3 Summary

This chapter has detailed the zonal mean zonal winds of the L50 and L70 analyses, comparing them against values from the SPARC climatology. For both the L50 and L70 analyses the structure of the zonal mean winds approaching each model lid is consistent with the biases seen in the temperature fields of Section 4. Here approaching the L50 and L70 model lids the zonal mean winds of the solstice seasons are considerably stronger than those of the SPARC climatology, with reduced negative and in some cases positive vertical wind shear in the lower and upper mesosphere. This misrepresentation of the zonal mean jets would again strongly suggest that the zonal gravity wave forcing supplied by the USSP scheme is underestimated in both the L50 and L70 models.

Chapter 6

Gravity Wave Parametrization

This chapter introduces the currently operational ultra simple spectral parameterization (USSP) gravity wave scheme used in the Met Office UM. Full details of how the scheme models the generation, propagation and dissipation of gravity wave momentum flux and hence momentum forcing is given. Such a detailed account is reproduced here to clarify the extension of the USSP scheme to include the direct heating rates from gravity wave dissipation. Such direct heating rates are commonly not included within modern GCM gravity wave schemes. The theory and assumptions behind applying such heating rates within the USSP scheme is also discussed.

6.1 USSP scheme

6.1.1 Background

The currently operational non-orographic gravity wave scheme used by the Met Office is the USSP scheme detailed in Warner and McIntyre [2001]. The USSP scheme uses the spectral approach to gravity wave parameterisation, where momentum deposition is determined from the dissipation of an initial spectrum of gravity waves launched in the troposphere as it propagates vertically through an atmosphere with varying density, static stability and background wind. As for all gravity wave parameterizations used in modern GCM's the

USSP is one-dimensional in the sense that wave propagation only occurs in the vertical direction. Also wave propagation is assumed to be instantaneous over the vertical domain.

The USSP scheme is based on the gravity wave scheme described by Warner and McIntyre [1996], hereafter (WM96), where gravity waves are modelled using a 3D energy spectrum dependent on vertical wavenumber m, the azimuthal angle of propagation ϕ and the intrinsic frequency $\hat{\omega}$. Here the intrinsic frequency is the frequency seen by an observer moving with the background wind. With wavevector $\mathbf{k} = (k, l, -m)$, horizontal projection $\mathbf{k_0} = (k, l, 0) = k_o(cos\phi, sin\phi, 0)$ and $k_o = (k^2 + l^2)^{1/2}$, general wave motion in the WM96 scheme is made specific to gravity waves by the use of the standard dispersion relation, where for vertically propagating gravity waves in a constantly stratified incompressible rotating atmosphere [Gill, 1982].

$$\hat{\omega}^2 = \frac{f^2 m^2}{k_o^2 + m^2} + \frac{N^2 k_o^2}{k_o^2 + m^2} \tag{6.1}$$

Here f is the Coriolis frequency and defining θ as the potential temperature the buoyancy frequency $N = \left(g\frac{\partial(\ln\theta)}{\partial z}\right)^{1/2}$. The WM96 scheme specifies a wave energy spectrum at the launch level identical to the empirical form used by Fritts and Vanzandt [1993].

$$\hat{E}(m, \hat{\omega}, \phi) = E_o A(m) B(\hat{\omega}) \Phi(\phi) \tag{6.2}$$

Here the constituent functions $A(m)$, $B(\hat{\omega})$ and $\Phi(\phi)$ are given by,

$$A(m) = A_o(s, t)\frac{m^s / m_*^{s+1}}{1 + (m/m_*)^{s+t}} \tag{6.3}$$

$$B(\hat{\omega}) = B_o(p)\hat{\omega}^{-p} \tag{6.4}$$

$$\Phi(\phi) = \Phi_o = \frac{1}{2\pi}, \tag{6.5}$$

and following Fritts and Lu [1993],

$$E_o = \frac{\beta N^2}{m^*} \tag{6.6}$$

Here $A_0(s,t)$, $B_o(p)$ and Φ_0 are normalisation coefficients chosen such that $A(m)$, $B(\hat{\omega})$ and $\Phi(\phi)$ all integrate to unity [Fritts and Vanzandt, 1993], full details of these coefficients can be seen in Warner and McIntyre [1996]. Here β is a constant known as the energy scale factor. The value of m_* is the vertical wavenumber m which corresponds to the maximum value of equation (6.3), and is known as the "characteristic" vertical wavenumber. The values of s, t and p are all constants chosen empirically. Both s and t determine the power law dependence for the vertical wavenumber spectrum while p determines the power law dependence for the intrinsic frequency spectrum.

The full 3D WM96 scheme was computationally too expensive for operational use, hence the operational USSP scheme involves making the following simplifications. Firstly following the work of Hines [1997] it is assumed that the Coriolis and non-hydrostatic effects in the wave motion of non-orographic gravity waves may be neglected, allowing the following so called mid-frequency approximation to equation (6.1) to be used.

$$\hat{\omega}^2 = N^2 \frac{k_o^2}{m^2} \tag{6.7}$$

As shown by Hines [1997] and Appendix A of Warner and McIntyre [2001] the use of equation (6.7) allows the $\hat{\omega}$ dependence in the WM96 gravity wave energy spectrum to be integrated out. Therefore vertical propagation of the launch energy spectrum can be modelled in terms of a reduced 2D spectrum depending on m and ϕ only, thus reducing computational expense. The USSP scheme assumes uniform and isotropic azimuthal spectral components at the launch level. Therefore integration of the energy spectrum over piecewise sectors centred on azimuthal directions ϕ_j allow the propagation of the launch energy spectrum to

be described in terms of m only. For convenience the operational USSP scheme uses four azimuthal directions (North, West, South and East), corresponding to $\phi_j = (\pi/2, \pi, 3\pi/2, 0)$.

The second simplification made in the operational USSP scheme follows from the work of Fritts and Lu [1993]. Here the launch energy spectrum for each ϕ direction is approximated by the use of idealised spectral shapes, or parts, which can be integrated analytically. The approximated launch spectrum has two distinct parts defined by the power laws of m, namely (s, t).

Small vertical wave numbers between a low cutoff value of m_{min} and m_* are known as "unsaturated" and have a power dependence $\propto m^s$. Large vertical wavenumbers between m_* and infinity are known as "quasi-saturated" and have a power law dependence $\propto m^{-t}$. Unsaturated parts of the spectrum correspond to gravity waves which are linear, while the quasi-saturated part represents waves which are non-linear due to various processes including wave-breaking and Doppler spreading effects [Warner and McIntyre, 1996]. Here (s, t) are the empirical constants chosen in equation (6.2) defining the initial energy spectrum of the WM96 scheme.

Given knowledge of the spectral characteristics of the initial launch energy spectrum used by the USSP scheme, the spectral density of the horizontal component of the wave pseudo-momentum at a specific altitude z is given by.

$$\hat{\mathbf{p}} = \hat{E}(z, m, \hat{\omega}, \phi)\frac{\mathbf{k_0}}{\hat{\omega}} \tag{6.8}$$

From equation (6.7) it is possible to define the vertical component of group velocity c_{gz} for gravity waves which are described by the mid-frequency approximation as,

$$c_{gz} = -\frac{\partial \hat{\omega}}{\partial m} = \frac{\hat{\omega}}{m}. \tag{6.9}$$

Using equation (6.9) and equation (6.8) it is possible to define the spectral density of the vertical flux of the horizontal component of the wave pseudo-momentum, hereafter for brevity "momentum flux", at an altitude z as.

$$\rho(z)\hat{\mathbf{F}}_{\mathbf{p}}(z, m, \hat{\omega}, \phi) = \rho(z)c_{gz}\hat{\mathbf{p}} = \rho(z)\hat{E}(z, m, \hat{\omega}, \phi)\frac{\hat{\omega}}{N(z)}\hat{\mathbf{k}}_{\mathbf{0}} \tag{6.10}$$

Here $\rho(z)$ is the atmospheric density at altitude z, $\hat{\mathbf{k}}_{\mathbf{0}}$ is the unit vector given by $\mathbf{k_o}/k_o = (\mathbf{i}\cos\phi + \mathbf{j}\sin\phi)$ and the mid-frequency dispersion relation has been used to eliminate k_o. Integration of equation (6.10) gives the total momentum flux at an altitude z as.

$$\rho(z)\mathbf{F_{ptot}}(z) = \rho(z)\int_0^{2\pi}\int_{\hat{\omega}_{min}}^{N}\int_{-\infty}^{\infty}\hat{\mathbf{F}}_{\mathbf{p}}(z, m, \hat{\omega}, \phi)\, dm\, d\hat{\omega}\, d\phi \tag{6.11}$$

Here $\hat{\omega}_{min}$ is the lower limit of the normalised intrinsic frequency spectrum, equation (6.4), and is fully detailed in Warner and McIntyre [1996]. From equation (6.11) the wave induced force on the back ground flow at an altitude z is given by.

$$\mathbf{G}(z) = -\frac{1}{\rho(z)}\frac{\partial\rho(z)\mathbf{F_{ptot}}}{\partial z} \tag{6.12}$$

The initial launch spectrum of momentum flux represents the sole source of gravity waves in the USSP scheme. Therefore the total momentum flux and hence wave induced force at higher altitudes can be calculated given prior knowledge of the launch spectrum and it's modification when propagating in the vertical direction.

6.1.2 Doppler Shifting

The propagation characteristics of the momentum flux launch spectrum are significantly altered when propagation occurs through an atmosphere with vertical wind shear, i.e., $\partial U/\partial z \neq 0$. Here the intrinsic frequency of a spectral element is related to its absolute frequency ω_0 relative to the ground by the Doppler shift relation.

$$\hat{\omega} = \omega_o - \mathbf{k}_o \cdot \mathbf{U}(z) = \omega_o - k_o U(z) \tag{6.13}$$

Here $\mathbf{U}(z)$ is the horizontal wind velocity and $U(z)$ is the component of the background wind in the $\mathbf{k}_o$ direction. Both ω_0 and k_0 are invariant under propagation in a steady horizontally homogeneous background flow [Lighthill, 1978]. Therefore a wave of given absolute frequency will have a varying intrinsic frequency as it ascends through an atmosphere with vertical wind shear. We shall adopt the abbreviation $\hat{\omega}_l = \hat{\omega}(z_l)$ for the intrinsic frequency at the launch altitude z_l and $\hat{\omega}_k = \hat{\omega}(z_k)$ for the intrinsic frequency at a higher altitude z_k, and similar notation for other z-dependent quantities such as m, N, ρ and U. Then using equation (6.13), propagation from altitude z_l to z_k results in the following expression for the change in intrinsic frequency.

$$\hat{\omega}_l - \hat{\omega}_k = k_o(U_k - U_l) \tag{6.14}$$

Using the mid-frequency dispersion equation (6.7) to substitute for $\hat{\omega}_l$ and $\hat{\omega}_k$ we obtain.

$$\frac{N_l}{m_l} - \frac{N_k}{m_k} = (U_k - U_l) \tag{6.15}$$

Rearranging equation (6.15) we obtain the following expression for how a given vertical wavenumber is Doppler shifted with height.

$$m_k = \frac{N_k m_l}{N_l - m_l[U_k - U_l]} \tag{6.16}$$

Note that equation (6.16) is independent of the horizontal wavenumber k_0, which allows the reduction of the full 3D spectrum used by WM96 to the 2D spectrum used by the USSP. An important property of the USSP is the ability to predict the evolution of a spectrum when conservatively propagated in the vertical direction. This requires knowledge of how to transform spectral densities that depend on vertical wavenumber.

$$\mathbf{f}(z_k, m_k, \phi_j) = \mathbf{f}(z_l, m_l, \phi_j)\frac{dm_l}{dm_k} \tag{6.17}$$

Here the dm_l/dm_k term can be expressed by rearranging (6.16) for m_l and differentiating with respect to m_k, obtaining the following.

$$\begin{aligned}\mathbf{f}(z_k, m_k, \phi_j) &= \mathbf{f}(z_l, m_l, \phi_j)\,\frac{N_l N_k}{[N_k + m_k(U_k - U_l)]^2} \\ &= \mathbf{f}(z_l, m_l, \phi_j)\,\frac{N_l N_k}{\left[N_k + m_k\left(\frac{N_l}{m_l} - \frac{N_k}{m_k}\right)\right]^2} \\ &= \mathbf{f}(z_l, m_l, \phi_j)\,\frac{N_k}{N_l}\left(\frac{m_l}{m_k}\right)^2 \end{aligned} \tag{6.18}$$

6.1.3 Launch Spectrum

Following Bushell et al. [2007] the momentum flux of each vertical wavenumber spectral element in the ϕ_j direction at a launch altitude of z_l is given by,

$$\begin{aligned}\rho_l \mathbf{F_{pl}}(z_l, m_l, \phi_j) &= \rho_l \int_0^{2\pi} \int_{\hat{\omega}_{min}}^{N} \hat{\mathbf{F}}_\mathbf{p}(z_l, m_l, \hat{\omega}, \phi)\, d\hat{\omega}\, d\phi \\ &= \rho_l \left(\tfrac{A_0}{m_*^2}\right) C_{l0} \hat{\mathbf{k}}_\mathbf{0}(\phi_j) \begin{cases} m_l^s / m_*^{s+1} & \ldots \quad m_{min} \le m_l < m_* \\ m_*^{t-1} / m_l^t & \ldots \quad m_* \le m_l < \infty \end{cases}\end{aligned} \tag{6.19}$$

Here C_{l0} is a constant isotropic scale factor, which is independent of location around the globe and equal to.

$$C_{l0} = \frac{\beta}{\pi} \sin(\Delta\phi/2) \left(\frac{p-1}{2-p}\right) \left[N_l^2 \frac{(\hat{\omega}_{min}/N_l)^{p-1}[1 - (\hat{\omega}_{min}/N_l)^{2-p}]}{[1 - (\hat{\omega}_{min}/N_l)^{p-1}]}\right] \tag{6.20}$$

The terms in equation (6.20) are obtained by the integration of equations (6.4) and (6.5) over their respective ranges, full details are discussed in Appendix A of Bushell et al. [2007]. The normalisation constant for the vertical wavenumber spectrum in equation (6.19) is given by.

$$A_o(s,t) = [(s+t)/\pi] \sin[\pi(s+1)/(s+t)] \tag{6.21}$$

Integration of equation (6.19) over the complete vertical wavenumber spectrum gives the total launch momentum flux in each ϕ_j direction as.

$$\begin{aligned}\rho_l \mathbf{F_{pltot}}(z_l, \phi_j) &= \rho_l \left(\frac{A_0}{m_*^2}\right) C_{l0} \left[\int_{m_{min}}^{m_*} \frac{m^s}{m_*^{s+1}} dm + \int_{m_*}^{\infty} \frac{m_*^{t+1}}{m^t} dm\right] \hat{\mathbf{k}}_\mathbf{0}(\phi_j) \\ &= \rho_l C_{l0} m_*^{-2} \hat{\mathbf{k}}_\mathbf{0}(\phi_j)\end{aligned} \tag{6.22}$$

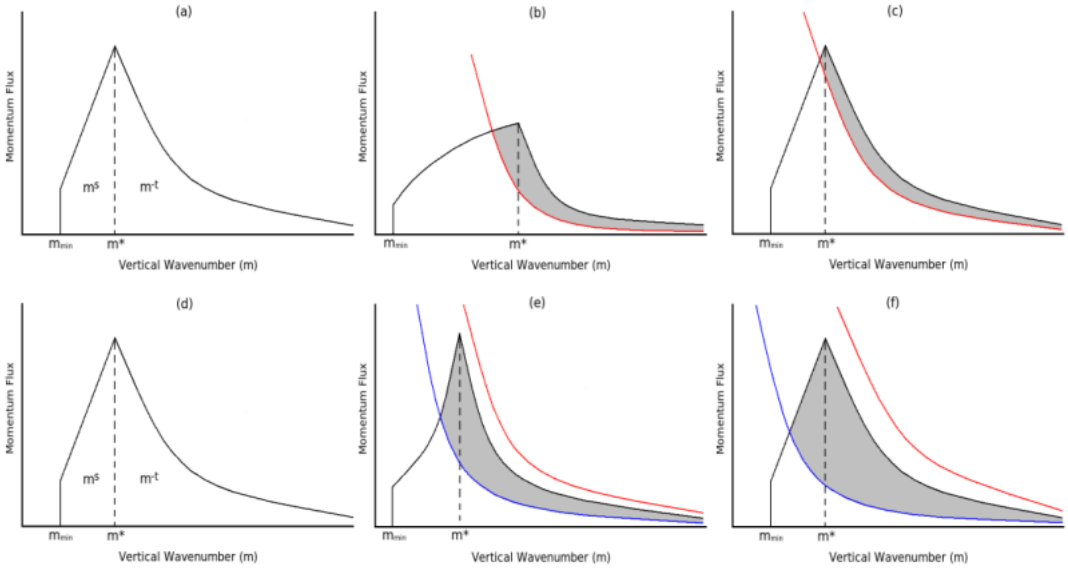

Figure 6.1: Schematic of USSP momentum flux spectral profile evolution. (a) Launch flux spectrum in the eastward direction at z_l; (b) Evolved eastward spectrum at z_k having propagated through positive wind shear; (c) Evolved eastward spectrum back-propagated to launch altitude z_l; (d) Launch flux spectrum in the westward direction at z_l; (e) Evolved westward spectrum at z_k having propagated through negative wind shear and (f) Evolved westward spectrum back-propagated to launch altitude z_l. Red line represents the saturation chop function at z_k, blue line represents saturation chop function at altitude greater than z_k. Grey shading represents portion of spectrum which is eroded.

The spectral representation of momentum flux for a typical launch profile can be seen in Figure 6.1 (a) and (d) for eastward and westward directions respectively. Here the low wavenumber and high wavenumber parts of the spectrum are clearly separated at the characteristic vertical wavenumber m^*. This approach results in each individual vertical wavenumber spectral elements having an associated momentum flux value. Under the mid-frequency approximation (equation 6.7) the vertical wave-number m can be very simply related to the hori-

zontal intrinsic phase speed $\hat{c}_h$ and background stability N [Fritts and Alexander, 2003].

$$|m| = \frac{N}{\hat{c}_h} \tag{6.23}$$

Here the horizontal intrinsic phase speed $\hat{c}_h = c_h - \bar{u}_h$, where c_h is the horizontal phase speed relative to the ground and $\bar{u}_h$ is the horizontal background flow. Thus the spectral approach used in the USSP scheme launches gravity waves with a continuous range of phase speeds, compared to discrete schemes such as that described in Lindzen [1981] which launch gravity waves with specific and finite phase speeds.

6.1.4 Conservative Propagation and the Quasi-Saturation Spectrum

Applying the conservative propagation transformation (6.18) to the momentum flux at the launch level z_l, we obtain the following expression for the conservatively propagated spectrum at altitude z_k.

$$\rho_k \mathbf{F_p}(z_k, m_k, \phi_j) = \rho_l \mathbf{F_p}(z_l, m_l, \phi_j) \frac{[N_l - m_l(U_k - U_l)]^2}{N_l N_k} \tag{6.24}$$

Here (6.16) has been used to substitute for m_k. In the absence of wave breaking the launch momentum flux spectrum will evolve according to equation (6.24), with the wavenumber limits m_{min} and m_* transforming according to (6.16). The schematic representation of this can be seen in Figure 6.1 (b) and (e) where the eastward (westward) launch flux profiles are assumed to have been vertically propagated through positive (negative) vertical wind sheer, i.e. increasing eastward flow with height where $U_k > U_l$.

Having conservatively propagated the launch momentum flux spectrum from altitude z_l to z_k, growth at the large m part of the spectrum is limited by wave

breaking. Wave breaking is modelled in the USSP scheme by imposing a "chopping function" or quasi-saturation spectrum, which has an identical form to the initial large m part of the launch spectrum and is $\propto m^{-t}$. Such a chopping function is applied as it brings empirical knowledge about the spectral characteristics of non-linear gravity wave-breaking [Warner and McIntyre, 1996], i.e. that the spectral slope is limited to the power law t. While there is no widely accepted theory for what determines the value of t, however observational evidence has show that for non-linear processes it is typically between ~2.5→3, with a value of 3 usually adopted in modern gravity wave schemes. Applying this chopping function we therefore require that the momentum flux spectrum propagated from the launch level to altitude z_k satisfies,

$$\rho_k \mathbf{F_p}(z_k, m_k, \phi_j) \leq \rho_k \mathbf{F_{pS}}(z_k, m_k, \phi_j) = \rho_k \left(\frac{A_0}{m_*^2}\right) C_{Sk} \frac{m_*^{t-1}}{m_k^t} \hat{\mathbf{k}}_\mathbf{0}(\phi_0) \tag{6.25}$$

where,

$$C_{Sk} \propto \frac{\beta}{\pi} \sin(\Delta\phi/2) \left(\frac{p-1}{2-p}\right) \left[N_k^2 \frac{(\hat{\omega}_{min}/N_k)^{p-1}[1-(\hat{\omega}_{min}/N_k)^{2-p}]}{[1-(\hat{\omega}_{min}/N_k)^{p-1}]} \right] \tag{6.26}$$

Note that the quasi-saturation chopping function is applied at the higher altitude z_k, where it has a simple spectral shape with power law dependence $\propto m^{-3}$. After conservative propagation and application of the appropriate chopping function the final spectrum at altitude z_k is known as the "evolved spectrum". From Figure 6.1 (b) we find that for the eastward launch flux profile, which has propagated through positive vertical shear, a portion of the evolved spectrum is transformed via equation (6.24) above the chopping function at altitude z_k denoted by the red line. As noted in Warner and McIntyre [2001] the sign convention of the USSP scheme results in positive wind shear corresponding to an increasing likelihood of wave breaking, i.e. $m \to \infty$. Here the momentum flux eroded from the launch spectrum at z_k is denoted by the gray shading. The opposite is true for the westward launch flux profile, which when propagating through negative wind shear is transformed away from wave breaking, as show by Figure 6.1 (e)

where the evolved spectrum is entirely below the saturation chopping function at altitude z_k (again denoted by the red line). However, since the saturation chopping function depends on the density (which dominates over the stability N) at higher chopping levels, the decrease in density with altitude results in an increasing likelihood of wave breaking. This is show in Figure 6.1 (e) where a chopping function defined at an altitude higher (denoted by the blue line) than z_k is able to erode the evolved spectrum even tho it has been transformed towards lower vertical wavenumbers.

6.1.5 Numerical Back-propagation

In order to calculate how momentum flux is deposited or propagated we need to know the impact on the launch spectrum of applying the quasi-saturation chopping function at altitude z_k. Here the simplest and most computationally cheap method is to conservatively back-propagate the chopping function at z_k to the launch level. Back-propagation can be viewed as the inverse of forward propagation in (6.16), where we are concerned with what initial spectrum at z_l is required to produce the appropriate evolved spectrum at z_k. Applying the conservative propagation transformation (6.18) to the quasi-saturation chopping function (6.25), after expressing m_k in terms of m_l by (6.16) and rearranging for the back propagated spectrum at z_l we obtain.

$$\rho_l \mathbf{F}^{\mathbf{b}}_{\mathbf{pS}}(z_l, m_l, \phi_j) = \frac{N_l N_k}{[N_1 - m_1(U_2 - U_1)]^2}\, \rho_k \mathbf{F}_{\mathbf{pS}}(z_k, m_k, \phi_j)$$

$$= \frac{N_l N_k}{[N_1 - m_1(U_2 - U_1)]^2}\, \rho_k \left(\frac{A_0}{m_*^2}\right) C_{Sk} \frac{m_*^{t-1}}{m_k^t} \hat{\mathbf{k}}_{\mathbf{0}}(\phi_0) \tag{6.27}$$

Here the superscript **b** is notation for the back-propagated spectrum. Again applying m_k in terms of m_l by (6.16) we obtain.

$$\rho_l \mathbf{F}^{\mathbf{b}}_{\mathbf{pS}}(z_l, m_l, \phi_j) = \frac{N_l N_k}{[N_1 - m_1(U_2 - U_1)]^2}\, \frac{[N_l - m_l(U_k - U_l)]^t}{N_k^t}\, \rho_k \left(\frac{A_0}{m_*^2}\right) C_{Sk} \frac{m_*^{t-1}}{m_l^t} \hat{\mathbf{k}}_{\mathbf{0}}(\phi_0)$$

$$= \frac{N_l}{[N_k]^{t-1}}[N_l - m_l(U_k - U_l)]^{t-2}\,\rho_k\left(\frac{A_0}{m_*^2}\right)C_{Sk}\frac{m_*^{t-1}}{m_l^t}\hat{\mathbf{k}}_{\mathbf{0}}(\phi_0) \tag{6.28}$$

Therefore with knowledge of U_k, ρ_k and N_k we have an expression for the back-propagated spectrum in terms of launch vertical wavenumbers m_l only. The back propagated evolved spectrums of Figure 6.1 (b) and (e) can be seen in Figures 6.1 (c) and (f). Here the portion of the launch flux profile, determined by the procedure described below, is again denoted by the gray shading.

6.1.6 Momentum Flux Deposition

Erosion of the initial launch momentum flux and hence deposited or propagated momentum flux at z_k is determined by calculating which parts of the launch spectrum will be subjected to chopping at z_k. This calculation is based on the point where the back-propagated quasi-saturated chopping function and launch spectra intersect at altitude z_l.

$$\rho_l\mathbf{F_{pl}}(z_l, m_l, \phi_j) = \rho_l\mathbf{F^b_{pS}}(z_l, m_l, \phi_j) \tag{6.29}$$

Defining the intersection point of (6.29) at z_l as m_{lx} with normalised value $m_{Nlx} = m_{lx}/m_*$, combining (6.19) and (6.28) we obtain.

$$\rho_k C_{Sk}\, m_{Nlx}^{-t}\frac{N_l}{[N_k]^{t-1}}\,[N_l - m_{lx}(U_k - U_l)]^{t-2} = \rho_l C_{l0}\begin{cases} m_{Nlx}^{s} & \ldots \quad m_{min}/m_* \le m_{Nlx} < 1 \\ m_{Nlx}^{-t} & \ldots \quad 1 \le m_{Nlx} < \infty \end{cases}$$

Rearranging gives,

$$\implies \rho_l C_{l0}\, B\, m_{Nlx}^{-t}\left(1 + \frac{A}{B}m_{Nlx}\right) = \rho_l C_{l0}\begin{cases} m_{Nlx}^{s} & \ldots \quad m_{min}/m_* \le m_{Nlx} < 1 \\ m_{Nlx}^{-t} & \ldots \quad 1 \le m_{Nlx} < \infty \end{cases} \tag{6.30}$$

Here,

$$\frac{A}{B} = \frac{m_*(U_l - U_k)}{N_l} \tag{6.31}$$

$$B = \frac{\rho_k C_{Sk}}{\rho_l C_{l0}} \left[\frac{N_l}{N_k}\right]^{t-1}. \tag{6.32}$$

Defining m_{NlX} and m_{NlY} respectively as intersect points in the low and high vertical wavenumber parts of the spectrum, we have the following solutions for (6.30).

$$\begin{aligned} m_{NlX}^{s+t} - B\left[1 + \tfrac{A}{B} m_{NlX}\right]^{t-2} = 0 \quad m_{min}/m_* \leq m_{NlX} < 1 \\ B\left[1 + \tfrac{A}{B} m_{NlY}\right]^{t-2} = 1 \quad 1 \leq m_{NlY} < \infty \end{aligned} \tag{6.33}$$

From (6.33) the solution for the high wavenumber intersect is obtained by.

$$m_{NlY} = \frac{1}{A}\left[B^{\frac{t-3}{t-2}} - B\right] \tag{6.34}$$

However the solution for the low wavenumber intersect m_{NlX} often requires an iterative solution. Full details of this iterative solution and the method used can be found in Bushell et al. [2007]. Note once found both m_{NlX} and m_{NlY} can respectively be mapped to m_{NkX} and m_{NkX} at altitude z_k via (6.16), as can the values of m_{min} and m_*. It should be noted that the intersect points can exist in either or both of the low and high wavenumber number regions, resulting in the different spectral chop type discussed in detail below.

Once m_{NlX} and m_{NlY} have been obtained the USSP scheme then calculates the total amount of momentum flux for each z_k level in a similar fashion to (6.22). However these calculations are simplified by evaluating the chopped portion of the spectrum at z_k, where it has a simple spectral shape, while the remaining conservatively propagating portion of the spectrum is obtained from the original launch spectrum at z_l.

Since the momentum flux spectrum at the launch level represents the sole source of gravity waves, i.e. no further generation of gravity waves above z_l, the vertically propagating momentum flux is restricted to never exceed the momentum

flux from the level beneath. Although the USSP scheme does not erode the launch spectrum in a level by level sequential manor, the shielding of upper levels from momentum flux deposited lower down does employ this concept in an integrated sense.

The extent to which the back-propagated chopping function can erode the initial launch spectrum is highly dependent on both the density and the strength of the Doppler shift between altitude levels. The relative velocity between levels z_l and z_k appears in (6.31). Using the sign of (6.31) and the fact that the $B\left[1+\frac{A}{B}\right]^{t-2}$ term in (6.33) must be less than one for the m_{NlX} intersect to exist, the USSP scheme is able to define the type of chopping function before the intersect points are calculated improving computational expense. In total there are five different chop types based on the above conditions, with some chop types having individual cases where chopping is limited to a specific region of the launch spectrum. For each case the momentum flux at each z_k level is defined as a fraction of the momentum flux at the launch level. Table VI details the conditions and definitions of momentum flux at z_k for each chop type case. Once the momentum flux in each ϕ_j direction has been calculated for all z_k levels in the vertical profile, the wave induced force at each level can be calculated via (6.12). Full details and the derivation of the momentum flux values for each chop type in Table VI can be seen in Appendix A.

6.1.7 Model Lid Condition

Remaining momentum flux at the level corresponding to the model lid is allowed one of two options. The first option is a transparent lid condition, where the remaining momentum flux is allowed to propagate through the model lid boundary. The second option is an opaque lid condition, where all remaining momentum flux is confined to the model domain by requiring that the momentum flux value at the model lid is always zero.

Chop-type	A/B	$B[1+A/B]^{t-2}$	Momentum Flux at z_k ($\rho_k \mathbf{F_{pktot}}$)
0	≥ 0	≥ 0	$\rho_{k-1}\mathbf{F_{p(k-1)tot}}$
I	< 0	≥ 0	$\rho_l \mathbf{F_{pltot}} \left[1 + \frac{A_0}{t-1}\frac{A}{B} m_{NlY}^{2-t}\right]$
IIa	≤ 0	< 1	$\frac{\rho_l \mathbf{F_{pltot}} A_0}{(s+1)(t-1)} \left[\left(s+t+(s+1)\frac{A}{B} m_{NlX}\right) m_{NlX}^{s+1} - (t-1) m_{Nlmin}^{s+1}\right]$
IIb ($m_{min} \geq m_{lX}$)	≤ 0	< 1	$\frac{\rho_l \mathbf{F_{pltot}} A_0}{(t-1)} \left[\frac{B}{m_{Nlmin}^{t-1}} \left(1 + \frac{A}{B} m_{Nlmin}\right)^{t-1} - \frac{A}{B} m_{NlY}^{2-t}\right]$
IIc ($A = 0$)	≤ 0	< 1	$\frac{\rho_l \mathbf{F_{pltot}} A_0}{(s+1)(t-1)} \left[(s+t) m_{NlX}^{s+1} - (t-1) m_{Nlmin}^{s+1}\right]$
IIIa	> 0	< 1	$\frac{\rho_l \mathbf{F_{pltot}}}{(s+1)(t-1)} \left[-(t-1) m_{Nlmin}^{s+1} + \left(s+t+(s+1)\frac{A}{B} m_{NlX}\right) m_{NlX}^{s+1} - (s+1)\frac{A}{B} m_{NlY}^{2-t}\right]$
IIIb ($m_{Nlmin} \geq m_{lX}$)	> 0	< 1	$\frac{\rho_l \mathbf{F_{pltot}} A_0}{(t-1)} \left[\frac{B}{m_{Nlmin}^{t-1}} \left(1 + \frac{A}{B} m_{Nlmin}\right)^{t-1} - \frac{A}{B} m_{NlY}^{2-t}\right]$
IV ($\lvert B/A \rvert \leq m_{Nlmin}$)	< 0		$\rho_k \mathbf{F_{pktot}} = 0$

Table VI: Summary of the chop types used in determination of momentum flux profiles in the USSP gravity wave scheme. Shown are the conditions and the definition of the momentum flux at z_k for each individual case.

In an ideal situation the model domain would extend to an altitude where the momentum flux remaining in the model is negligible. For obvious reasons this is not possible and a choice of opaque or transparent lid condition is needed. Since the launch momentum flux is isotropic in each azimuthal direction, the net momentum flux introduced at the launch level is zero. Theoretically the opaque lid condition is preferred as it prevents any anisotropy of momentum flux leaving the model lid and consequently results in momentum being conserved throughout the model domain. The conservation of momentum throughout the model domain has been shown by Shepherd and Shaw [2004] to avoid any spurious influence of radiative perturbations in the mesosphere on the atmosphere below through gravity wave drag feedbacks. As shown by Shepherd and Shaw [2004] such feedbacks predominantly have an important impact over seasonal time scales and hence do not constrain the choice of model lid condition for NWP forecasts. However since above ~1.0 hPa the L50 model is essentially free running, no observational data is assimilated, it is possible that the gravity wave feedbacks noted in Shepherd and Shaw [2004] could still apply. Thus the opaque lid condition is theoretically preferred for the L50 model. However, as noted by Scaife et al. [2002] when modelled gravity waves approaching the model lid contain significant amounts of momentum flux, the requirement of depositing all momentum at this level can lead to model instability. For this reason the transparent lid condition is applied in creation of the L50 analyses [Bushell et al., 2007], where due to the altitude of the model domain being located in the lower mesosphere large portions of the gravity wave spectrum are unsaturated and carry significant momentum flux at the model lid. It should be noted that the transparent lid condition is also applied in the L70 analyses. However the higher altitude of the L70 model lid results in larger proportions of the launch momentum flux being deposited throughout the mesosphere, hence the issue of numerical instability when applying the opaque lid condition is less severe for the L70 model. As such the influence of applying this condition is investigated in the UM experiments described in Chapter 8.

Name	Description	Value
p	Intrinsic Frequency Spectrum Power Index	5/3
β	Energy Scale Factor	0.10227987
m_*	Characteristic vertical wavenumber at launch level	$2\pi/(4300)m^{-1}$
$\Delta\phi$	Angle of Azimuthal Sector	$\pi/2$
m_{min}	Minimum Vertical Wavenumber	$2\pi/(20\times10^3)m^{-1}$
s	Low Vertical Wavenumber Spectrum Power Index	1
t	High Vertical Wavenumber Spectrum Power Index	3
C_{l0}	Launch Spectrum Scale Factor	$3.4191\times10^{-9}s^{-2}$

Table VII: Summary of adjustable parameters in the USSP gravity wave scheme, with given standard values.

6.1.8 Standard Parameters Values

As detailed above the USSP scheme has several parameters which can be freely adjusted. Table 6.1.7 describes these parameters and states the standard values used since the introduction of the USSP in October 2003, as detailed in Bushell et al. [2007]. These standard parameters were tuned by Scaife et al. [2000, 2002] to give a stratospheric quasi-biannual oscillation (QBO) signal in the model which was realistic, specifically this was achieved by tuning the launch spectrum scale factor and the characteristic vertical wavenumber using an idealised one dimensional model of zonal wind [Scaife et al., 2000]. It was found that the amplitude of the wind oscillation is largely determined by the characteristic vertical wavenumber value while the period depends on both parameters. Here

sensitivity tests suggested that only one set of parameters would give a realistic QBO signal, namely $C_{l0} = 3.4191 \times 10^{-9} s^{-2}$ and $m_* = 2\pi/(4300) m^{-1}$, thus defining the value of momentum flux in each azimuthal direction at the launch level through equation (6.22) to values which give appropriate forcing and period of the QBO for tropical latitudes.

In addition to those described in Table 6.1.7, the final parameter in the USSP scheme is the choice of the initial launch spectrum altitude level. As described in Section 3.1.3 the UM uses a height based vertical coordinate. The launch altitude of the initial spectrum is "hard wired" to be located at the model theta level which is closest to the value of $\eta = 0.045$ [Bushell et al., 2007], corresponding to launching the spectrum close to the Earth's surface in the lower troposphere.

For the L50 analyses this hard wired value results in the initial spectrum being launched at the $\eta = 0.0457734$ model level, corresponding to ~2880 m geopotential height. For the L70 analyses the hard wired value results in the initial spectrum being launched at the $\eta = 0.044$ model level, corresponding to ~3520 m geopotential height.

6.2 Direct Heating Rates due to Gravity Wave Dissipation

As detailed in previous chapters breaking gravity waves are known to force the mean circulation and thermal structure of the middle atmosphere, a process which involves the large scale transport of heat, momentum and mass [Holton and Alexander, 2000]. As well as this large scale forcing gravity waves can also cause localised small scale turbulent mixing and transport of the above quantities [Fritts and Werne, 2000]. The generation of such localised turbulence can result in a direct heating of the atmosphere [Becker, 2004].

Previous studies have shown that direct heating rates due to gravity wave breaking have magnitudes of the order 10^{-1} K day^{-1} below ~65 km. Such magnitudes represent an insignificant contribution to the overall heating budget of the lower mesosphere and stratosphere, hence they were not included for the USSP scheme used in the L50 analyses. In the upper mesosphere direct heating from gravity wave breaking does contribute significantly to the overall heating budget [Lübken, 1997], however most middle atmospheric GCM's do not include this contribution. This is true of the L70 analyses which only employ the momentum forcing from the USSP code [Bushell et al., 2007]. Thus incorporating additional heating rates from gravity wave breaking further enhances the realism of physical processes modelled by the USSP scheme. As such the impact of employing direct heating rates due to gravity wave breaking within the UM is investigated in Chapter 8. The theoretical approach that we have applied to determine such direct heating rates within the operational USSP code is described below.

6.2.1 Turbulent Diffusion and Dissipation

The idea that breaking gravity waves are able to generate turbulence was first suggested by Hodges [1969]. Gravity waves are believed to generate localised turbulence if they become unstable under timescales sufficiently shorter than those describing wave propagation [Fritts and Alexander, 2003]. The primary mechanisms for turbulent generation occurs through wave-breaking due to both Kelvin-Helmholtz (KH) shear and convective instability. The onset of turbulence acts to restore the wave stability and by limiting (damping) the amplitude growth of vertically propagating waves results in a convergence of vertical momentum flux [Holton, 1982]. Thus, gravity wave instability can generate both turbulence and a net forcing of the mean flow. While these processes are ultimately related they are however distinct, thus knowledge of gravity wave drag can not be obtained through approximating the effects of turbulence on the mean flow [Lindzen, 1981]. Hence both sub-grid processes must be parameterised simultaneously in middle atmospheric GCM's.

Turbulence generated from gravity wave instability occurs over far smaller scales than the gravity wave motions themselves, and has an assumed characteristic "inertial range". The upper bound of this inertial range is defined as the turbulence scale not influenced by the background static stability, known as the "buoyancy" or "outer" scale and is given by $L_B \propto (\epsilon/N^3)^{1/2}$ [Weinstock, 1981]. Here ϵ is the turbulent energy dissipation rate and is discussed in further detail below. At the smallest scales the inertial range of turbulence is constrained by viscosity μ with an "inner" scale given by $l_o \propto (\upsilon^3/\epsilon)^{1/4}$, here $\upsilon = (\mu/\rho)$ is known as the kinematic viscosity.

While a complete description of turbulence is beyond our current knowledge, numerical simulations are typically based on the general characteristics that all turbulent flows posses. These characteristics include unpredictability, diffusion, fluctuating values of vorticity and the dissipation of kinetic energy [Roberts and Webster, 2002].

Turbulence is unpredictable in the sense that fluctuations of velocity within the flow can not be determined from knowledge of an instantaneously observed velocity at a particular time, i.e. velocity fluctuations are irregular or chaotic. These irregular velocity fluctuations in turbulent flow produce local mixing and transport of heat, momentum and mass commonly known as turbulent diffusion. In the stratosphere and mesosphere turbulent diffusion dominates (by ~1-2 orders of magnitude) over molecular diffusion, which is generally not an important mechanism until the turbopause at ~110 km where the exponential growth of kinematic viscosity allows it to dominate. Small scale mixing and transport, commonly known as eddy diffusion, is generally parameterized in middle atmospheric models through a quantity known as the eddy diffusion coefficient. Here, the turbulent diffusion of heat can act to both cool and heat the atmosphere.

Accompanying the deposition of gravity wave momentum there is also a deposition of gravity wave energy into the mean flow at the onset of wave breaking. As shown by Lindzen [1973] this energy deposition can lead to both increased heat and kinetic energy of the mean flow. The turbulent outer scale eddies draw kinetic energy from the mean flow [Roberts and Webster, 2002] which is then transferred to successively smaller scales via complex interactive vortex dynamics [Andreassen et al., 1998, Fritts et al., 1998, 2003]. This downward "energy cascade" continues until the scale of the eddies is small enough that molecular viscosity becomes important, i.e. the inner scale of turbulent motion. At this inner scale the eddy kinetic energy is transferred to heat via the frictional effects of viscous dissipation. This frictional heating therefore represents the transfer of kinetic energy from the mean flow to turbulent scales. The rate at which energy is dissipated to heat is the quantity ϵ introduced above.

6.2.2 Theoretical Approach

A complete and generally accepted theory for the direct heating associated with energy deposition and turbulent diffusion is currently unavailable [Becker, 2004]. Therefore several different approaches to parameterizing direct heating from gravity wave breaking have been applied in previous studies [Becker, 2004]. The approach taken in this study follows the theory detailed in Warner and McIntyre [2001], where as in Fritts and Vanzandt [1993] we assume that for wave breaking all of the energy removed from the gravity wave field is available for dissipation to heat. Thus we assume the potential energy increase due to gravity wave energy deposition can be ignored, based on the results of McIntyre [1989] which states even under the most violent wave breaking at least half the wave energy is available for dissipation to heat, and very likely more than half.

Gravity waves have a total (kinetic plus potential) wave energy per unit mass typically denoted as E. As such waves propagate vertically their energy is transported, where the wave energy flux per unit mass F_E is related to E via [Fritts

and Werne, 2000],

$$F_E = c_{gz} E. \tag{6.35}$$

Here c_{gz} is the group velocity in the vertical direction. Previously Hines and Reddy [1967] had shown that the total gravity wave energy flux, which we shall denote ρF_{ETOTAL}, available for dissipation is given by,

$$\rho(z) F_{ETOTAL} = \rho(z) F_E + \rho(z) \overline{u'w'} \cdot \mathbf{U}. \tag{6.36}$$

Where the second term on the r.h.s of equation (6.36) represents the advection of the mean flow kinetic energy by the gravity wave field. As noted above the spectral description of the wave energy per unit mass $\hat{E}$ adopted by the USSP scheme is identical to the empirical form used by Fritts and Vanzandt [1993], and is given by equation (6.2). Rearranging equation (6.10) to write $\rho\hat{E}$ in terms of $\rho\hat{\mathbf{F}}_\mathbf{p}$, with the $\hat{\omega}$ dependence integrated out and using equations (6.7) and (6.35) we obtain the following expression, equation (12) of Warner and McIntyre [2001], for the 2D wave energy flux at an altitude z,

$$\rho(z) F_E(z, m, \phi_j) = \rho(z) \left| \mathbf{F}_\mathbf{p}(z, m, \phi_j) \right| \frac{N}{m} \frac{\Delta\phi}{2\sin(\Delta\phi/2)}. \tag{6.37}$$

Since energy flux is a scalar quantity and momentum flux is a vector, the unit vector $\hat{\mathbf{k}}_\mathbf{o}$ in equation (6.10) is replaced by the $\Delta\phi/[2\sin(\Delta\phi/2)]$ term in equation (6.37), which arises from integrating over each azimuthal sector. The error associated with this discretization is insignificant compared to the empirical uncertainty in the USSP parameters used to define the initial wave energy spectrum [Warner and McIntyre, 1996]. Integration of equation (6.37) with respect to m allows the calculation of the one dimensional energy flux $\rho F_E(z, \phi)$ in each azimuthal direction, full details are given below. Following Warner and McIntyre [2001] we can then calculate the one dimensional net wave energy flux, summed over all azimuthal directions, at an altitude z as,

$$\rho F_E(z) = \sum_{j=0}^{n-1} \rho F_E(z, \phi_j). \tag{6.38}$$

Where $n = 4$ is the number of azimuthal directions. Under the mid-frequency approximation the N/m term in equation (6.37) is equal to the intrinsic horizontal phase speed of the vertically propagating gravity wave. Hence, ignoring the discretization over each azimuthal sector, equation (6.37) is equivalent to the first theorem of Eliassen and Palm [Eliassen and Palm, 1961] which associates the vertical flux of gravity wave energy to the momentum flux via,

$$\rho(z)F_E = -\rho(z)(\mathbf{U} - c)\overline{u'w'} = \overline{p'w'}. \tag{6.39}$$

Here $\mathbf{U}$ is the mean flow, c is the gravity wave horizontal phase speed and $\overline{p'w'}$ is a quantity known as the gravity wave pressure flux, where over bars refer to the average over one wave cycle. In the USSP formulation the gravity wave energy flux per unit mass is therefore equal to the wave pressure flux. For conservative wave propagation there is no deposition of gravity wave energy into the mean flow, and we therefore require that the vertical divergence of equation (6.39) be equal to zero [Fritts and Vanzandt, 1993]. Since under conservative propagation the vertical divergence of $\overline{u'w'}$ is zero [Eliassen and Palm, 1961], it is clear that for such conditions the vertical divergence of equation (6.39) is non zero in the presence of mean shear. This paradox occurs as the wave energy flux, defined in terms of the wave pressure flux, is not conserved in the presence of mean shear. Here changes in energy flux occur because when momentum flux propagates through vertical mean shear there is a continuous transfer of energy between the mean flow and the wave [Bretherton, 1966]. This wave-mean flow interaction therefore alters the amount of gravity wave energy available for dissipation and must be accounted for when computing ϵ. This paradox was solved by Hines and Reddy [1967] with the addition of the term describing advection of the mean flow kinetic energy in equation (6.36). As noted by Lindzen [1973] it is the vertical divergence of equation (6.36) that equals zero under conservative propagation

in the presence of mean shear. Following equation (13) of Lindzen [1973] we can equate the vertical divergence of equation (6.36) to the increase of kinetic energy and heat of the mean flow via,

$$\rho(z)\frac{\partial}{\partial t}(\frac{1}{2}\mathbf{U}^2) + \rho(z)\epsilon = -\frac{\partial}{\partial z}(\rho(z)F_E + \rho(z)\overline{u'w'} \cdot \mathbf{U}). \tag{6.40}$$

Here the first term on the l.h.s of equation (6.40) is the change of mean flow kinetic energy, while the second term on the l.h.s is the increase of heat denoted by ϵ times the density at altitude z. The mean flow kinetic energy term in equation (6.40) can be written as $\rho\mathbf{U}(\partial\mathbf{U}/\partial t)$. Expanding the r.h.s of equation (6.40) and expressing $(\partial\mathbf{U}/\partial t)$ in terms of momentum flux via equation (6.12), we have the following expression for the turbulent energy dissipation rate in terms of the momentum flux and the USSP formulation (wave pressure flux) of the net wave energy flux at an altitude z [Warner and McIntyre, 2001].

$$\epsilon(z) = -\frac{1}{\rho(z)}\frac{\partial}{\partial z}\rho F_E(z) - \frac{\partial\mathbf{U}}{\partial z} \cdot \mathbf{F_p}(z) \tag{6.41}$$

The second term on the r.h.s of equation (6.41) has been referred to as the "residual work" [Becker and Schmitz, 2002]. The turbulent energy dissipation rate can then easily be converted to a local temperature increase of the mean flow via [Lübken, 1997],

$$\frac{\partial T}{\partial t} = \frac{\epsilon(z)}{c_p}. \tag{6.42}$$

Where c_p is the heat capacity of dry air, which has a value of ~1005 $JK^{-1}kg^{-1}$. Using equation (6.41) it is therefore possible to determine the local heating of the mean flow due to gravity wave breaking.

6.2.3 One Dimensional Wave Energy Spectrum

As noted above, to calculate the local heating of the mean flow due to gravity wave breaking it is necessary to determine vertical profiles of the one dimensional wave energy flux at each altitude z_k. Applying the previous definitions

of the momentum flux spectrum at z_k to (6.37) and integrating with respect to m_k, we obtain the following expressions for the one dimensional wave energy flux at z_k for each of the chop types described in Section 6.1.6. While the following calculations are an extension of the momentum flux profiles determined in Bushell et al. [2007] using the theory of Warner and McIntyre [2001], their formal definitions are original to this work.

Chop Type I

$$\rho_k F_{Ek}(z_k, \phi_j) = \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} \left[1 - A_0 \left(\int_{m_{lY}}^{\infty} \frac{m_*^{t-1}}{m^{t+1}} dm - \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \int_{m_{kY}}^{\infty} \frac{m_*^{t-1}}{m^{t+1}} dm \right) \right]$$

$$= \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l F_{pltot} \left[1 - \frac{A_0}{t m_*} m_{NlY}^{-t} \left(1 - \frac{N_l}{N_k} B \left(1 + \frac{A}{B} m_{NlY}\right)^t\right)\right] \tag{6.43}$$

Chop Type IIa

$$\rho_k F_{Ek}(z_k, \phi_j) = \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} A_0 \left[\int_{m_{min}}^{m_{lX}} \frac{m^{s-1}}{m_*^{s+1}} dm + \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \int_{m_{kX}}^{\infty} \frac{m_*^{t-1}}{m^{t+1}} dm \right]$$

$$= \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} \frac{A_0}{m_*} \left[\frac{1}{s} \left(m_{NlX}^s - m_{Nlmin}^s \right) + \frac{1}{t} \frac{N_l}{N_k} B m_{NlX}^{-t} \left(1 + \frac{A}{B} m_{NlX}\right)^t \right] \tag{6.44}$$

Chop Type IIb

$$\rho_k F_{Ek}(z_k, \phi_j) = \frac{\Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} \frac{A_0 N_l}{t m_*} B m_{Nlmin}^{-t} \left(1 + \frac{A}{B} m_{Nlmin}\right)^t \tag{6.45}$$

Chop Type IIc

$$\rho_k F_{Ek}(z_k, \phi_j) = \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} \frac{A_0}{m_*} \left[\frac{1}{s} \left(m_{NlX}^s - m_{Nlmin}^s \right) + \frac{1}{t} \frac{N_l}{N_k} B m_{NlX}^{-t} \right] \tag{6.46}$$

Chop Type IIIa

$$\rho_k F_{Ek}(z_k, \phi_j) = \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} A_0 \ *$$

$$\left[\int_{m_{min}}^{m_{lX}} \frac{m^{s-1}}{m_*^{s+1}} dm + \frac{\rho_k}{\rho_l}\frac{C_{Sk}}{C_{l0}}\left(\int_{m_{kX}}^{\infty} \frac{m_*^{t-1}}{m^{t+1}} dm - \int_{m_{kY}}^{\infty} \frac{m_*^{t-1}}{m^{t+1}} dm\right) + \int_{m_{lY}}^{\infty} \frac{m_*^{t-1}}{m^{t+1}} dm\right]$$

$$= \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} \frac{A_0}{m_*}\ * \tag{6.47}$$

$$\left[\frac{1}{t} m_{NlY}^{-t}\left(1 - \frac{N_l}{N_k} B\left(1 + \frac{A}{B} m_{NlY}\right)^t\right) + \frac{1}{s}\left(m_{NlX}^s - m_{Nlmin}^s\right) + \frac{1}{t}\frac{N_l}{N_k} B m_{NlX}^{-t}\left(1 + \frac{A}{B} m_{NlX}\right)^t\right]$$

Chop Type IIIb

$$\rho_k F_{Ek}(z_k, \phi_j) = \frac{N_k \Delta\phi}{2\sin(\Delta\phi/2)}\ \rho_l \mathbf{F_{pltot}} \frac{A_0}{m_*}\ * \tag{6.48}$$

$$\left[\frac{N_l}{N_k}\frac{B}{t} m_{Nlmin}^{-t}\left(1 + \frac{A}{B} m_{Nlmin}\right)^t \frac{1}{t} m_{NlY}^{-t}\left(1 - \frac{N_l}{N_k} B\left(1 + \frac{A}{B} m_{NlY}\right)^t\right)\right]$$

Chop Type IV

$$\rho_k F_{Ek}(z_k, \phi_j) = 0 \tag{6.49}$$

Applying the above expressions to the USSP code to calculate the one dimensional wave energy flux in each azimuthal direction allows the calculation of the local heating rates at each altitude increment z via equation (6.41).

6.2.4 Implementation of the Heating Rate Code

As stated above the currently operational USSP only employes momentum forcing from gravity wave breaking, and not direct heating rates. Therefore to investigate the impact of such direct heating in Section 8 the USSP code was modified and compiled into an experimental build of the UM. Here the above expressions for each chop type describing the one dimensional energy flux profile were added to the USSP code. For each time step of the model integration equation (6.41) was used to determine the direct heating rates, which were then transformed to potential temperature increments and passed to the appropriate UM routine. New diagnostic variables were created within the UM system to

archive the direct heating rate and potential temperature increments averaged over one day.

6.3 Summary

This chapter has fully detailed the currently operational USSP gravity wave scheme used in the L50 and L70 analyses. Details of the free parameters and initial launch level of momentum flux in the operational scheme were given, along with a description of the transparent and opaque lid conditions. Information on the theoretical and practical approach of extending the operational scheme to include direct heating rates from gravity wave breaking was given, with associated expressions for the one dimensional wave energy flux profile for each of the momentum flux profile chop types fully detailed.

Chapter 7

Off-line Estimates of Gravity Wave Drag and Direct Heating Rates

The evidence presented in Sections 4 and 5 suggests that throughout the mesosphere the gravity wave forcing supplied by the USSP scheme for both the L50 and L70 models is too weak. Unfortunately model fields of gravity wave drag are not archived for the stratospheric assimilated data set. However it is possible to obtain estimates of gravity wave drag in the model by applying the USSP scheme "off-line" to the temperature and wind fields of both the L50 and L70 analyses. Using this method it is possible to conduct off-line (computationally inexpensive) sensitivity studies with respect to the free parameters of the USSP, thus determining the logical parameters, those that produced additional gravity wave forcing, to perturb in the UM (computationally expensive) sensitivity experiments in Section 8. Before these off-line sensitivity experiments are detailed, the uncertainty in the off-line calculations, primarily due to the launch altitude of the initial momentum flux spectrum, and comparison of off-line forcing values with previous observational and modelling studies are both discussed. In addition off-line estimates of the momentum flux and the direct heating rates due to gravity wave breaking for the L70 model are also calculated, and compared to previous observational and experimental values.

7.1 Off-line Drag Uncertainty

While calculated off-line values are expected to present a realistic representation of the actual model fields, they are subject to a certain degree of error, as discussed below. Since the temperature and winds fields used in the off-line calculations are from the daily analyses at 1200 UTC, the results are essentially for single time-steps of the forecast model integration at this particular time. For model forecasts, the induced gravity wave forcing is calculated and applied to model fields after each time step throughout each day. Hence model fields represent the impact that applying forcing at each time step has on the temperature and wind field, thus representing the feedback which occurs due to differences in upward propagating wave filtering arising from changes in the temperature and wind fields produced from gravity wave forcing. This feedback is obviously not accounted for in the single time step off-line calculations. However since the middle atmospheric circulation and temperature are known to vary slowly in time below the lower mesosphere, where the majority of wave filtering occurs, it is expected that the error introduced from using a single timestep for each day will be negligible, particularly when averaging over monthly timescales.

The largest source of discrepancy between model and off-line fields is expected to arise from the difference in altitudes of the initial launch spectrum. Previous studies of Ern et al. [2004, 2006] have shown using that the momentum flux and gravity wave distributions of the stratosphere, obtained using the USSP scheme, are sensitive to the launch altitude of the initial momentum flux spectrum, and in fact the appropriate choice of launch level is open for debate [Fritts and Alexander, 2003]. Here differences in the above fields could occur due to the differences in wave filtering by the tropospheric jets. The results of Ern et al. [2004] show that when comparing absolute values of momentum flux produced by applying the USSP to Met Office stratospheric analyses with those from the CRISTA missions, horizontal correlation between the two distributions in the stratosphere are more realistic when launching the initial spectrum from the 681 hPa (~2.7 km) and 464

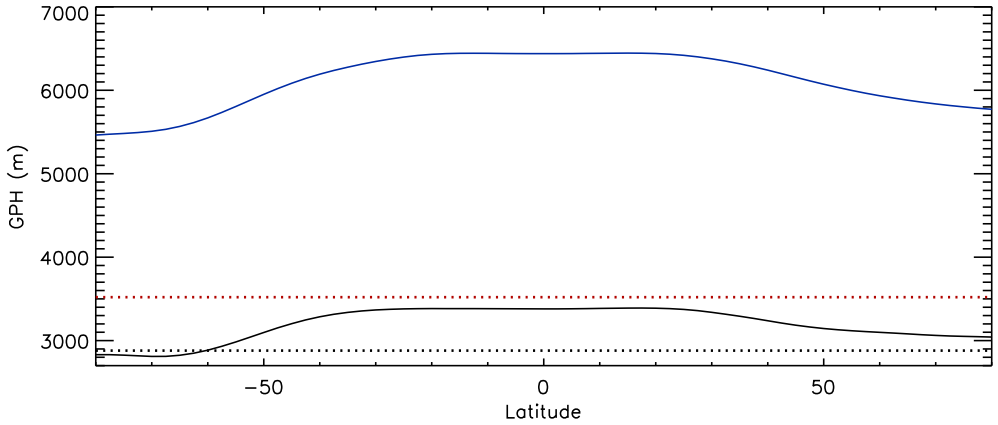

Figure 7.1: Geopotential heights for model and off-line launch altitudes of USSP initial momentum flux. Solid black and blue lines denote 2008 yearly zonal average for off-line L50 calculations at 681 hPa and 464 hPa respectively. Dashed black and red lines denote USSP launch level for L50 and L70 analyses respectively.

hPa (~5.4 km) pressure levels. Furthermore Ern et al. [2006] noted that this horizontal correlation was optimised when launching from the 464 hPa level, and holds true for varying values of the characteristic vertical wavenumber m^* and the spectral power index s. However as the hard-wired launch levels for the L50 and L70 analyses are ~2880 m and ~3520 m respectively, the launch level for both the L50 and L70 off-line calculations was set to 681 hPa. The yearly zonal averages of geopotential height (GPH) at the 681 hPa and 464 hPa pressure levels of the L50 analyses for 2008 are shown in Figure 7.1, also shown for comparison are the hard-wired launch levels for both the L50 and L70 models. Note the zonal averages of GPH at the above pressure levels for the L70 analyses were nearly identical to those of the L50, and hence are not shown. The above values of GPH for the L50 2008 analyses are typical of all other years.

From Figure 7.1 we find that at high southern latitudes the average GPH of the L50 analyses at 681 hPa is nearly identical to the hard-wire L50 model value. While for high northern latitudes the analyses values at 681 hPa are ~200 m

higher than the hard-wired value, with this difference increasing to ~400→600 m for middle and tropical latitudes of both hemispheres. To estimate the impact this difference has on the accuracy of the off-line calculations we shall first determine the impact of launching the initial spectrum in the off-line calculations from the 464 hPa level, which is ~2600 m higher than the 681 hPa case. Figure 7.3 shows the monthly zonal mean off-line gravity wave values for March, June, September and December 2008 when launching the initial spectrum from 681 hPa and 464 hPa, also shown is the residual between the two distributions.

For an appreciation of the differences in wave filtering which occur when launching from the two pressure levels, the L50 monthly zonal mean winds in the troposphere for associated months are shown in Figure 7.2. Here the vertical range not propagated through when using the 464 hPa launch level is denote by the area between the two dashed black lines. For extra-tropical latitudes, where the zonal winds of all seasons are eastward, there is significant positive eastward vertical wind shear (see equation (6.31)). From wave filtering we would therefore expect a larger erosion of the eastward spectrum when launching at 681 hPa compared to 464 hPa. Therefore launching from 464 hPa under solstice conditions, one would expect in the winter (summer) hemisphere a decrease (increase) of asymmetry in the zonal forcing resulting in a decrease (increase) of westward (eastward) drag values in the lower mesosphere compared to the 681 hPa case. This is indeed what is observed in Figure 7.3. Similar arguments also apply to solstice tropical latitudes and all other seasons.

From Figure 7.3 we find that the maximum differences in drag values occur in the lower mesosphere at latitudes which approximately correspond to the location of the tropospheric jet maximums for that season. This correlation is expected as we would anticipate any differences in the asymmetry of zonal forcing due to the launch altitude height to have a larger impact for regions which have a considerable amount of zonal asymmetry above, i.e near the tropospheric jet cen-

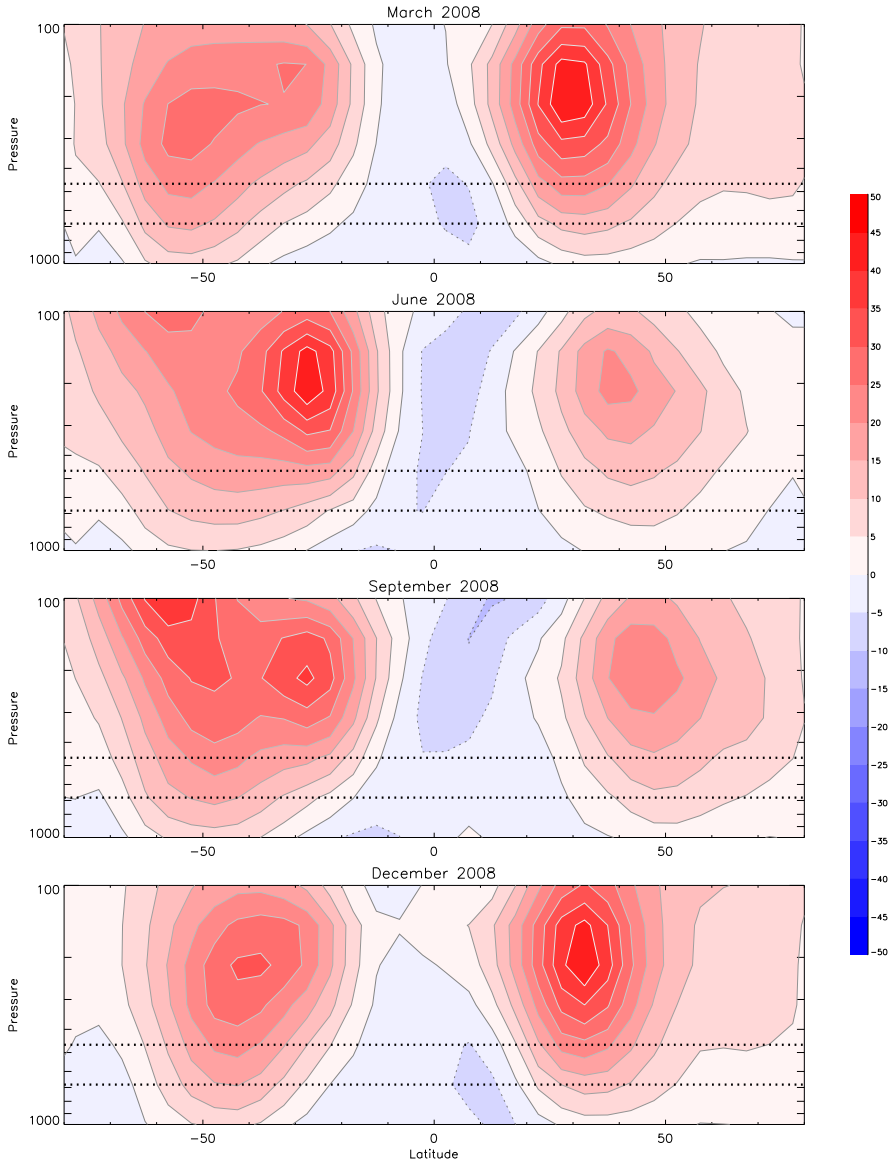

Figure 7.2: L50 analyses monthly zonal mean winds in the troposphere for March, June, September and December 2008. Contour interval is 5 ms^{-1}, negative values are dashed. Dashed black lines denoted the 681 hPa and 464 hPa launch levels.

tres where the vertical wind shear is greatest. Largest differences in forcing are seen in the SH winter and summer hemispheres with magnitudes of ~7.5→12.5 $ms^{-1}day^{-1}$ and ~5.0→7.5 $ms^{-1}day^{-1}$ respectively. Difference magnitudes of all other seasons and latitudes are generally below ~5.0 $ms^{-1}day^{-1}$. Since the above

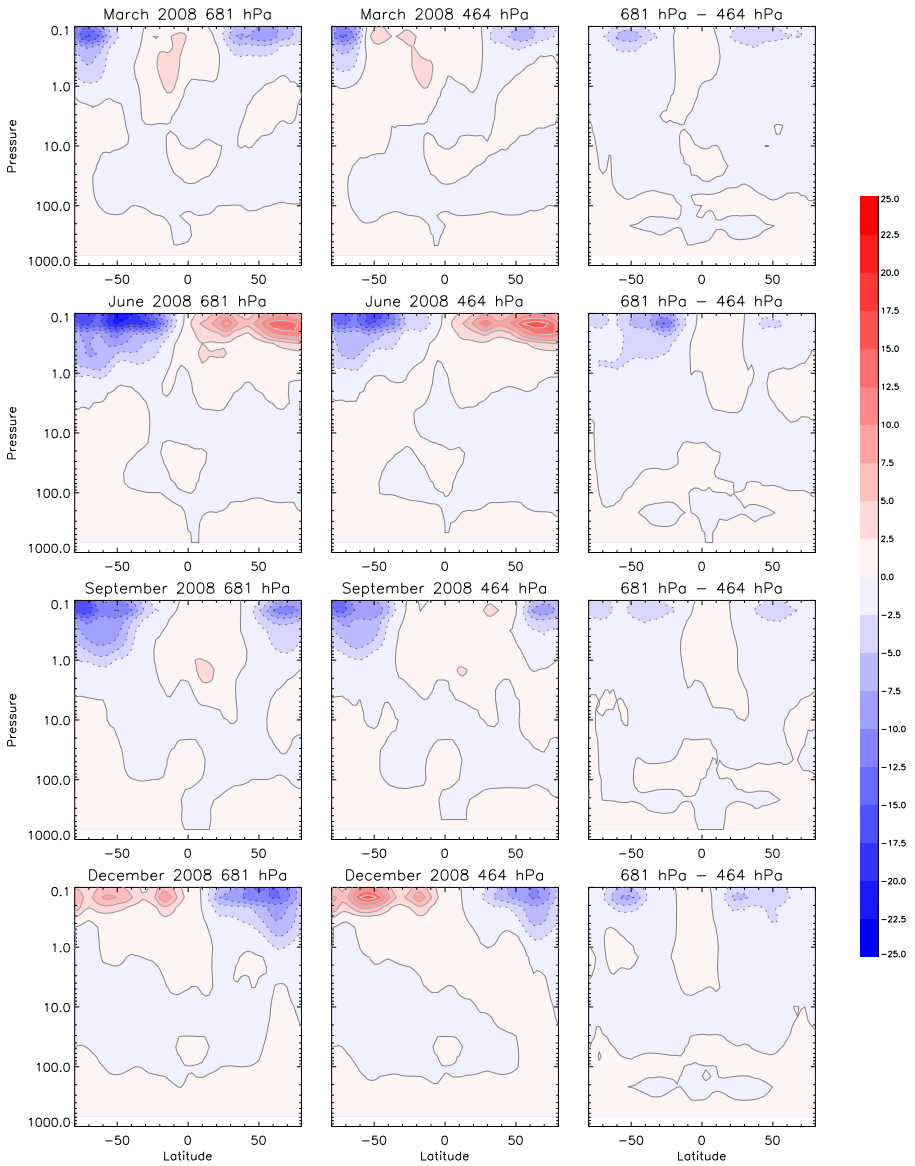

Figure 7.3: Monthly zonal averages of off-line gravity wave drag using L50 analyses for March, June, September and December 2008 launching initial momentum flux from 681 hPa (left column) and 464 hPa (middle column), positive (negative) values correspond to eastward (westward) forcing. The residual (681 hPa minus 464 hPa) is shown in the right column. Units ($ms^{-1}day^{-1}$).

differences in forcing are found when the launch altitudes are ~2600 m apart, it is reasonable to expect that the smaller differences (maximum ~600 m for the L50

analyses) between the hard-wired model and pressure GPH values will have an even smaller impact. Based on simple scaling arguments, since the vertical wind shear is approximately constant with height for the altitude range in question, it is reasonable to expect that a ~600 m difference in GPH would impact the off-line calculations by ~25% of the above values for ~2600 m, i.e below ~3 $ms^{-1}day^{-1}$.

Therefore we can state that the off line calculations of gravity wave drag for the L50 analyses, discussed below, most likely underestimate (overestimate) westward (eastward) forcing. However the magnitude of this inaccuracy is generally less than ~10% of the presented values. Differences of monthly zonal mean drag when launching from 681 hPa and 464 hPa for the L70 off-line calculations (not shown) were found to be of the same order in the upper mesosphere as those described for the L50 case. Since the launch altitudes in the L70 off-line calculations are below the hard-wired model value, we therefore expect that the results presented below most likely overestimate (underestimate) eastward (westward) forcing in the lower and upper mesosphere. Again this inaccuracy is generally less than ~10%, but could possibly be slightly larger for the high latitude SH where GPH differences between model and pressure levels are ~700 m.

7.2 Off-line Momentum Flux Values

As previously noted in Section 4 it is possible that the application of the transparent lid condition may contribute to temperature errors of the analyses fields. Before investigating the impact of applying the opaque lid condition in the UM experiments of Section 8 off-line studies were performed to asses the magnitude of remaining momentum flux approaching the L70 model lid. Figure 7.4 details the latitude time-series of the off-line zonal average zonal momentum flux at 0.01 hPa, using the L70 temperature and wind analyses from October 2009 to June 2010. Here the standard parameters of Table 6.1.7 are used with the transparent lid condition set and the initial momentum flux spectrum launched from 681 hPa. From Figure 7.4 we find significant portions of the initial launch spectrum

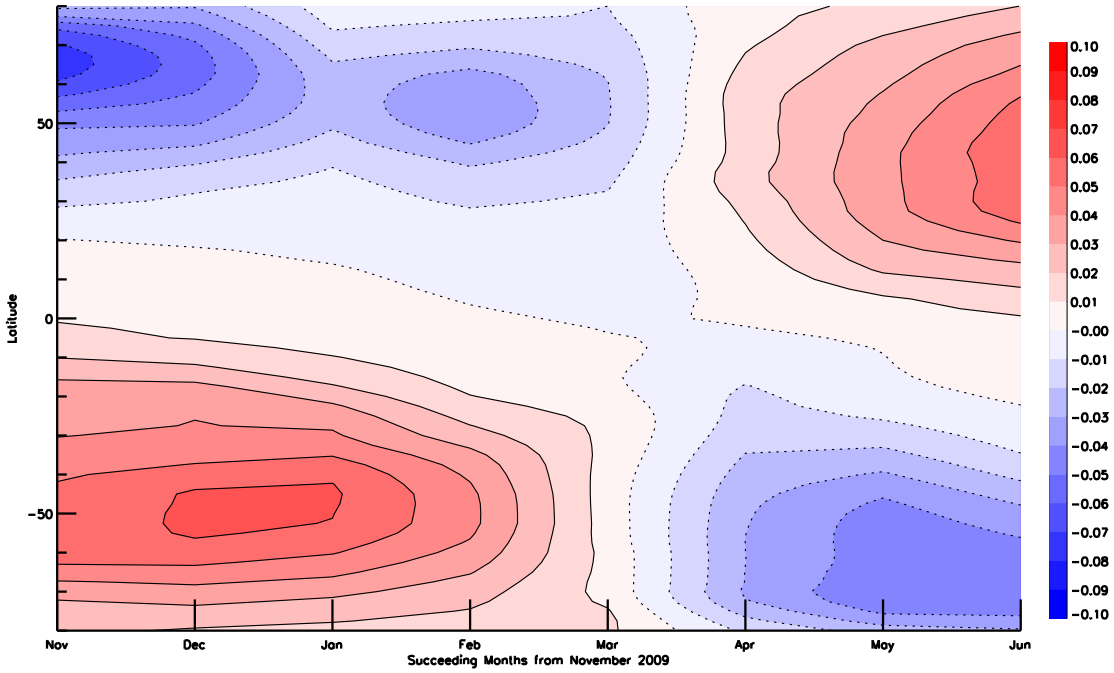

Figure 7.4: Latitude timeseries of off-line zonal average zonal momentum flux at 0.01 hPa for the L70 analyses from November 2009 to June 2009. Contour interval is 0.01 mPa with negative values dashed.

are still unsaturated at the model lid, where largest values are seen at middle and high latitudes of each hemisphere having magnitudes of ~0.05→0.06 mPa. This is consistent with the previous studies of Fritts and Vincent [1987], Fritts and Yuan [1989], Vincent and Fritts [1987] who showed that momentum flux carried by gravity waves can propagate through the mesopause and into the lower thermosphere.

7.3 Off-line Drag Values

7.3.1 L50 Model

Calculations of the off-line gravity wave drag values were performed using daily analyses of temperature, wind and geopotential height. The USSP scheme was

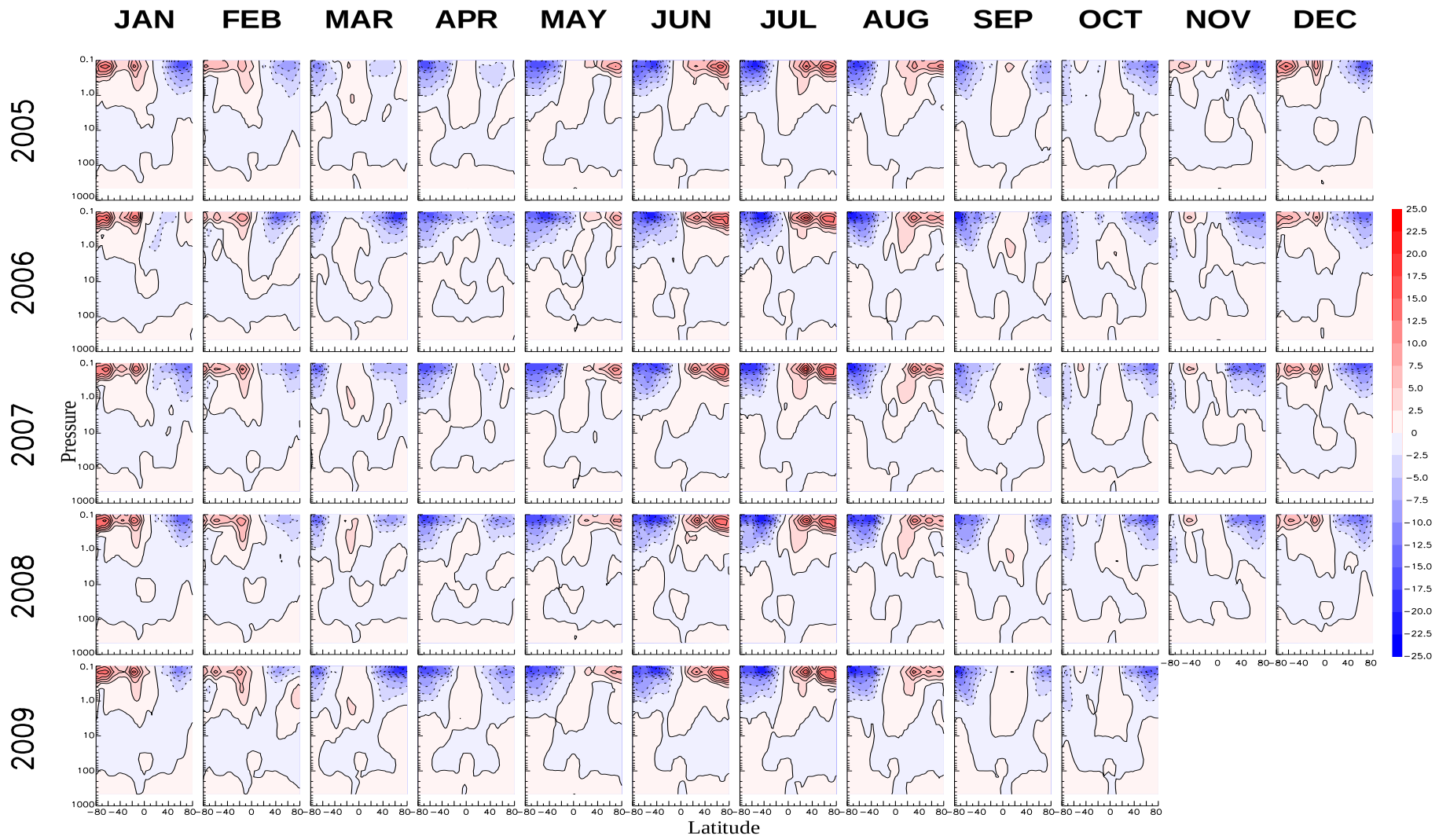

Figure 7.5: Off-line monthly zonal mean zonal gravity wave drag values from January 2005 to October 2009. Drag values have contour interval of 5 $ms^{-1}day^{-1}$ with negative dashed.

applied to these daily analyses fields using the standard parameter values detailed in Table 6.1.7; as for the operational USSP scheme the transparent lid condition is used. These single time-step daily values were then averaged over the month and in the zonal direction. The zonal component of these monthly zonal mean off-line calculations from January 2005 to October 2009 are shown in Figure 7.5, discussion follows.

The majority of forcing occurs at the upper levels of the model domain, with magnitudes greater than 10 $ms^{-1}day^{-1}$ generally only seen in the lower mesosphere above ~0.4 hPa. Under solstice conditions there is eastward forcing in both summer hemispheres. There are two distinct regions of forcing, one occurring in the sub-tropics at ~30-40^o latitude and a stronger second region at polar latitudes. Maximum values of forcing are identical for each hemisphere and typically occur in January and July with magnitudes of ~12.5→17.5 $ms^{-1}day^{-1}$ between ~0.2→0.1 hPa. For winter seasons there is a westward forcing in each hemisphere, with maximum values seen at middle to high latitudes of ~50→70^o between ~0.2→0.1 hPa. For the SH winter months the magnitude of this westward forcing is generally greater than the eastward forcing seen in the opposing summer hemisphere, with typical values of ~17.5→22.5 $ms^{-1}day^{-1}$. For the NH winter months the magnitude of the westward forcing is comparably to that seen in the opposing summer hemisphere, with magnitudes of ~12.5→17.5 $ms^{-1}day^{-1}$. For equinox months of both hemispheres the maximum forcing is westward and occurs at polar latitudes of both hemispheres with typical values of ~10→15 $ms^{-1}day^{-1}$. The above distributions and magnitudes of off-line forcing by the USSP in the lower mesosphere are qualitatively similar to those obtained by Scaife et al. [2002], with the only exception being that the westward off-line drag values of the winter SH are ~5 $ms^{-1}day^{-1}$ larger. This consistency places confidence in the results of the off-line values. In should be noted that there are very small amounts of inter-annual variability in the above values of off-line forcing, the exception is in January 2006 where in the NH winter the

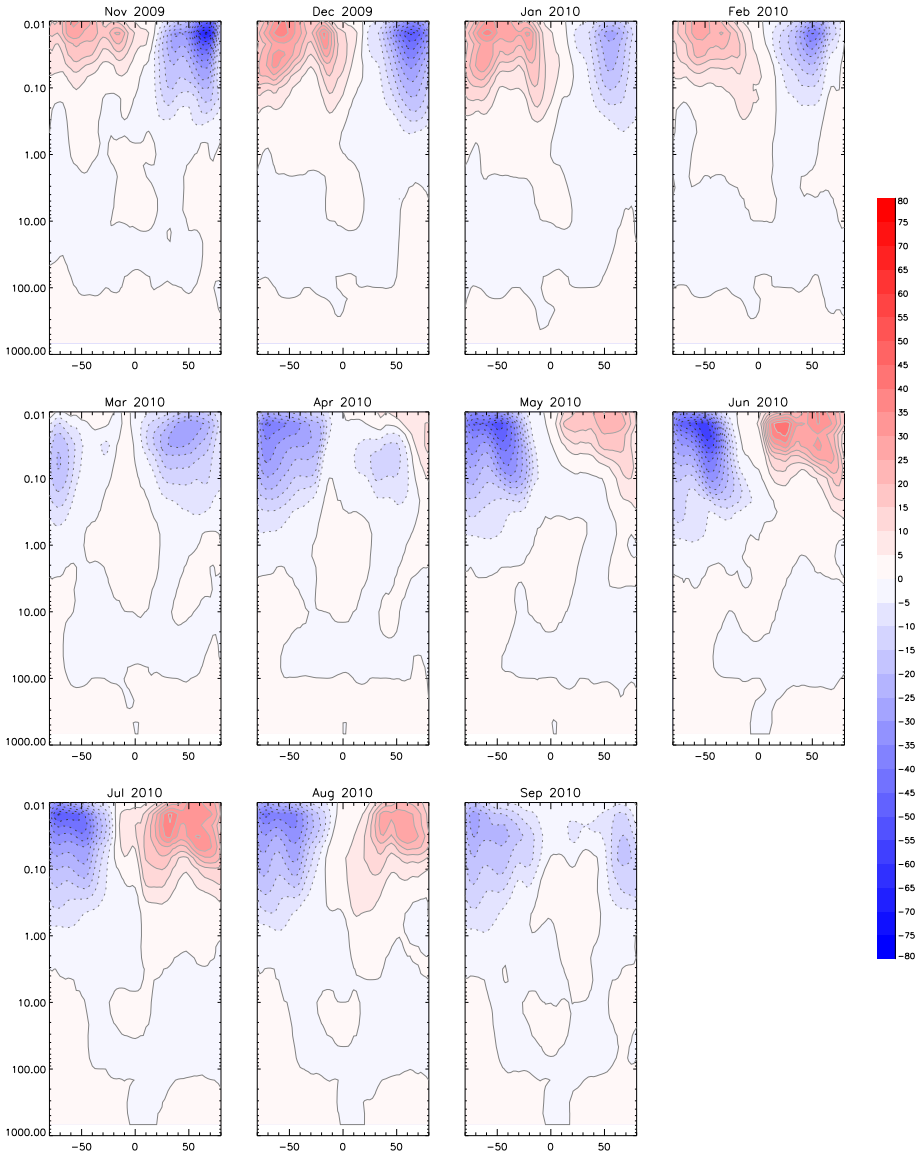

Figure 7.6: Monthly zonal mean zonal gravity wave drag values for the L70 analyses from November 2009 to June 2010. Drag values have contour interval of 5 $ms^{-1}day^{-1}$ with negative values indicted by dashed contours.

westward forcing is noticeably smaller than other years. Such a difference is most likely a consequence of the weaker eastward winds resulting from the major sudden stratospheric warming which occurred in this month, as detailed in Section 5.1.

7.3.2 L70 Model

Figure 7.6 details the monthly zonal mean zonal drag values for the off-line calculations using the L70 analyses from November 2009 to June 2010; discussion follows.

As for the lower mesophere in the L50 analyses discussed above, the gravity wave forcing supplied by the USSP in the L70 analyses tends to oppose the seasonal zonal mean winds of each hemisphere in the upper mesosphere. Here the magnitudes of forcing supplied at ~0.1 hPa are qualitatively similar to those seen in the L50 analyses for each season. Maximum forcing occurs near the model lid for each season between ~0.02→0.01 hPa.

For the summer months of both hemisphere the upper mesosphere has two distinct regions of large forcing, occurring at sub-tropical latitudes of ~15→30^o and extra-tropical latitudes of ~50→60^o. For the SH summer months maximum forcing occurs at extra-tropical latitudes in the solstice month of December with magnitudes of ~35→40 ms^{-1}day^{-1}. Similar magnitudes are also seen in the NH summer month of June, however here maximum values occur at tropical latitudes.

For the winter season of both hemispheres the magnitudes of forcing seen at tropical latitudes are smaller than those seen in the opposing tropical summer hemisphere. Winter season maximum values generally occur at extra-tropical latitudes of ~50→70^o. For the NH winter season maximum forcing of ~60→65 ms^{-1}day^{-1} and ~45→50 ms^{-1}day^{-1} are seen for the months of November and December respectively. For the SH winter season maximum forcing of ~ ~55→60 ms^{-1}day^{-1} is seen for June and ~45→50 ms^{-1}day^{-1} for May and July.

As noted in Sections 4.6.1 and 5.2 the NH winter months of November and December 2009 have an uncharacteristic temperature structure, with a lower

mesospheric polar vortex that is too cold, resulting in zonal means winds which are too strong and a jet maximum that occurs too high. Based on wave filtering arguments, these excessive winds would result in westward propagating gravity being able to propagate to higher altitudes and thus through the density decrease have larger amplitudes when wave breaking occurs. Therefore the off-line calculations in November and December 2009 are not characteristic of the USSP forcing of this season, with larger values than would be obtained if the lower mesospheric temperature and winds were more realistic. Hence it is impossible to comment on the degree of asymmetry between NH and SH winter forcing produced by the USSP scheme in the L70 analyses.

7.3.3 Comparison with Previous Studies

Estimation of gravity wave drag values in the middle atmosphere are generally performed using the following techniques: (1) Momentum balance studies where calculations of atmospheric heating rates can be used with the knowledge of zonal temperature and wind fields to estimate the resolved EP-flux values and all terms on the l.h.s of equation (2.6), from which the residual of the above terms is assumed to be small scale forcing from gravity waves [Hamilton, 1983, Marks, 1989, Shine, 1989, Smith and Lyjak, 1985]. (2) Observational studies of momentum flux, from which gravity wave drag values can be inferred [Jiang et al., 2006, Nakamura et al., 1996, Tsuda et al., 2000]. (3) Model studies using gravity wave schemes which are constrained by observations [Holton, 1983]. (4) Data assimilation where inverse techniques are used to estimate gravity wave drag from knowledge of the atmospheric flow and differences in the observed and predicted forcing [Pulido and Thuburn, 2005, 2006]. (5) High resolution modelling where it is assumed that the majority of small scale forcing is resolved in the mean fields and hence can be estimated by decomposing values of EP-flux divergence into appropriate length scales [Hamilton et al., 1999, Watanabe et al., 2008].

The momentum balance study of Marks [1989] calculated that small scale forcing peaks at ~55 $ms^{-1}day^{-1}$ in the SH winter at ~0.1 hPa and between ~50→60^o. The NH winter has smaller peak values of ~40 $ms^{-1}day^{-1}$ centred at similar pressure and latitude values. For both summer hemispheres the peak drag values occur over larger latitude ranges than winter values, from ~20→70^o, with magnitudes of ~15→20 $ms^{-1}day^{-1}$ at ~0.1 hPa. Similar magnitudes and distribution of forcing were also found by [Shine, 1989]. Qualitatively consistent with the above estimates were the results of Smith and Lyjak [1985] and Hamilton [1983], which separately calculated ~25 $ms^{-1}day^{-1}$ between ~60→70^o at ~0.4 hPa in the NH winter season.

Observational studies of momentum flux and inferred gravity wave drag provide qualitatively consistent results to those obtained by momentum balance studies. In the extra-tropics of both hemispheres mesospheric drag has been shown to range between ~10→70 $ms^{-1}day^{-1}$ [Fritts and Yuan, 1989, Reid and Vincent, 1987, Tsuda et al., 1990]. Also consistent with momentum balance studies the inferred drag values of Nakamura et al. [1996] showed larger values for the SH winter season compared to the NH winter [Fritts and Alexander, 2003]. This asymmetry in mesospheric forcing was also noted in Jiang et al. [2006], where winter drag values peak at ~40 $ms^{-1}day^{-1}$ at ~60 km in the NH and ~100 $ms^{-1}day^{-1}$ at 55 km for the SH.

Modelling studies generally produce larger mesospheric drag values than the above two methods [Fritts and Alexander, 2003]. Here the studies of Holton [1983], Garcia and Solomon [1985] both estimate that peak forcing at mid-latitudes under solstice conditions has magnitudes of ~30→120 $ms^{-1}day^{-1}$. Uncertainty in these results is relatively large due to limited accuracy of observational data used to tune certain free parameters in the gravity wave schemes used. Recently operational middle atmospheric analyses (up to 0.01 hPa) created by the ECMWF Integrated Forecast System (IFS) have adopted the spectral

gravity wave scheme by Scinocca [2003]. The study of Orr et al. [2010] details the monthly zonal mean temperature and wind fields averaged over a 6-year climate run of the IFS model when the Scinocca gravity wave scheme in employed. They found that realistic fields, compared to the SPARC temperature climatology and ECMWF ERA-Interim wind reanalysis [Simmons et al., 2007], were produced when forcing from the gravity wave scheme in the summer months of both hemispheres was ~40 $ms^{-1}day^{-1}$ from ~50→60^o latitude approaching ~0.01 hPa. However it was noted that the magnitude of this upper mesospheric forcing in the SH summer is too weak. In the SH winter the forcing supplied by the Scinocca scheme has maximum values of ~60 $ms^{-1}day^{-1}$ from ~50→60^o latitude at ~0.04 hPa. Also as noted in Orr et al. [2010] the forcing supplied by the Scinocca scheme is placed too high in each hemisphere for both solstice seasons when compared with previous studies.

A recent and novel technique for estimating drag is the so called assimilation system for drag estimation (ASDE) detailed in [Pulido and Thuburn, 2005, 2006]. The study of Pulido and Thuburn [2008] used ASDE to evaluated the seasonal cycle of gravity wave drag using Met Office stratospheric analyses for 2002. They found peak values of gravity wave drag in the SH winter of ~50 $ms^{-1}day^{-1}$ between ~60→70^o latitude at ~0.4 hPa. Peak values of NH winter drag were found at similar locations with magnitudes of ~35 $ms^{-1}day^{-1}$, again noticeably smaller than SH winter values. In both summer hemispheres the drag values peaked at ~30 $ms^{-1}day^{-1}$ between ~20→40^o latitude and at a lower altitude of ~0.6 hPa. However as noted in Pulido and Thuburn [2008] it is not possible to obtain a definitive conclusion about the height of peak drag values as they occur at the top of the analyses used in the study.

The high resolution modelling study of Watanabe et al. [2008] reproduced realistic values of temperature and circulation in the mesosphere. They found that gravity wave forcing in both the NH and SH summer months had maximum

values of ~50→60 $ms^{-1}day^{-1}$ approaching 0.01 hPa located between ~50→60^o latitude. For these latitudes the forcing decreases with decreasing height, where values have fallen off to ~10 $ms^{-1}day^{-1}$ at ~0.04 hPa. Furthermore they found that for sub-tropical summer latitudes the gravity wave forcing extends further downwards, where for both hemispheres forcing of ~20 $ms^{-1}day^{-1}$ can be seen at ~0.1 hPa. The sub-tropical forcing occurs above the summer season secondary jet maximum discussed in Section 2.2.2. In the winter months they found that the gravity wave forcing has maximum values of ~70→80 $ms^{-1}day^{-1}$ in the NH, and ~80→100 $ms^{-1}day^{-1}$ in the SH. The SH winter maximum forcing occurs over a large latitude range from ~30→60^o and extends from ~0.07→0.01 hPa. In contrast the NH winter forcing maximum values are contained over latitudes of ~50→60^o from ~0.1→0.04 hPa. For both hemispheres the maximum forcing in the extra-tropical summer hemisphere occurs at higher altitudes than maximum forcing in the extra-tropical winter hemisphere, consistent with the previous results of Marks [1989].

The upper mesospheric forcing values of Watanabe et al. [2008] are also qualitatively consistent with the previous estimates of Holton [1982], who found forcing of ~70→100 $ms^{-1}day^{-1}$ were needed in each hemisphere under solstice conditions to reproduce realistic wind profiles in their model which used the Lindzen [1981] parameterization. The study of Fomichev et al. [2002] using the extended Canadian Middle Atmosphere Model (CMAM) model, employing the Hines doppler spread parameterization (DSP) of gravity waves, also estimated gravity wave drag in the upper mesosphere to have values approaching ~100 $ms^{-1}day^{-1}$ above ~75 km.

In general the values of the off-line calculations in Figure 7.5 show lower mesospheric drag values which are considerably smaller, or at the lower limit, of those diagnosed by the previous studies discussed above. This disagreement is less for the summer hemisphere where the off-line values are generally smaller than

previous estimates by less than ~10 $ms^{-1}day^{-1}$. In the winter hemispheres the disagreement is generally between ~20→40 $ms^{-1}day^{-1}$ or larger, although some results from these previous studies are subject to a high degree of uncertainty. Since the error of the off-line calculations is only ~10% of the presented values, we can confidently state that differences between these values and previous estimations of middle atmospheric drag values, accounting for known biases of these estimations, is due to deficiencies in the USSP scheme, e.g. values of free parameters or deficiencies caused by employing a transparent lid, and not inaccuracies of the off-line calculations.

Furthermore, the noted asymmetry of the NH and SH winter drag values present in previous estimates are considerably larger than the differences found in the off-line values. This asymmetry exists due to the increased amount of convective and frontal generation of gravity waves at tropical latitudes during the SH winter compared to the NH winter (evidence for this is based on averages of cloud cover [Shine, 1989]), the stronger winter zonal winds in the SH stratosphere producing an increased amount of asymmetry in the zonal gravity wave spectrum, and the higher stability of the SH stratospheric polar vortex allowing the free transmission of gravity waves due to the reduction of critical layers for mountain waves [Shine, 1989, Whiteway et al., 1997]. The inability of the USSP scheme to accurately capture the extent of this asymmetry is most likely associated with the use of the constant homogeneous launch spectrum scale factor. Applying this constant forcing at the launch level will not capture the increased convective and frontal generation of gravity waves in the SH. The constant forcing at the launch level will also not capture the localised generation of mountain waves which as discussed above can carry large values of momentum flux. This missing source could be partly responsible for the under-estimation of lower mesospheric forcing by the USSP scheme. Also a significant proportion of the deficiency in the extra-tropics will occur as the parameters which control the amount of momentum flux at the launch level are tuned for the QBO signal in the tropics.

As for the values of forcing in the lower mesosphere of the L50 analyses, the forcing supplied by the USSP scheme in the upper mesosphere is smaller, or on the lower limit, of values diagnosed by the previous studies detailed above. The results from the off-line study do however show similar magnitudes of forcing to those seen in the study of Orr et al. [2010]. However large regions of eastward forcing in the L70 analyses occur at sub-tropical summer latitudes, while maximum summer forcing in Orr et al. [2010] is contained at extra-tropical latitudes. Furthermore, as explained above the off-line calculations for the L70 analyses are likely to overestimate eastward forcing.

7.4 Off-line Turbulent Heating Rates

As for the off-line drag calculations detailed above it is possible to obtain off-line estimates of the heating rates due to gravity wave breaking. In order to obtain an estimate of the heating rates in realistic middle atmospheric temperature and wind fields, the USSP heating rate code was first applied off-line to the CIRA'86 data. Secondly the heating rate code was applied to the L70 analyses fields, with attention drawn to the effects of applying the heating rate code to the biases in temperature and wind fields in the upper mesosphere. Full details of these off-line calculations and results are detailed below.

7.4.1 CIRA'86 Off-line Heating Rates

The offline turbulent heating rate code was applied to the CIRA'86 dataset for the solstice months of June and December. The values of the free parameters in the USSP scheme were set to the standard values given in Table 6.1.7. For these off-line calculations the initial spectrum was launched at ~4200 km (~615 hPa) in the lower troposphere. Figure 7.7 details the total heating rates for the CIRA June and December fields. Also shown for each month are the individual contributions from both terms on the r.h.s. of equation (6.41), i.e the convergence

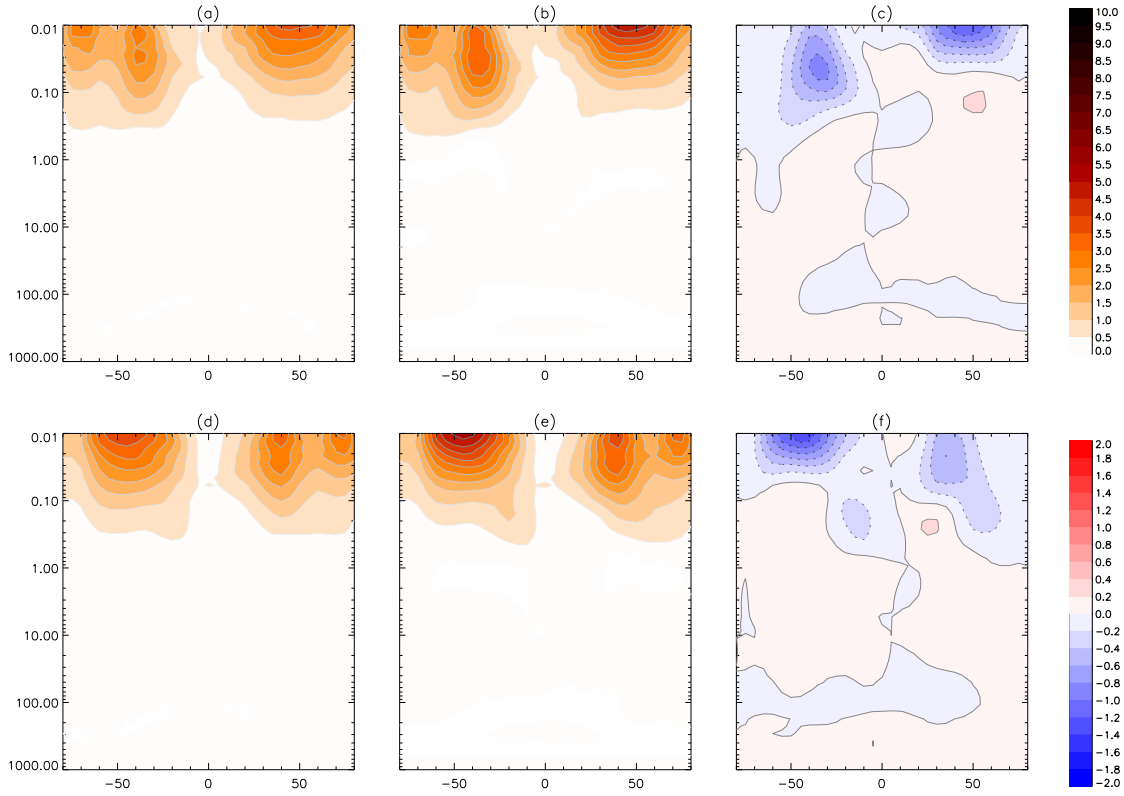

Figure 7.7: Gravity wave heating rates using CIRA data for (a) June (Total); (b) June (Pressure flux convergence); (c) June (Residual work); (d) December (Total); (e) December (Pressure flux convergence); (f) December (Residual work). Contour intervals for (a),(b),(d) and (e) are 0.5 K day^{-1}, while intervals for (c) and (f) are 0.2 K day^{-1} with negative values dashed.

of the wave pressure flux and the residual work.

From Figure 7.7 we find that for both hemispheres the summer months have maximum heating rates of ~3.5→4.0 K day^{-1} approaching ~0.01 hPa between ~40→60^{o} latitude. The winter seasons of each hemisphere have slightly smaller magnitudes of ~2.5→3.0 K day^{-1} approaching ~0.01 hPa which are located nearer the equator at ~35→45^{o}. While the above heating rate values are qualitatively consistent with those from previous studies there is a noticeable difference in the asymmetry between seasons, particularly when comparing against the results

of Becker [2004] which show that summer hemisphere heating rates are twice those seen in the winter season. This seasonal asymmetry has also been found in direct observations of turbulent energy dissipation rates [Lübken, 1992, 1997]. Figure 7.8 shows the average observed dissipation rates from Lübken [1997], Tables 3 and 4, compared to the off-line CIRA calculations at 70^o North for both July and December. The results of Lübken [1997] show that turbulence in the summer hemisphere is confined to a relative small altitude region of ~80→100 km (~0.01→0.001 hPa) where peak heating rates of ~13 K day^{-1} are found at ~90 km, while in the winter hemisphere turbulence is found to occur through out the mesosphere from ~60→100 km with heating rates having maximum values of only ~2 K day^{-1}. In comparison the CIRA calculations show that turbulence extends throughout the entire mesosphere for both summer and winter hemispheres, with both seasons having maximum heating rate values at ~80→85 km (~0.01 hPa). Here the peak values in the winter season are stronger than those observed by Lübken [1997] while in the summer season the peak values are considerably less. For both seasons the peak values also occur at a lower altitude than the observed values. Similar discrepancies are also found between the modelled dissipation rates of Becker [2004], Warner and McIntyre [2001] and the observed measurements of Lübken [1997].

It should be noted that the previous study of Warner and McIntyre [2001], Figure (5), yielded peak energy dissipation rates significantly closer to those from Lübken [1997] when compared to the results of Figure 7.8, however they launched the initial spectrum above the tropospheric jets at an altitude of 19.2 km with alternative values of m_* and β than those used in this study. However the results of Warner and McIntyre [2001] were not as localised in altitude as the observational measurements, noting that a possible explanation for this is the fact that the zonally averaged climatology of the CIRA data maybe a poor approximation in the mesosphere. It should be noted that when applying the standard parameters in Table 6.1.7 to the CIRA data using a launch altitude of

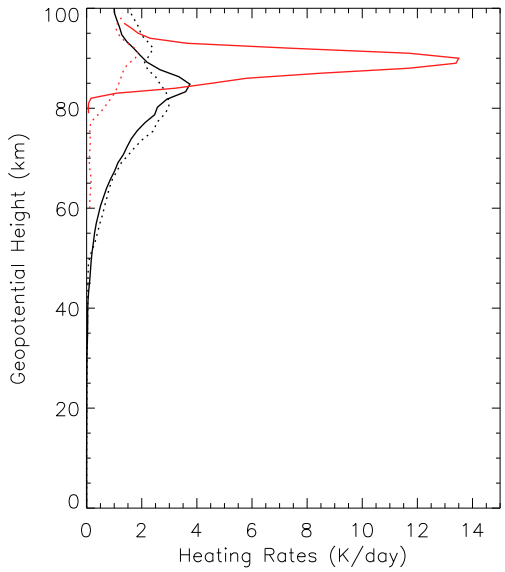

Figure 7.8: Off-line heating rates at 70^o north using CIRA data for summer (Black solid line) and winter (Black dotted line). Observational heating rates from[Lübken, 1997] at 70^o north for summer (Red solid line) and winter (red dotted line).

19.2 km we obtained comparable dissipation rates, not shown, to those described in Warner and McIntyre [2001].

7.4.2 L70 Analyses Off-line Heating Rates

The heating rate code was also applied to the temperature and wind fields of the L70 analyses for June 2010. The values of the free parameters in the USSP were set to the standard values, and as for the L70 off-line drag calculations the initial gravity wave spectrum was launched from 681 hPa and the transparent lid condition was employed. Figure 7.9 details the total heating rates for the L70 June analyses, along with the individual contributions from the wave pressure flux and the residual work.

From Figure 7.9 we find that approaching 0.01 hPa the total heating rates in the

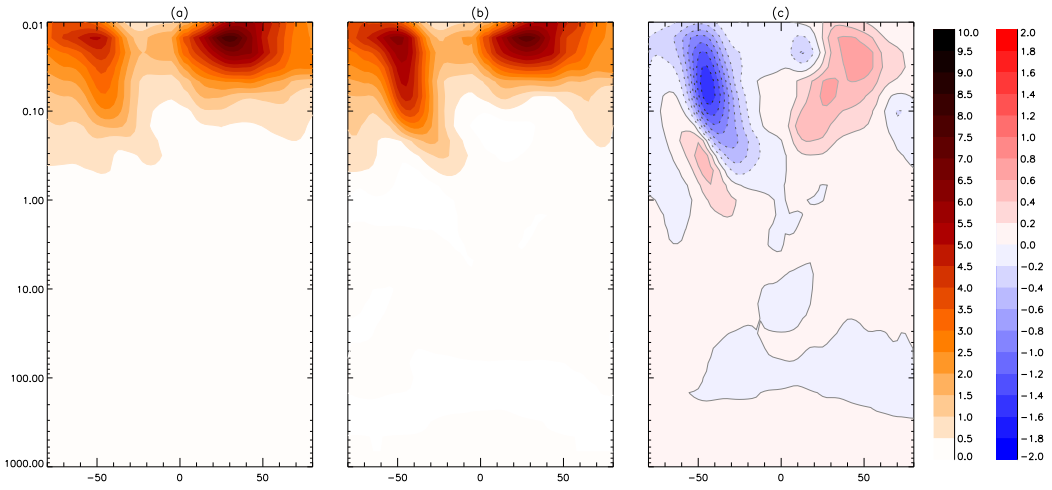

Figure 7.9: Gravity wave heating rates using L70 analyses for June (a) Total; (b) Pressure flux convergence; (c) Residual work. Contour intervals for (a) and (b) are 0.5 K day^{-1}, while intervals for (c) are 0.2 K day^{-1} with negative values dashed.

SH winter hemisphere have maximum magnitudes of ~4→4.5 K day^{-1} centred at ~50^{o} latitude. In the NH summer hemisphere approaching 0.01 hPa the total heating rates have maximum magnitudes of ~7.5→8.0 K day^{-1} centred between ~30→40^{o} latitude. These results show the difference between maximum heating rates of both seasons is ~4 K day^{-1}, noticeably larger than the difference seen in Figure 7.7. Such a large difference is not predicted by previous numerical estimates of gravity wave heating rates at ~80 km, where the studies of Akmaev [1994], Becker [2004], Fomichev et al. [2002] all show an asymmetry of ~2 K day^{-1} between seasons under solstice conditions. However, as for all gravity wave schemes these values are obtained by somewhat arbitrary settings of free parameters and therefore have high uncertainty. Additional confidence in the suggestion that the asymmetry of heating between seasons in Figure 7.9 is overestimated can be obtained by comparison with the observational values of Lübken [1992, 1997] given in Figure 7.8. Here such a large degree of asymmetry in the observational values is only seen above the L70 model lid at ~85 km altitude. The SH winter season heating rates of ~4→4.5 K day^{-1} at ~0.01 hPa are

qualitatively similar to those diagnosed by previous studies. It is therefore most probable that the larger asymmetry described above is due to overestimated NH summer heating rates when using the June L70 analyses.

One possible reason for the over estimation of the NH summer heating rates when using L70 analyses is the known temperature and wind biases of these fields, which as shown in Sections 4.6.3 and 5.2 have larger magnitude approaching the model lid and for the summer hemisphere. From Figure 7.9 (c) we find that the residual work contribution to the total heating rates in the NH summer upper mesosphere is positive, compared to negative in Figure 7.7 (c). This positive residual work contribution is a direct result of the biased zonal winds in the summer upper mesosphere for June 2010, where as shown by Figure 5.3 the westward winds of this region have positive westward vertical wind shear compared to the negative sign of climatological values. It is this incorrect vertical wind shear that results in the positive residual work contribution to the total NH summer heating rates of Figure 7.9, thus increasing the asymmetry between winter and summer seasons. This result highlights the sensitivity of heating rate calculations when applying equation (6.41) to bias wind fields, such as those of the L70 analyses in the upper mesosphere.

7.5 Off-line Drag Sensitivity Experiments

All of the evidence presented in previous sections suggests that the gravity wave forcing supplied by the USSP scheme is too weak. As stated above the USSP scheme has several free parameters which can be tuned, within certain limits of constraint. It is therefore possible to tune the USSP via the free parameters to produce larger magnitudes of gravity wave forcing. Before performing expensive numerical experiments with the UM, see Section 8, the sensitivity of forcing to certain free parameters can be estimated using off-line experiments as discussed below.

The limit to which each of the free parameters in the USSP scheme can be tuned depends on both observational and theoretical constraints. Below is a discussion of the constraints placed on the free parameters in Table 6.1.7. This discussion separates the free parameters into those which are tightly and loosely constrained by either observations or theory, thus highlighting certain candidates to study via off-line sensitivity experiments.

7.5.1 Previous Experiments and Constraints

Of all the free parameters in the USSP scheme the power law t is most well constrained by observational studies of temperature and horizontal wind fields. As detailed in equation (6.2) and shown by numerous studies, gravity wave energy spectra have a slope at large vertical wavenumbers $\propto m^{-t}$, which tends to be roughly independent of time, location and altitude. The studies of Allen and Vincent [1995], Hertzog et al. [2001], Smith et al. [1987], VanZandt [1982] all show that $t \approx 3$, i.e the operational value.

Increasing the number of azimuthal sectors, i.e. decreasing the angle of each azimuthal sector, will increase the computational expense of the USSP scheme. Off-line tests (not shown) proved that increasing the number of azimuthal sectors resulted in relatively small differences in the magnitude of gravity wave forcing, consistent with the previous full 3D model results of Warner and McIntyre [1996].

In the operational USSP scheme the characteristic vertical wavenumber m_* has an associated characteristic vertical wavelength λ_*, given by the relation $m_* = 2\pi/\lambda_*$. As shown by Ern et al. [2006] is well an initial characteristic vertical wavelength launch value, defined by λ_*^{launch}, will increase with height due to the growth of gravity wave amplitudes with decreasing density. This results in an extension of the quasi-saturated part of the gravity wave spectrum and shifts m_* towards lower vertical wavenumbers [Fritts and Vanzandt, 1993]. Previous studies by Ern et al. [2006] have calculated the variation of λ_* throughout the stratosphere

for various values of λ_*^{launch}. By comparing these results with the observational vertical wavelength values of $\sim 2-5$ km in the lower stratosphere reported by Allen and Vincent [1995], Hertzog et al. [2001], Tsuda and Hocke [2002], constraints on the realistic values of λ_*^{launch} could be inferred. It was concluded that optimum values of λ_*^{launch} should be between ~2-4 km, otherwise values of λ_* in the lower stratosphere would not be within the observational range. Also Ern et al. [2006] measured the sensitivity to zonal mean zonal GWD to changes in various values of λ_*^{launch}. The results showed increasing λ_*^{launch} resulted in larger GWD values for lower altitudes. This is a consequence of larger λ_*^{launch} values producing initial launch spectrums with larger quasi-saturated parts, hence greater momentum flux deposition at lower altitudes. As previously noted, the value of m_* has been tuned to achieve a realistic QBO signal with the currently operational value at the upper limit of the constraints suggest by Ern et al. [2006].

Previous experiments at the Met Office [Bushell, 2010] have shown that the magnitude of momentum forcing supplied by the USSP scheme is highly sensitive to the choice of launch spectrum scale factor C_{l0}. While this parameter is only loosely constrained, based on the magnitude of the forcing produced, as for m_* the operational value of C_{l0} has been tuned to achieve a realistic QBO signal at tropical latitudes.

Observational studies of temperature and horizontal wind variances [Cot, 2001, Hertzog and Vial, 2001, Tsuda et al., 2004] have shown that values of p vary between $\sim 1 \rightarrow 2$, with the majority of results obtaining values close to $p = 5/3$. Hence this parameter is fairly well constrained and as in Fritts and Lu [1993] the value of $p = 5/3$ is used in the standard operational USSP scheme.

The operational value of m_{min} was chosen to give energy dissipation rates which are comparable with those from the observational studies of Lübken [1992, 1997]. As noted in Warner and McIntyre [1999] the momentum forcing is very sensitive

to the choice of m_{min}, however only at altitudes above ~90 km (above the L70 model lid) where the momentum forcing supplied by the scheme begins to decrease. They noted that momentum forcing starts to decrease at lower altitudes if a larger value of m_{min} is chosen. Therefore, typically the sensitivity related to m_{min} is only seen above the L70 model lid. **Based on the above considerations the number of azimuthal sectors and the values of t, p, m_*, C_{l0} and m_{min} are not suitable parameters to use in tuning of the USSP scheme.**

In the USSP scheme the energy scale factor β for the launch spectrum is directly proportional to the amount of gravity wave momentum flux at all altitude levels. Hence variation in β scales the magnitude of deposited momentum flux and derived GWD, without changing the overall distribution. From theoretical assumptions the value of β is $\sim 10^{-1}$ [Warner and McIntyre, 1996, Fritts and Alexander, 2003], with an uncertainty factor of ~ 2. As noted in Ern et al. [2006] the value of β simply scales the magnitude of momentum flux in the USSP scheme.

The power law s is not well constrained theoretically or by observation [Fritts and Alexander, 2003], as corresponding vertical wavelengths are too large to be resolved practically. This part of the gravity wave spectrum corresponds to deep gravity waves which are believed to carry significant proportions of the momentum flux spectrum into the mesosphere. Using the USSP scheme Ern et al. [2006] measured the sensitivity of zonal mean zonal GWD to changes in the standard value of $s = 1$. While the magnitudes of peak GWD values were unaffected, differences did occur for the vertical distributions. For $s < 1$, where momentum flux is increased for the unsaturated part of the spectrum, larger parts of the gravity wave spectrum are deposited at lower altitudes, resulting in larger values of GWD at lower altitude when compared to results for $s = 1$. By similar arguments when $s > 1$, the quasi-saturated part of the launch spectrum is reduced, resulting in larger GWD values being shifted to higher altitudes.

Based on the above considerations the parameters most ideally suited to tune the USSP scheme are β and s. The sensitivity of momentum forcing produced by the USSP scheme to these parameters using off-line calculations is discussed below.

7.5.2 Drag Sensitivity to Energy Scale Factor β

Figure 7.10 details the off-line gravity wave drag values when using perturbed values of β for the L70 analyses fields of June 2010. The results shown are the perturbed drag values when β is varied from 0.11→0.19 minus those obtained using the standard value of β=0.10227987. Here blue contours in the SH winter hemisphere correspond to an increase of the seasonal westward forcing, and red contours in the NH summer hemisphere correspond to an increase of the seasonal eastward forcing.

From Figure 7.10 we find that for increasing values of β the drag seen in each hemisphere increases, as expected [Ern et al., 2006]. Here β=0.14 results in increased forcing with maximum magnitudes of ~12→16 ms^{-1} day^{-1} in the winter upper mesosphere, and ~8→12 ms^{-1} day^{-1} in the summer upper mesosphere. Further increasing β=0.18 results in further increases of maximum forcing magnitudes, with ~24→28 ms^{-1} day^{-1} in the winter and ~12→16 ms^{-1} day^{-1} in the summer upper mesosphere. As noted above increasing the value of β effectively scales the amount of momentum flux for each model level, hence we find that the values of the drag in the winter hemisphere have a higher sensitivity to increasing β since they initially contained a larger vertical divergence of momentum flux.

7.5.3 Drag Sensitivity to Un-saturated Spectral Power Law s

Figure 7.11 details the off-line gravity wave drag values when using perturbed values of spectral power law s for the L70 analyses fields of June 2010. The results shown are the perturbed drag values when s is varied from 0.6→0.15

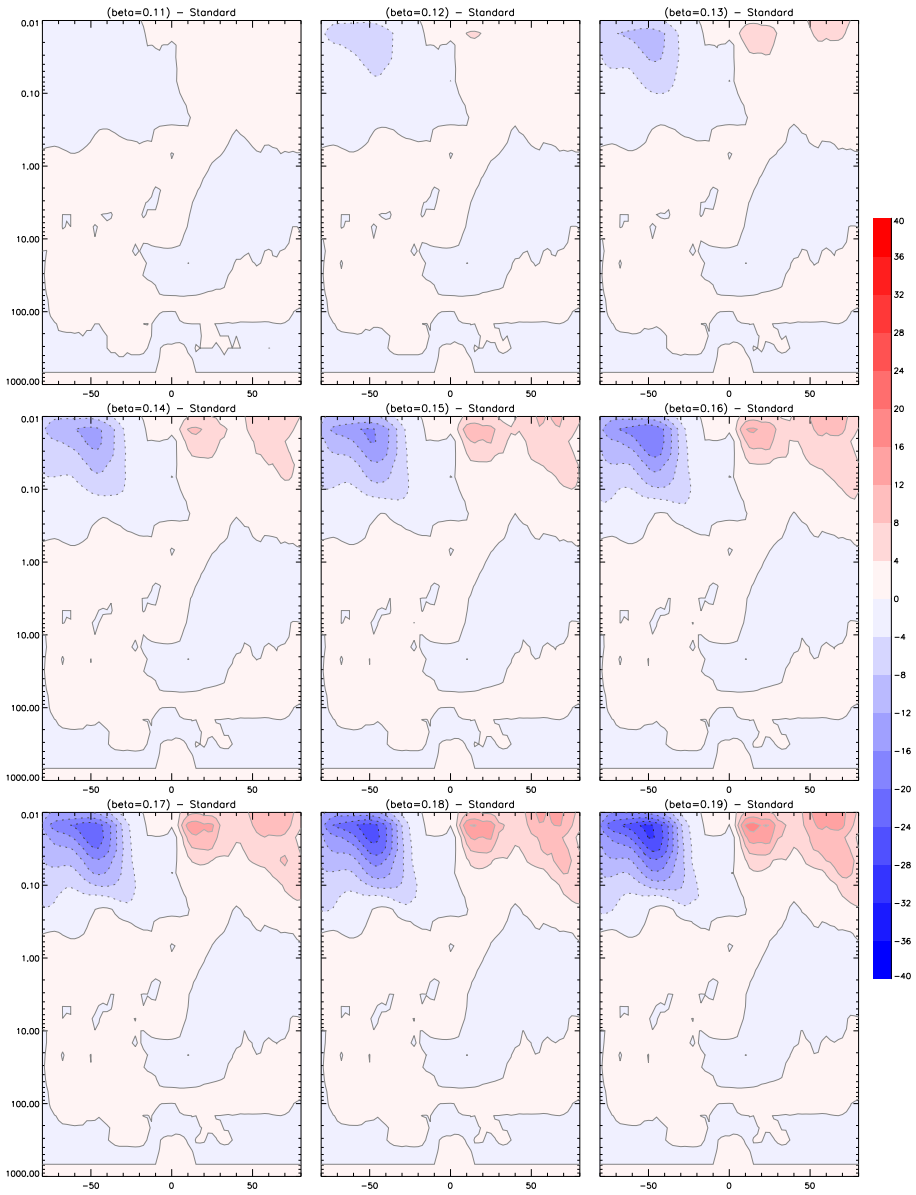

Figure 7.10: Off-line sensitivity of monthly zonal mean zonal gravity wave drag to perturbed values of β using L70 analyses fields for June 2010. Plots show the perturbed drag values minus those using standard values for β varying from 0.11→0.19. Contour interval is 4 ms^{-1} day^{-1} with negative values dashed.

minus those obtained using the standard value of s=1. Here blue (red) contours in the SH winter hemisphere correspond to an increase (decrease) of the seasonal westward forcing, and red (blue) contours in the NH summer hemisphere cor-

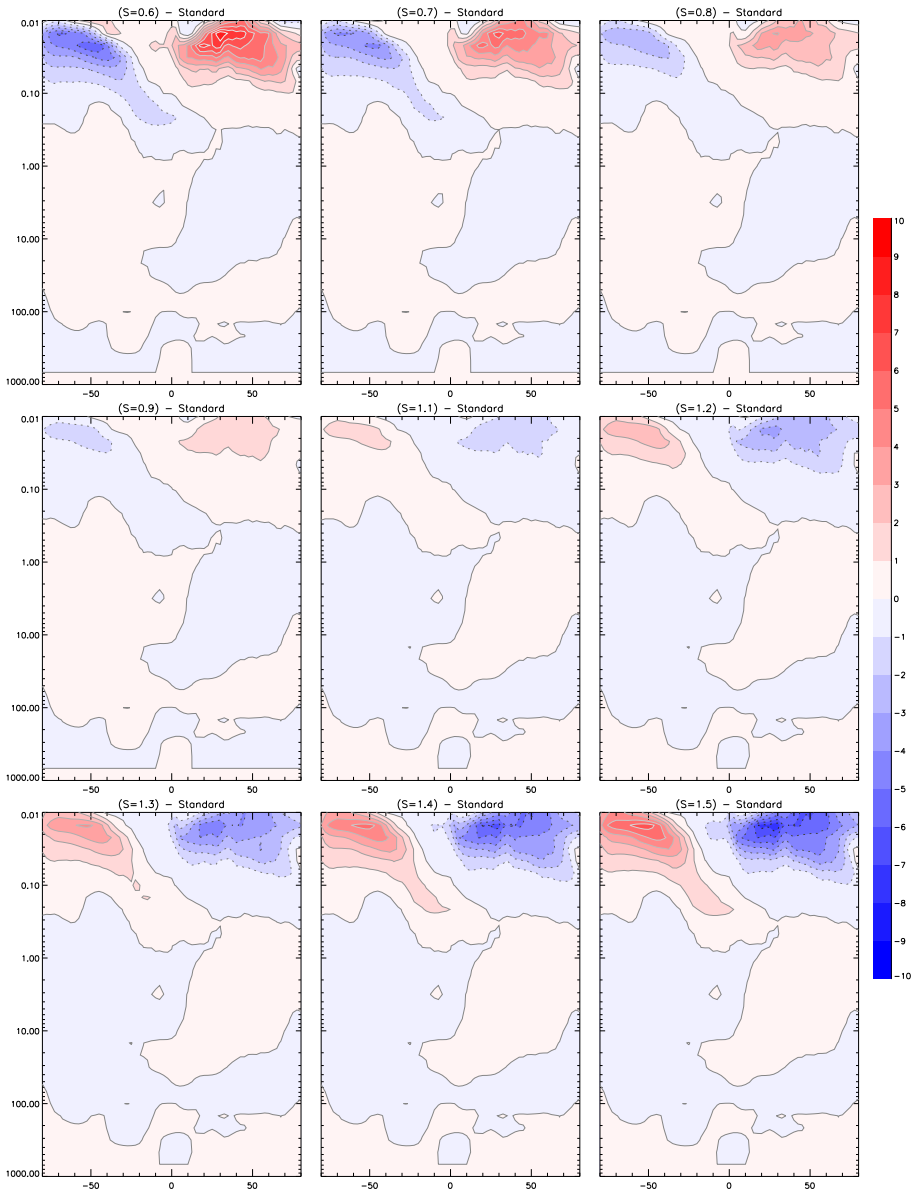

Figure 7.11: Off-line sensitivity of monthly zonal mean zonal gravity wave drag to perturbed values of spectral power law s using L70 analyses fields for June 2010. Plots show the perturbed drag values minus those using standard values for s varying from 0.6→1.5. Contour interval is 1 ms^{-1} day^{-1} with negative values dashed.

respond to an increase (decrease) of the seasonal eastward forcing.

From Figure 7.11 we find that reducing the value of s produces an increase of the

seasonal westward (eastward) forcing in the winter (summer) hemisphere. The magnitude of the such differences are ~3→4 ms^{-1} day^{-1} for $s = 0.8$ and ~6→8 ms^{-1} day^{-1} for $s = 0.6$. Consistent with the previous results of Ern et al. [2006] lowering the value of s produces large values of momentum forcing at lower altitudes, however notable differences only occur for the largest perturbations of s. Similar magnitudes in the differences of momentum forcing are seen when increasing the value of s. Here an increase of s produces a decrease of the seasonal westward (eastward) forcing in the winter (summer) hemisphere.

7.5.4 Drag Sensitivity to Opaque Lid Condition

In addition to the results for the free parameters β and s described above it is also possible to asses the impact of applying the opaque lid condition off-line for the L70 analyses, where as shown by Figure 7.4 significant proportions of the launch momentum flux remain at the model lid. Such off-line experiments (not shown) produced an increase in the gravity wave forcing seen at the model lid when compared to off-line magnitudes when using the transparent lid condition. Here increases in magnitude were ~10 ms^{-1} day^{-1}, however as expected differences are only seen at the model lid. The impact of such increases are expected to have a significant impact on the atmosphere below through downward control and subsequent feedbacks in gravity wave forcing. As such feedbacks are not represented by the single timestep off-line studies a fully detailed account of applying the opaque lid condition is not given here, however is reported in the UM experiments performed in Section 8.

7.6 Off-line Direct Heating Sensitivity Experiments

As for the values of momentum forcing discussed above the sensitivity of the direct heating rates to tuning free parameters in the USSP scheme can be examined off-line. As for the momentum forcing attention is restricted to the parameters β and s, results and discussion follow.

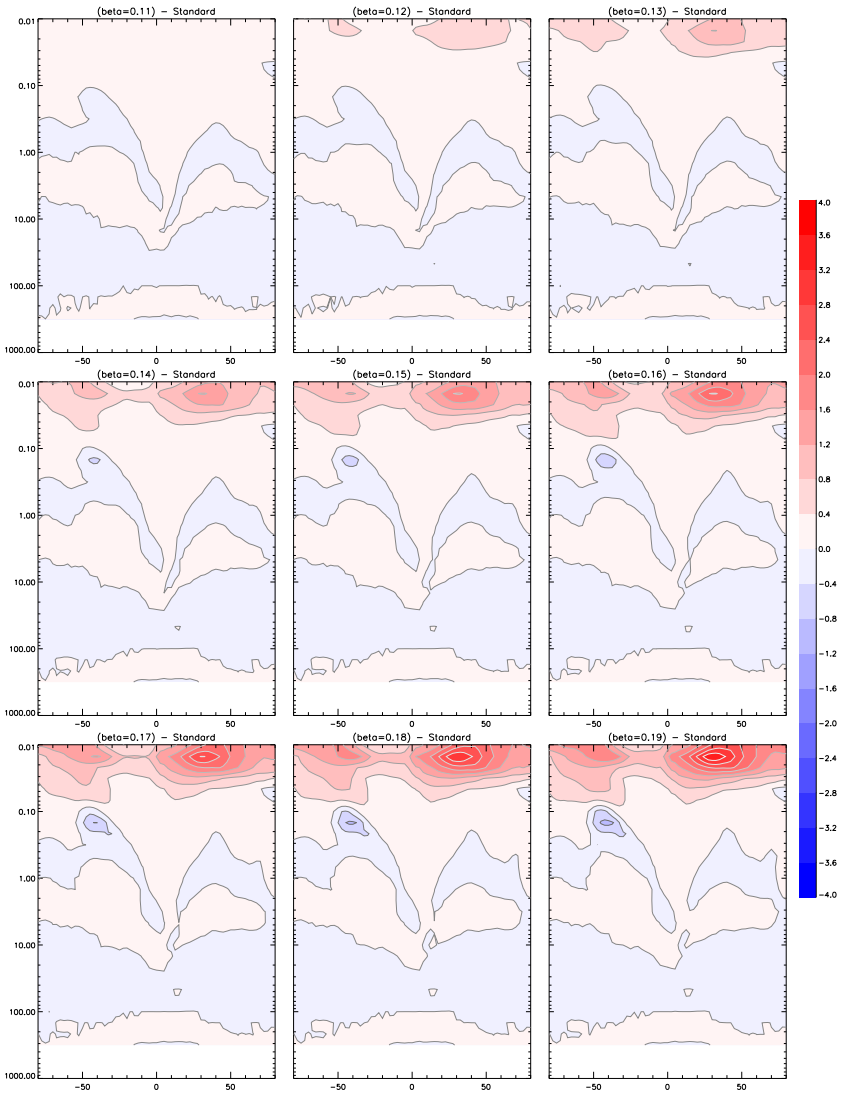

Figure 7.12: Off-line sensitivity of monthly zonal mean gravity wave direct heating to perturbed values of β using L70 analyses fields for June 2010. Plots show the perturbed heating rates values minus those using standard values for β varying from 0.11→0.19. Contour interval is 0.4 K day^{-1} with negative values dashed.

7.6.1 Heating Rate Sensitivity to Energy Scale Factor β

Figure 7.12 details the off-line gravity wave heating rate values when using perturbed values of β for the L70 analyses fields of June 2010. The results shown are the perturbed heating rate values minus those obtained using standard values

for β varying from 0.11→0.19. Here blue contours correspond to a decrease and red contours to an increase of the direct heating rates obtained using the standard value of β in Table 6.1.7.

From Figure 7.12 we find that increasing the value of β increases the heating rates seen at the upper levels of the model domain. Largest increases are seen in the upper mesosphere of the summer NH where β=0.14 and β=0.18 increases the heating rate by ~1.2→1.6 K day^{-1} and ~2.4→2.8 K day^{-1} respectively. The upper mesosphere of the winter SH is less sensitive that the NH where β=0.14 and β=0.18 increases the heating rate by ~0.8→1.2 K day^{-1} and ~1.2→1.6 K day^{-1} respectively.

7.6.2 Heating Rate Sensitivity to Un-saturated Spectral Power Law *s*

Figure 7.13 details the off-line gravity wave heating rate values when using perturbed values of spectral power law s for the L70 analyses fields of June 2010. The results shown are the perturbed heating rate values minus those obtained using standard values for s varying from 0.6→0.15. Here blue contours correspond to a decrease and red contours correspond to an increase of the direct heating rates obtained with the standard operational value of s=1.0.

From Figure 7.13 we find that reducing the value of s typically produces an increase of the direct heating rates in the upper mesosphere. Here a value of s=0.6 produces a maximum increase of ~0.6→0.8 K day^{-1} in the SH and ~0.8→1.2 K day^{-1} in the NH. Increasing the value of s results in a decrease of the direct heating rates in the upper mesosphere, where a value of s=1.5 produces a decrease of ~0.6→0.8 K day^{-1} in both hemispheres.

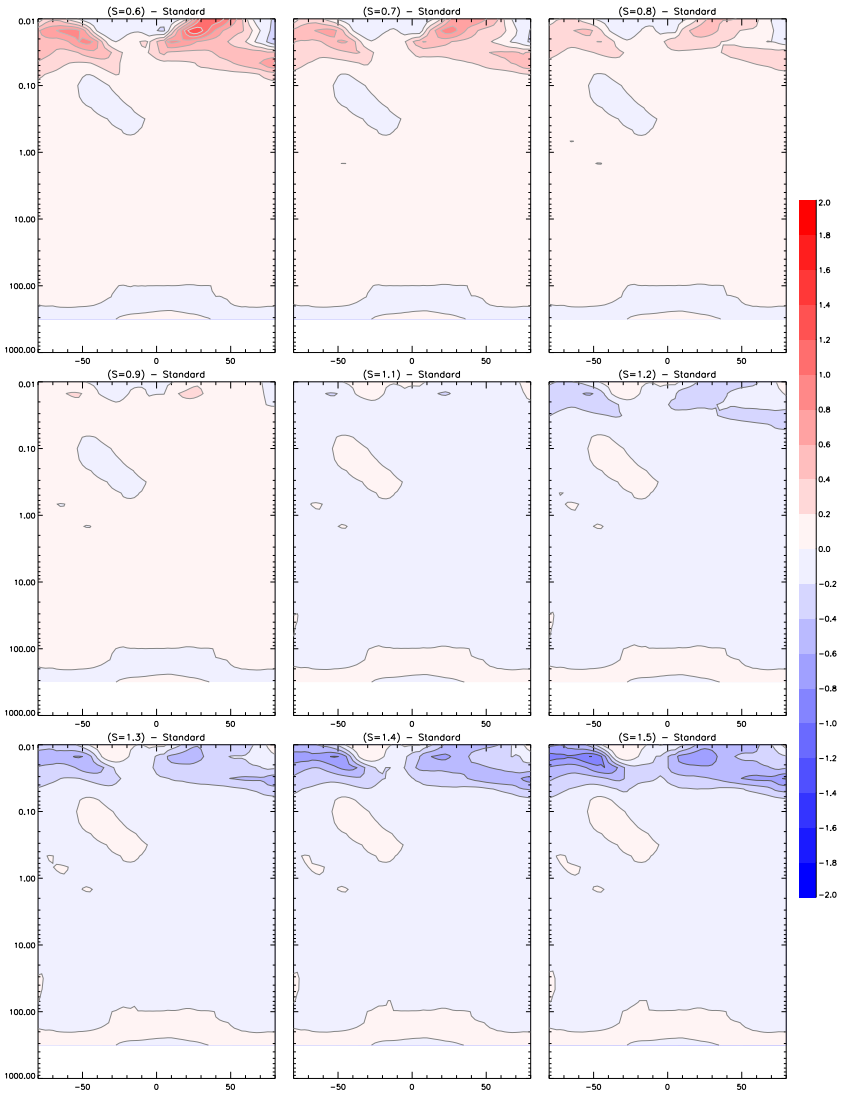

Figure 7.13: Off-line sensitivity of monthly zonal mean gravity wave direct heating to perturbed values of spectral power law s using L70 analyses fields for June 2010. Plots show the perturbed heating rates values minus those using standard values for s varying from 0.6→1.5. Contour interval is 0.2 K day^{-1} with negative values dashed.

7.7 Summary

In this chapter we have detailed values of the gravity wave forcing in the L50 and L70 analyses using off-line estimates. For both the L50 and L70 analyses we have shown that the off-line values are likely within ~1-3 ms^{-1} day^{-1} of actual

model fields.

For the L50 analyses gravity wave forcing approaching the model lid has maximum magnitudes of ~12.5→17.5 ms^{-1} day^{-1} for both summer hemisphere seasons. In the NH winter season maximum forcing magnitudes are also ~12.5→17.5 ms^{-1} day^{-1}, while the SH winter season has larger magnitudes of ~17.5→22.5 ms^{-1} day^{-1}. These forcing magnitudes are weaker, or at the lower limit, of previous estimates for gravity wave forcing of the lower mesosphere. This result is consistent with the analyses of the temperature biases in the lower mesosphere, discussed in Section 4.3.1.

For the L70 analyses gravity wave forcing approaching the model lid has maximum magnitudes of ~35→40 ms^{-1} day^{-1} for both summer hemisphere seasons. The SH winter season has larger maximum forcing magnitudes of ~50→55 ms^{-1} day^{-1} approaching the model lid, while the NH winter season has larger magnitudes of ~60→65 ms^{-1} day^{-1}. However the forcing magnitudes of the NH winter are likely to be unrealistic of the L70 analyses due to the uncharacteristic cold polar vortex in the lower mesosphere of this season, as discussed in Section 4.5. As for the L50 analyses, these estimates of gravity wave drag are weaker, or at the lower limit, of previous estimates in the upper mesosphere. Again, this weak forcing is consistent with the analysis of L70 temperature biases given in Section 4.6.3.

The sensitivity of the gravity wave forcing produced by the USSP to perturbing free parameters of the scheme was also investigated. Based on observational and theoretical constraints the parameters β and s were chosen for the sensitivity study. It was found, compared to standard values, that using β=0.14 increased the forcing by ~12→16 ms^{-1} day^{-1} in the winter upper mesosphere and ~8→12 ms^{-1} day^{-1} in the summer upper mesosphere. The sensitivity to perturbed values of s was smaller than the above values for β, where using a large perturbation

of s=0.6 increased the forcing by ~6→8 ms^{-1} day^{-1} in both the summer and winter seasons. Based on these results the obvious choice of parameter to vary in the UM experiments is β, further reasoning and details are given in Section 8.

Off-line calculations of the direct heating rates due to gravity wave breaking were also performed using standard values of the free parameters in the USSP scheme. When applied to the L70 analyses fields for June 2010 maximum heating rates of ~4→4.5 K day^{-1} were seen in the SH winter and ~7.5→8.0 K day^{-1} in the NH summer approaching the model lid. While these winter season heating rates are consistent with previous studies and observational estimates, summer value are noticeably larger by ~2~2.5 K day^{-1}. This difference is most likely attributed to the incorrect vertical wind shear of the summer season zonal mean winds, which act to increase the heating rate magnitudes through the residual work term of equation (6.41). This result highlights the influence that bias wind fields may have on the magnitude of heating rates obtained using the USSP formulation. As for the momentum forcing, the sensitivity of direct heating rates to perturbed free parameters in the USSP scheme was also investigated. Here the largest sensitivity was again found when perturbing the free parameter β, where using β=0.14 results in an increase of ~1.2 K day^{-1} in both winter and summer hemispheres.

Chapter 8

Unified Model Experiments

This chapter details several numerical experiments performed with the Met Office Unified Model. Based on the results of previous sections it is apparent that small scale gravity wave forcing provided by the USSP scheme is insufficient and such a deficiency results in significant systematic temperature biases in the middle atmosphere. Hence experiments were designed to examine the impact on the temperature and wind fields of the middle atmosphere when the gravity wave forcing was increased, through both tuning of the USSP scheme and applying the opaque lid condition forcing momentum flux to be conserved throughout the model domain. In addition to this, experiments were also designed to examine the impact on temperature and wind fields when employing gravity wave turbulent heating rates in the USSP scheme, full details follow.

8.1 Experimental Background

All of the experiments described in this section were performed using a free running forecast version of the Unified Model, i.e. no data assimilation cycle included. Using a free running version allows the impact of each experiment on the model forecast to be assessed without biases being corrected by increments from the assimilation cycle. This approach is further justified as the currently operational Met Office system only assimilates data up to the stratopause, i.e. is

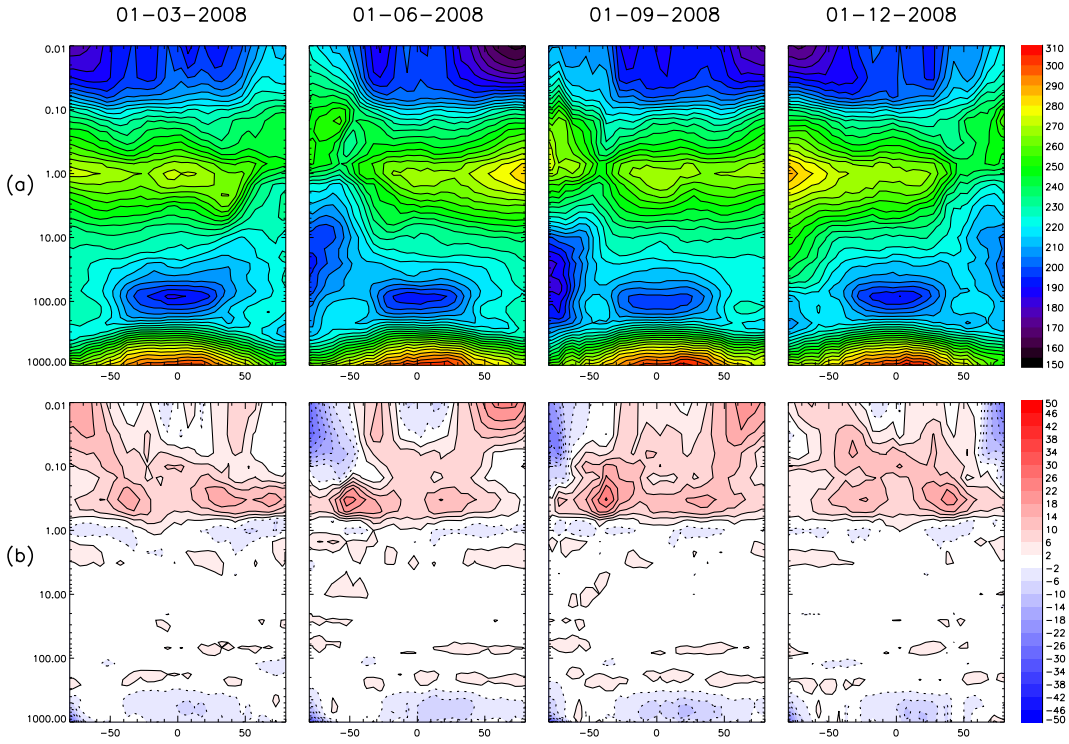

Figure 8.1: (a) Zonal mean temperature and (b) zonal mean temperature bias w.r.t to EOS MLS data of ECMWF operational analyses for the 1^{st} March, June, September and December of 2008. Contour interval for zonal mean temperatures is 5 K. Contour interval for temperature bias is 4 K with negative values dashed.

a free running model throughout the mesosphere.

The experimental model domain extends from the ground to 85 km (~0.01 hPa) employing the 85 level vertical grid configuration detailed in Figure 3.1 (c). The 85 level configuration was chosen as the increased vertical resolution has been shown to reduce stratospheric biases for forecasts over ~10-15 days [Mahmood, 2010], and is the currently preferred configuration for Met Office global seasonal and climate forecasting models. Furthermore it is anticipated that the operational global analyses will soon be up-graded to 85 levels.

To investigate the model response during the solstice and equinox seasons of

both hemispheres the experimental forecasts were initialised using ECMWF operational analyses [Gobiet et al., 2005] for the 1^{st} of March, June, September and December of 2008. The zonal mean temperature structure and zonal mean temperature bias with respect to EOS MLS data of these initial conditions is shown in Figure 8.1. From Figure 8.1 we find that for all seasons throughout the mesosphere the ECMWF analyses have temperature biases which are noticeably different from those found in the corresponding operational L70 analyses of Figure 4.10. A "spin up" time of one month was therefore employed to evolve the biases present in the initial conditions to those more characteristic of the Met Office Unified Model. As such the model fields from all experiments are analysed from the 1^{st} of April, July and October of 2008 and January of 2009 for a two week period. This two week period was chosen due to the fact that model forecasts over 10→15 days are unable to effectively predict the occurrence of sudden stratospheric warmings [Jackson, 2010]. As sudden stratospheric warming events involve a reversal of the zonal wind the occurrence of such an event will significantly alter the gravity wave filtering that occurs in the winter stratosphere, and subsequently the gravity wave forcing and temperature structure of the upper mesosphere winter hemisphere. A major sudden stratospheric warming occurred in the NH during the latter half of January 2009 over ~50 days from the start of the model forecast, and thus will not be well reproduced in the model fields for this period. To limit the influence of poorly representing the above stratospheric warming the analysis of experimental results is restricted to the two week periods noted above, thus allowing for a clearer comparison between the northern and southern hemisphere winter seasons. Full details of each model experiment follow.

8.2 Control Run

In order to compare the model response for the different experiments a control run for each season was performed. The control integrations employed the Li and Shine ozone climatology and identical solar spectrum and spectral files used

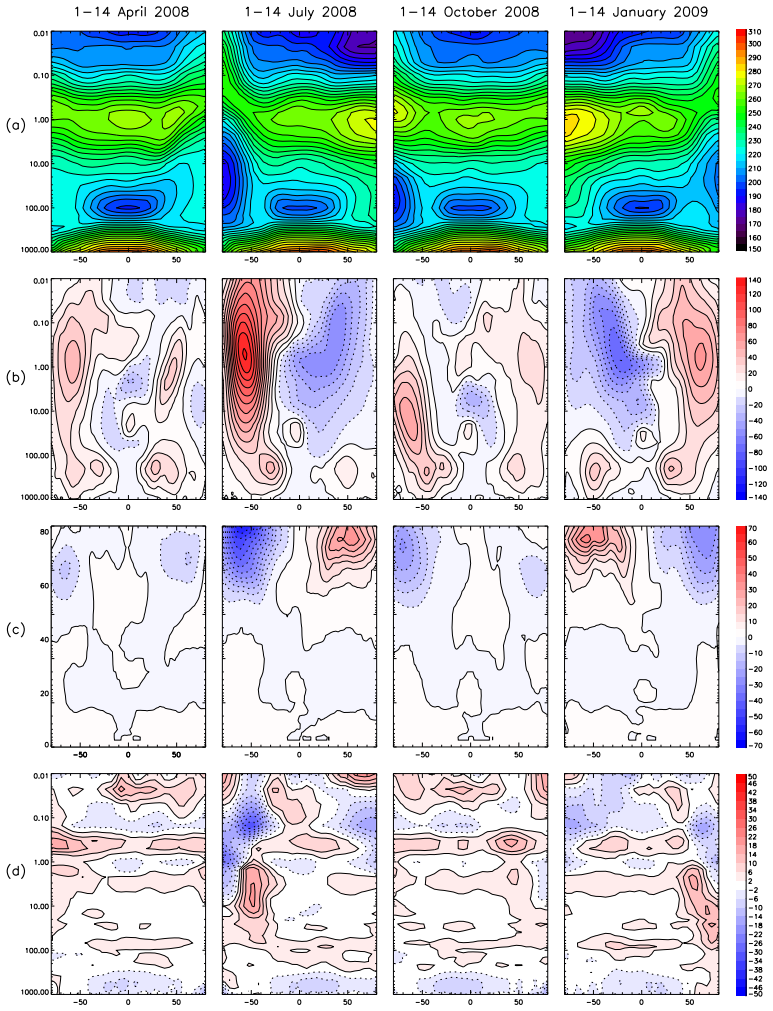

Figure 8.2: Control run model fields (a) zonal mean temperature; (b) zonal mean zonal wind; (c) zonal mean gravity wave drag and (d) zonal mean temperature error with respect to EOS MLS of the control run averaged from the 1^{st} to 14^{th} of April, July, October 2008 and January 2009. Contour interval is 5 K for temperature. Contour intervals are 10 ms^{-1} for zonal wind, 5 ms^{-1} day^{-1} for gravity wave drag and 4 K for temperature error with negative values dashed.

in the operational Edwards-Slingo radiation scheme. The free parameters of the USSP scheme were set to the operational values of Table 6.1.7 and the transparent lid condition was employed. Figure 8.2 details the zonal mean temperature, zonal mean zonal wind, zonal mean zonal gravity wave drag and the zonal mean

temperature errors of the control run averaged from the 1^{st} to the 14^{th} of April, July and October 2008 and January 2009. Figure 8.3 details the divergence of EP flux, first term on the r.h.s of Equation 2.6, averaged over the 1^{st} to 14^{th} for July 2008 and January 2009 for the control experiment, and all of the subsequent experiments described below. As noted in Section 2.3.3 the divergence of EP flux represents the contribution of resolved forcing, usually attributed to Rossby waves, in the zonal momentum budget of the TEM equations. Also detailed in Figure 8.3 are the zonal mean meridional and vertical components ($\bar{v}^*, \bar{w}^*$) of the TEM equations averaged over the 1^{st} to 14^{th} for July 2008 and January 2009 for all experiments.

We are primarily concerned with responses of the UM experiments compared to the control run and how such responses could possible affect the operational analyses. However, it is possible that the magnitude and distribution of such responses maybe influenced by the initial state of the control run fields and or affected by differences between the model configurations, i.e. assimilation of data or vertical resolution. Thus, in order to qualitatively assess the response the different UM experiments may possibly have on the operational analyses, it is first necessary to compare the temperature biases of the control run against those of the L70 analyses. A comparison of Figure 8.2 (d) and Figure 4.10 reveals that in the upper stratosphere and lower to middle mesosphere temperature biases for all seasons are qualitatively similar, with both latitude independent warm and cold biases from ~0.9→0.3 hPa and ~0.3→0.05 hPa respectively having comparable magnitudes. However the above cold biases of the control run are generally ~2→10 K smaller than those seen in the L70 analyses at tropical latitudes. The similarity of the above latitude independent biases in the control run and the L70 analyses, following the spin up time of one month, would place additional confidence they are caused by the radiative deficiencies discussed in Chapter 4. In the upper mesosphere the temperature biases of the equinox control months are qualitatively similar to those of the L70 analyses for similar seasons, however

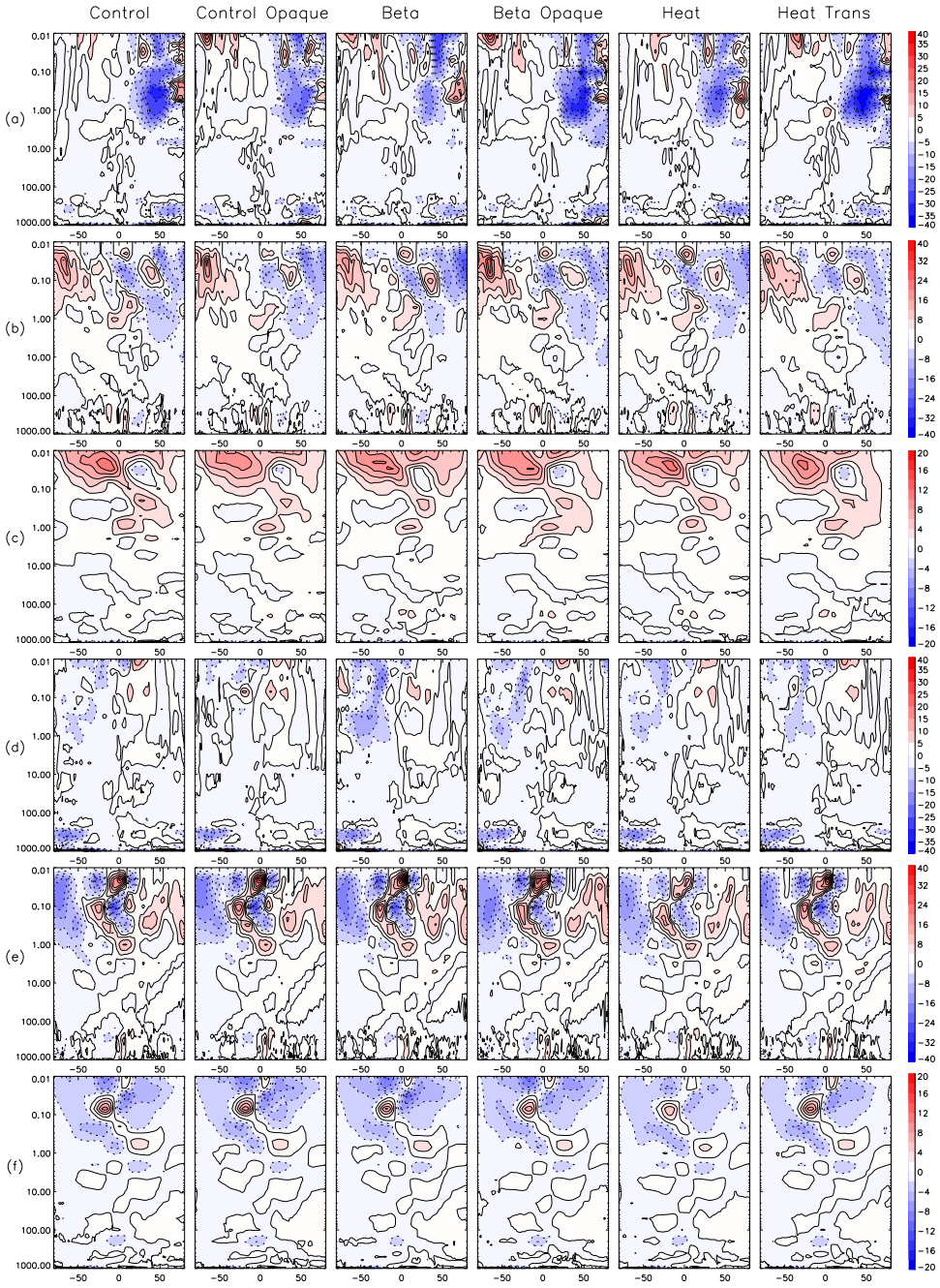

Figure 8.3: Diagnostic fields for all UM experiments zonally averaged from the 1^{st} to 14^{th} of July 2008 and January 2009. (a) Divergence of EP flux for January; (b) $\bar{w}^*$ for January; (c) $\bar{v}^*$ for January; (d) Divergence of EP flux for July; (e) $\bar{w}^*$ for July and (f) $\bar{v}^*$ for July. Contour interval is 5 ms^{-1} day^{-1} for (a) and (d), 4 mms^{-1} for (b) and (e) and 4 ms^{-1} for (c) and (f). Negative contour values are dashed.

there are noticeable differences in the extra-tropical upper mesosphere under solstice conditions. Here the control run warm biases of the SH winter extra-tropical upper mesosphere are ~10→18 K, contrasting with the ~10→14 K cold bias seen in the L70 analyses. In the NH winter extra-tropical upper mesosphere the control run has temperature biases which are ~2~6 K colder than those seen in the L70 analyses. For the control run summer seasons warm bias of ~18→22 K and ~2→10 K are seen the extra-tropical upper mesosphere of the NH and SH respectively. These warm biases are characteristic of the L70 analyses, however magnitudes are noticeably smaller (by ~36→40 K for the NH and ~46→50 K for the SH) than those presented in Figure 4.10.

The discrepancies between control run and L70 analyses biases in the SH winter upper mesosphere are most likely due to the differences seen in gravity wave forcing of this region. Here we find from Figure 8.2 (c) that maximum values of westward forcing in the control run are ~10→15 ms^{-1} day^{-1} stronger than the ~45→50 ms^{-1} day^{-1} values seen in the L70 off-line studies of Figure 7.6. Following the theory of Section 2.3.5 we would therefore expect the control run temperatures of the SH extra-tropical upper mesosphere during July to be warmer than those of the L70 analyses, as noted. A probable explanation for this discrepancy in gravity wave forcing is the assimilation of data throughout the stratosphere in the L70 analyses, where through gravity wave filtering the constrained temperature and wind fields result in a different forcing of the mesospheric circulation, compared to the free running control run. Here we find that for the July control run there is a warm bias of ~14→18 K from ~40→60^{o} in the SH, which is not present in the L70 analyses. This warm bias results in control run stratospheric meridional temperature gradients, which decrease towards the cold SH winter pole, having larger values than those seen in the L70 analyses and as expected from thermal wind balance stronger seasonal eastward winds. From previously explained wave filtering arguments the stronger zonal mean winds of the control run in the stratosphere result in increased westward gravity

wave forcing in the upper mesosphere.

Similar arguments to those above also follow for the NH winter season, where as detailed in Figure 8.2 (d) there is a warm bias of ~2→10 K in the extra-tropical NH stratosphere. With smaller warm bias magnitudes, compared to the opposing SH winter stratosphere values, one would expect to see a smaller increases between control run and L70 analyses temperatures in the NH winter upper mesosphere. However the control run temperatures of the NH winter upper mesosphere are ~2~6 K colder than those seen in the L70 analyses, and the westward gravity wave forcing of this season has qualitatively similar magnitudes for both the control run and L70 analyses. Thus it is possible that the differences in biases seen in the NH winter upper mesosphere are primarily a consequence of inter-annual variability, which is known to be larger in the NH winter compared to the SH winter due to enhanced Rossby wave activity.

The strong influence of Rossby wave forcing on the NH winter mesosphere can be seen from Figure 8.3 (a), where we find that for the control run there is a convergence of EP flux resulting in ~30→35 ms^{-1} day^{-1} at extra-tropical latitudes of the lower mesosphere, with ~15→20 ms^{-1} day^{-1} westward forcing seen at similar latitudes in the upper mesosphere. This resolved forcing in the control run NH winter upper mesosphere is ~30% of the total forcing, i.e resolved plus small scale gravity wave forcing. Furthermore, in the upper mesosphere of the control run NH winter there is also a divergence of EP flux resulting in eastward forcing between ~20→30^{o} and at polar latitudes, resulting in the dipole structure of EP flux convergence/divergence detailed in Section 2.3.5. The existence of such a dipole structure strongly suggests that planetary waves are being generated in-situ through baroclinic or barotropic instabilities in the control run NH winter upper mesosphere. Therefore, since resolved forcing has a strong impact on the dynamical forcing of the control run NH winter mesosphere, and this forcing is known to have high inter-annual variability, it is reasonable and expected to find

temperature bias differences between the control run and L70 analyses for this region. Furthermore it is also possible that the differences in vertical resolution between the L70 and L85 models could influence the highly variable resolved forcing of the NH winter mesosphere. However as the values of resolved forcing for the L70 analyses are not available it is impossible to comment on the magnitude (or even existence) of such an influence.

From the off-line studies presented in Figure 7.6 we find that the magnitude of the gravity wave forcing for extra-tropical latitudes in the upper mesosphere of both summer hemispheres for the control run is qualitatively similar to that of the L70 analyses. With the above similarities in extra-tropical gravity wave forcing it is thus surprising that as noted above there are such large temperature differences between the control run and L70 analyses in the upper summer mesosphere of both hemispheres. As noted in Section 2.3.5 Rossby waves are primarily absorbed in the stratosphere of the summer hemisphere and generally do not propagate into the summer upper mesosphere. This is evident in the control run fields by the small magnitudes, generally less than $\sim$5 ms^{-1} day^{-1}, of EP flux convergence seen in the upper mesosphere of both summer season in Figure 8.3 (a) and (d). Thus it is unlikely that any differences in the resolved forcing of the summer upper mesosphere could alone account for the large differences in the upper mesospheric summer temperatures.

One possible reason for the colder temperatures of the control run summer upper mesosphere and those from the L70 analyses is the inter-hemispheric coupling that exists between the winter stratosphere and summer mesosphere. Following the reasoning of Section 2.3.5 it is plausible that the unconstrained stratospheric SH winter temperature and wind fields, which as noted above result in stronger gravity wave forcing of the upper mesosphere, could result in a cooling at polar latitudes of the NH summer upper mesosphere for the control run, which would not occur in the constrained L70 analyses. Since the warm bias of the control

SH winter stratosphere manifests in the spin up time from the 15^{th} of June 2008 (not shown), and as noted in Section 2.3.5 temperature responses in the polar summer mesosphere typically occur ~15-20 days after changes in the structure of the winter stratosphere, it is plausible that such a response would be captured when averaging the control run from the 1^{st} to the 14^{th} of July 2008. However it is unlikely that the above mechanism would alone explain the large temperature differences between the NH summer upper mesosphere of the control run and the L70 analyses, since as noted in Karlsson et al. [2009] the magnitude of the temperature response (cooling) in the summer upper mesosphere to an increase of ~5 ms^{-1} day^{-1} in gravity wave drag at mid-latitudes of the opposing winter hemisphere is ~2→5 K. However as such a temperature response has been shown to be linear to anomalous forcing, it is reasonable that the stronger forcing seen in the control run SH winter upper mesophere (~10→15 ms^{-1} day^{-1} compared to the L70 analyses) could result in a cooling of ~6→15 K in the opposing summer hemisphere. The above inter-hemispheric coupling also applies between the NH winter and SH summer upper mesosphere. Here relatively small differences in resolved forcing between the control run and L70 analyses, could possibly have a significant impact on SH summer upper mesospheric temperatures as the correlation between the NH winter and SH summer is typically stronger that found between the SH winter and the NH summer [Karlsson et al., 2007], due to the higher inter-annual variability of the NH winter season.

In addition to the inter-hemispheric coupling detailed above, differences between the control run summer upper mesospheric temperatures and those of the L70 analyses could occur from variations in the radiative heating/cooling of the upper mesosphere due to the different vertical resolutions of the L70 and L85 model configurations. As discussed in Section 4.6.3 the Edwards-Slingo radiation scheme does not account for the breakdown of LTE, and this is a likely cause of the latitude warm bias found in the L70 analyses approaching the model lid. As the L85 configuration has a higher vertical resolution approaching ~0.01

hPa, compared to the L70 model, it is likely that any warm bias resulting from not representing the breakdown of LTE would be reduced for the control run. However, as for the values of EP flux, the radiative heating rates for the L70 analyses are not archived, thus it is impossible to comment on the magnitude or existence of such a difference in radiative heating.

8.3 USSP Tuning Experiment

From the off-line sensitivity experiments performed in Section 7.5 we found that both increasing β and decreasing s produced an increase of the momentum forcing in both winter and summer hemispheres. Based on previous experimental results, which detail the polar temperature response to increased gravity wave forcing in the upper mesosphere [Garcia and Boville, 1994, Hamilton, 1995, Hitchman et al., 1989, Orr et al., 2010], it was estimated that a forcing increase of 10 ms^{-1} day^{-1} would provide a significant strengthening of the mesospheric meridional circulation, and hence temperature response of polar temperatures in the upper mesosphere of the winter and summer seasons. Restricting the increase of the control run gravity wave forcing to 10 ms^{-1} day^{-1} would result in total forcing magnitudes which are within previous estimates detailed in Section 7.3.3, thus this increase is physically reasonable based on the high level of uncertainty in these previous estimates.

As detailed in Section 7.5 a perturbed value of s=0.6 provides an increase of only $\sim$6$\rightarrow$8 ms^{-1} day^{-1}. Since s is not well constrained and large perturbations from the operational value are required to produce additional forcing of sufficient magnitude, the β parameter was chosen for tuning the USSP scheme to provide the additional 10 ms^{-1} day^{-1} gravity wave drag. From Figure 7.10 we find that using a perturbed value of β=0.14 provides an increase of $\sim$10 ms^{-1} day^{-1} in both winter and summer hemispheres, hence this value was chosen for the UM experiment described below which we shall refer to as the β run.

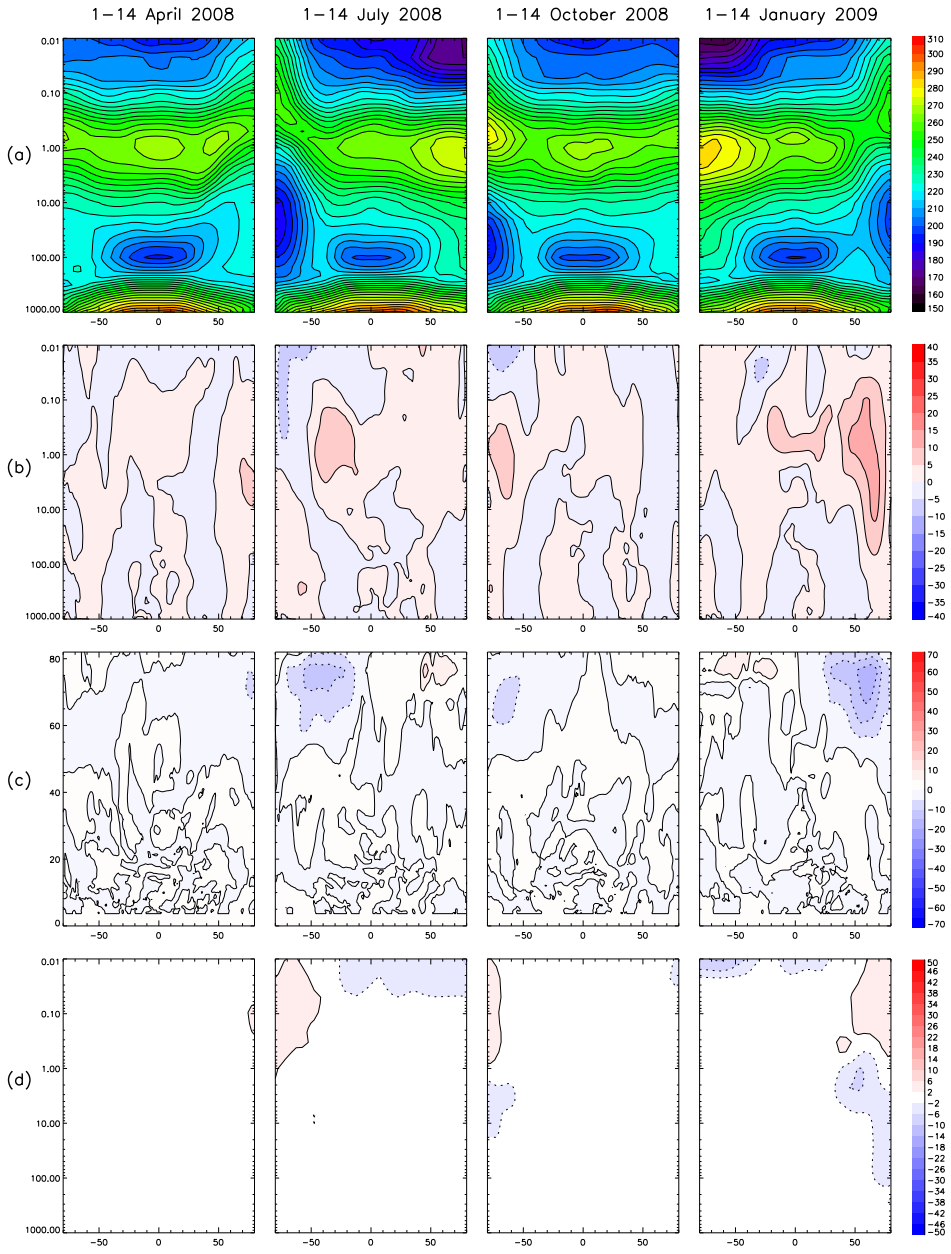

Figure 8.4: (a) Zonal mean temperature fields of the β run; (b) zonal mean wind for control minus β; (c) zonal mean gravity wave drag for control minus β and (d) zonal mean temperature error with respect to EOS MLS for control minus β for the 1^{st} to 14^{th} of April, July, October 2008 and January 2009. Contour interval is 5 K for temperature. Contour intervals are 5 ms^{-1} for zonal wind difference, 5 ms^{-1} day^{-1} for gravity wave drag difference and 4 K for temperature error difference with negative values dashed.

Figure 8.4 details the zonal mean temperature fields of the β run along with the differences, β minus control, for the zonal mean zonal wind, gravity wave forcing and zonal temperature errors averaged over the 1^{st} to the 14^{th} of April, July, October 2008 and January 2009. From Figure 8.4 (c) we find that setting β=0.14 has increased the gravity wave drag values in the upper mesosphere for all seasons. The westward drag of the winter season has increased by ~10→15 ms^{-1} day^{-1} in the SH and ~15→20 ms^{-1} day^{-1} in the NH, while in the eastward drag of summer season has increases by ~5→10 ms^{-1} day^{-1} for both hemispheres. Such increases of the winter season drag are larger than those predicted in the off-line calculations of Section 7.5. However this would be expected since the parameter β scales the value of forcing at each model level, and as noted above the initial forcing in the control run has larger magnitudes compared to the L70 analyses estimates. In the equinox seasons there is a smaller increase of drag values with magnitudes primarily under ~5 ms^{-1} day^{-1}, except for the SH of October 2008 where the westward forcing has increased by ~5→10 ms^{-1} day^{-1}.

The temperature response to this increased forcing is primarily seen at polar latitudes in the upper mesosphere approaching the model lid. Here winter polar temperatures for the β run are warmer than those of the control run, resulting in a ~2→6 K increase of the warm biases seen at the model lid and a ~2→6 K decrease of the cold biases present from ~0.4→0.05 hPa. In the summer hemisphere the polar temperatures of the β run are lower than those seen in the control run. The resulting decrease of the warm biases seen in the control run for this season is ~2→6 K for the SH and ~2→10 K for the NH. For equinox months the smaller increase of momentum forcing results in a smaller response in β temperature fields, where noticeable difference of ~2→6 K in temperature biases between the control and β runs are only seen for high polar latitudes of both hemispheres for the 1^{st} to 14^{th} of October 2008. Here a larger temperature response is seen as the SH winter jet structure persists longer into the spring season than in the NH winter, as such there is increased gravity wave forcing and hence temperature

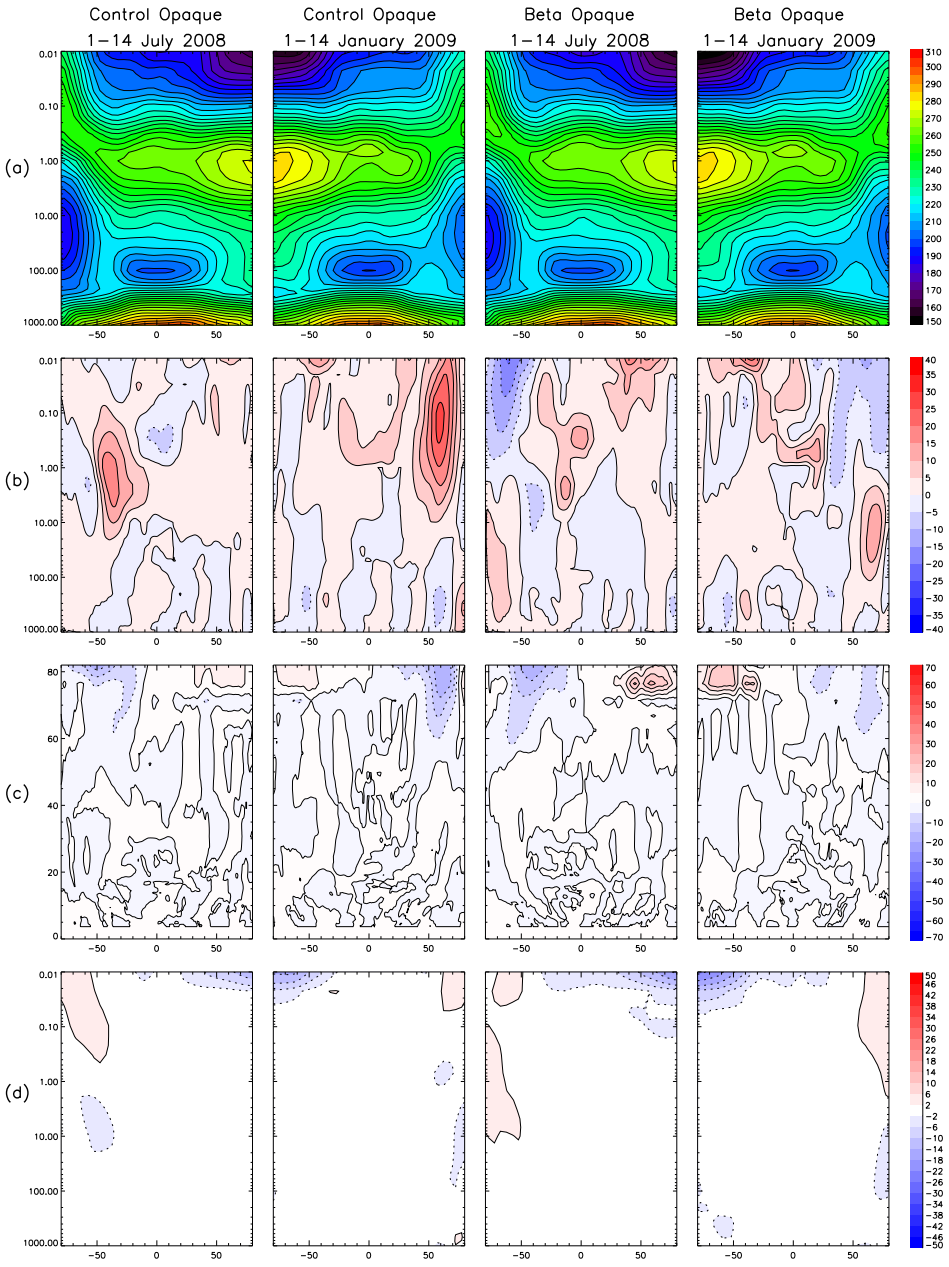

Figure 8.5: (a) Zonal mean temperature fields of the control opaque and β opaque runs; (b) zonal mean wind for control minus control opaque and control minus β opaque; (c) zonal mean gravity wave drag for control minus control opaque and control minus β opaque (d) zonal mean temperature error with respect to EOS MLS for control minus control opaque and control minus β opaque for the 1^{st} to 14^{th} of July 2008 and January 2009. Contour intervals are the same as Figure 8.4

response compared to the NH spring season. The increased winter and decreased summer polar temperatures in the upper mesosphere of the β run are consistent with a strengthened meridional circulation, as expected when applying larger magnitudes of zonal gravity wave forcing. Consistent with this result we find from Figure 8.3 (c) and (f) that the residual meridional velocities of the β run are stronger than those for the control run approaching the model lid, most noticeable in the winter seasons of both hemispheres.

8.4 Opaque Lid Condition Experiments

In addition to the β experiment described above the influence of applying the opaque lid condition to the USSP scheme was also investigated. As noted previously, approaching ~0.01 hPa the momentum flux modelled by the USSP scheme has significant (non-zero) values, therefore, through the downward control argument, applying the opaque lid condition will employ a measure of the momentum forcing above the model domain. Thus it is expected that the magnitude of the momentum forcing produced by applying the opaque lid condition will be greater than that seen when the transparent lid condition is employed. Both the impact of applying the opaque lid condition to the control run and the increased β experiment was investigated. As shown by the β experiment above, increasing the gravity wave forcing primarily results in significant impacts of the temperature and wind fields for the winter and summer hemispheres only, therefore we restrict our attention to these seasons for these experiments. We shall refer to the control and β experiments using the opaque lid condition as *control opaque* and β *opaque* respectively.

Figure 8.5 details the zonal mean temperature fields of the control opaque and β opaque experiments along with the differences, control opaque minus control and β opaque minus control, for the zonal mean zonal wind, gravity wave forcing and zonal temperature errors averaged over the 1^{st} to the 14^{th} of July 2008 and January 2009. From Figure 8.5 we find applying the opaque lid condition

to the control run experiment results in an increase of summer eastward forcing by ~5→10 ms^{-1} day^{-1} for both hemispheres approach the model lid, while the winter westward forcing increases by ~10→15 ms^{-1} day^{-1} in the SH and ~15→20 ms^{-1} day^{-1} in the NH. Note that the increase of gravity wave forcing seen by imposing the opaque lid condition has similar magnitudes to that seen in the previous β experiment, however the increased forcing typically occurs at higher altitudes closer to the model lid. Compared to the control run the β opaque experiment results in an increase of summer eastward forcing by ~15→20 ms^{-1} day^{-1} for both hemispheres approach the model lid, while the winter westward forcing increases by ~15→20 ms^{-1} day^{-1} in the SH and ~5→10 ms^{-1} day^{-1} in the NH. As expected, applying the opaque lid condition in the β opaque run increases the forcing compared to the β run.

As the increase of momentum forcing compared to the control run is similar for both the β and control opaque experiments, similar responses are also seen in the control opaque temperature errors. However the control opaque run results in colder temperatures for polar latitudes in the upper mesosphere of both summer hemispheres when compared to the β experiment, where the resulting decrease of the warm biases seen in the control run for this season is ~6→10 K for the NH and ~6→14 K for the SH. The fact that employing the opaque lid condition results in a larger decrease of the warm polar mesospheric temperature biases when compared to the β experiment highlights the importance and impact of conserving momentum throughout the model domain. The strong increase of forcing in the β opaque experiment produces the largest difference in temperature errors compared to the control run. Here the resulting decrease of the warm biases seen in the control run summer polar upper mesosphere is ~10→18 K for the NH and ~14→22 K for the SH. Here the vertical extent of the change in polar temperature is larger for the β opaque experiment compared to the control opaque and β runs, where noticeable changes in bias values are seen between ~0.05→0.04 hPa. In the winter hemisphere the β opaque experiment has similar

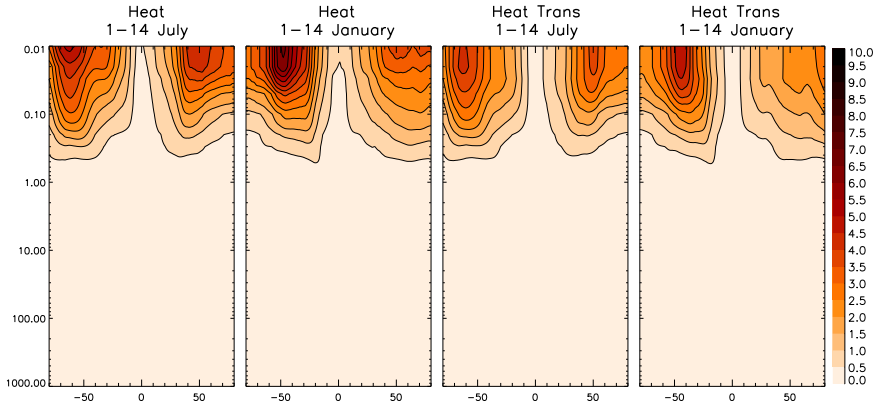

Figure 8.6: Zonal mean gravity wave heating for the Heat and Heat Trans experiments averaged from the 1^{st} to 14^{th} of July 2008 and January 2009. Contour interval is 0.5 K day^{-1}.

impacts on the control run biases to those seen in the previous two experiments however with larger vertical extent, most noticeable in the SH where warmer temperatures of ~2→6 K are seen in the upper stratosphere from ~10→1.0 hPa.

Again the diagnostic fields of Figure 8.3 are consistent with the above change in temperature biases. Here the largest change in temperature biases, seen in the summer polar upper mesosphere of both hemispheres for the control opaque and β opaque runs, coincide with regions of the largest increase in upward values of $\bar{w}^*$, while smaller changes in the values of downward $\bar{w}^*$ are seen at polar latitudes of the winter upper mesosphere for both hemispheres.

8.5 Gravity Wave Heating Rate Experiments

The final experiments using the Unified Model involved applying the direct heating rates due to gravity wave breaking discussed in Section 6.2. Here the β opaque experiment, which produced the largest decrease of temperature biases in the polar summer mesosphere of all previous experiments, was run with the addition of the gravity wave heating rates. When using the opaque lid condition

for calculation of the gravity wave momentum forcing and the direct heating rates, i.e. the momentum and energy flux profiles used to calculate the direct heating using Equation 6.41 have zero values at the model lid, the experiment is referred to as the *Heat* run. As for the previous opaque lid condition experiments we limit our attention to the winter and summer seasons of both hemispheres.

8.5.1 Opaque Lid Direct Heating Rates

Figure 8.6 details the zonal mean gravity wave heating rates of the Heat experiment averaged over the 1^{st} to 14^{th} for July 2008 and January 2009. Figure 8.7 details the zonal mean temperature, zonal wind, zonal gravity wave drag and zonal temperature error for both the β opaque and Heat experiments averaged from 1^{st} to 14^{th} for July 2008 and January 2009. Also shown are the differences, β opaque minus Heat fields, for the zonal mean zonal wind, gravity wave drag and zonal temperature bias.

From Figure 8.6 we find that the gravity wave heating rates of the Heat run have maximum magnitudes in the extra-tropics of the upper mesosphere with winter season values of ~4.5→5.0 K day^{-1} in the SH and ~3.5→4.0 K day^{-1} in the NH, while summer seasons have magnitudes of ~5.5→6.0 K day^{-1} in the SH and ~4.0→4.5 K day^{-1} in the NH. In the lower mesosphere from ~0.3~0.1 hPa heating rates of the SH winter and summer are ~1.0→2.0 K day^{-1}, while in the NH winter and summer smaller magnitudes of ~0.5→1.0 K day^{-1} are for this altitude range. The above heating rates of the Heat run are qualitatively consistent with those seen in the off-line study of Section 7.4. However the magnitudes of heating in both seasons of the SH lower mesosphere is ~1.0 K day^{-1} larger than those in the previous studies of Becker [2004], Fomichev et al. [2002]. Such discrepancy is likely due to the strength of the gravity wave forcing in the lower mesosphere. Here the gravity wave forcing of the SH winter lower mesosphere for the β opaque experiment has magnitudes of ~45~50 ms^{-1} day^{-1}. Such values are noticeably stronger, or at the upper limit, of previous estimates

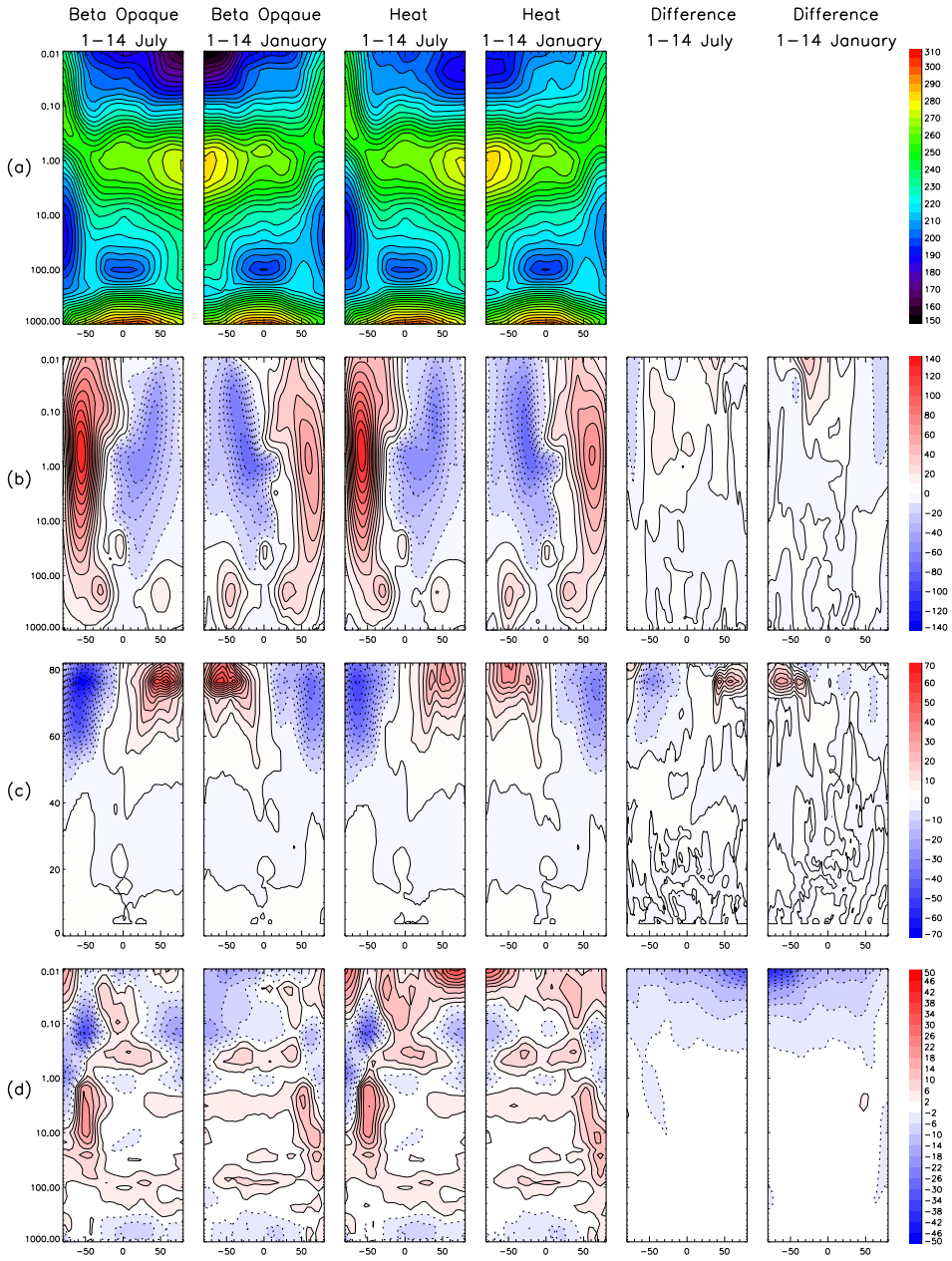

Figure 8.7: (a) Zonal mean temperature; (b) zonal mean wind; (c) zonal mean gravity wave drag and (d) zonal mean temperature error with respect to EOS MLS for β opaque and Heat experiments averaged from the 1^{st} to 14^{th} of July 2008 and January 2009. Differences, β opaque minus Heat, shown for (b), (c) and (d) fields. Contour interval is 0.5 K for gravity wave heating and 10 ms^{-1} for wind fields, all other contour intervals are identical to Figure 8.4.

from the studies detailed in Section 7.3.3. Such strong gravity wave forcing in the lower mesosphere of this season for the β opaque experiment is a consequence of further increasing the forcing seen in the control run, which as detailed above already has large magnitudes due to the influence of stratospheric temperature biases not being constrained by assimilation.

As the radiative timescale in the upper mesosphere is ~4→5 days, it is expected that the above gravity wave heating rates would likely result in a temperature increase at extra-tropical latitudes of the upper mesosphere of ~20→25 K in the SH winter, ~15→20 K in the NH winter, ~25→30 K in the SH summer and ~15→20 K in the NH summer hemisphere. From Figure 8.7 (a) and (d) we find that the difference between β opaque and Heat temperature fields is qualitatively similar to these expected values. However temperature differences are larger than the above estimated values in the summer hemispheres and lower than the estimated values in the winter hemispheres. Furthermore, the temperature differences of all seasons show a distinctly different structure and distribution to the direct heating rates of Figure 8.6. Here maximum difference in the temperature biases of the Heat run compared to those of the β opaque run for all season are located at polar regions in the upper mesosphere, i.e. warmer by ~6→14 K in the both winter hemispheres, and warmer by ~34→38 K in the SH summer and ~26→30 K in the NH summer hemispheres. While maximum direct heating rates typically occur between ~40→50^o latitude in the summer hemispheres and between ~55→70^o latitude in the winter hemispheres. This non-local response, with temperature increases larger in the summer and smaller in the winter than the expected values above, would strongly suggest that the introduction of the direct heating rates has weakened the mesospheric meridional circulation compared to the β opaque run, resulting in decreased adiabatic warming (cooling) for the polar winter (summer) latitudes.

The reduction of the Heat run meridional circulation approaching the model

lid is also evident by the $\bar{v}^*$ and $\bar{w}^*$ distributions of Figure 8.3. Here largest differences in the meridional circulations of the Heat and β opaque runs occur for seasons with the largest heating rates, i.e both summer hemispheres. Here the magnitudes of $\bar{v}^*$ for the Heat run are $\sim$2$\rightarrow$4 ms^{-1} smaller compared to β opaque values for both summer hemispheres in the extra-tropics approaching the model lid. From Figure 8.7 (c) we find that there is a reduction of the gravity wave forcing in both hemispheres of the Heat run summer upper mesosphere when compared to β opaque values, here the maximum differences in forcing is $\sim$25$\rightarrow$30 ms^{-1} day^{-1} in SH summer and $\sim$20$\rightarrow$25 ms^{-1} day^{-1} in NH summer. Since for the summer upper mesosphere the dynamical forcing is known to be dominated by gravity waves, it is therefore most likely that here the weakening of the meridional circulation is due to the reduced gravity wave forcing caused by the introduction of the direct heating rates. This argument is supported by the evidence of Figures 8.6 and 8.7 which show that in the upper summer mesosphere the largest decreases in gravity wave forcing occur for the hemisphere which has the largest heating rates, i.e in the SH.

For the SH winter season the $\bar{v}^*$ and $\bar{w}^*$ distributions of Figure 8.3 also show a reduction of the Heat run meridional circulation compared to the β opaque experiment, consistent with the lower observed difference in upper mesospheric temperatures compared to the expected values noted above. As for both summer hemispheres this weakening of the circulation in the SH winter upper mesosphere is most likely due to the decrease in gravity wave forcing, where from Figure 8.7 (c) we find that gravity wave forcing in the SH winter upper mesosphere of the Heat run is $\sim$25$\rightarrow$30 ms^{-1} day^{-1} weaker than in the β opaque run.

However in the NH winter season the meridional circulation of the Heat run in the polar upper mesosphere is marginally stronger than that seen in the β opaque experiment, even though there is a reduction, albeit noticeably smaller than that for all other seasons ($\sim$5$\rightarrow$10 ms^{-1} day^{-1}), in the gravity wave forcing

of this region. This strengthening of the Heat run meridional circulation results in the observed temperature increases of the NH winter upper mesosphere being marginally closer (warmer) to the expected increases due to the inclusion of direct heating rates, when compared to the observed temperature increases seen in the SH winter upper mesosphere. As shown in Section 8.2 resolved forcing contributes significantly to the dynamical forcing of the NH winter mesosphere, and from Figure 8.7 (a) we find that the resolved eastward forcing in the NH upper mesosphere of the β opaque run, apparent at latitudes poleward of ~65^o, has reduced by ~5→10 ms^{-1} day^{-1} for the Heat experiment. Thus, it is plausible that this reduction maybe responsible for the strengthening of the meridional circulation in the NH winter upper mesosphere of the Heat run.

From Figure 8.7 (b) we find that for the Heat run the upper mesospheric zonal mean winds have increased/decreased magnitudes either side of the location of maximum heating rate compared to the β opaque experiment. These differences can be seen as dipole structures in the plots of difference between wind fields of the Heat and β Opaque runs. These dipole structures have strongest magnitudes where the meridional gradient of direct heating rates are also strongest, suggesting that the changes in wind distributions are attributed to the thermal wind balance response from changes to the meridional temperature gradient, induced by a combination of the response to the additional heating rates and the subsequent changes in dynamical adiabatic heating from the reduced meridional circulation.

From Figure 8.7 (d) we find that the application of direct heating rates in the Heat run has a detrimental impact on the temperature biases of the β opaque run in the upper mesosphere. Here the introduction of the direct heating rates, and the resulting weakening of the meridional circulation, results in warm biases throughout the upper mesosphere which at polar latitudes have larger magnitudes than biases seen at similar regions of the β opaque run. To address this

issue of excessive warm temperatures in the upper mesosphere of the Heat run the Heat Trans experiment was designed, full details are given below.

8.5.2 Transparent Lid Direct Heating Rates

As noted above the application of direct heating rates in the Heat run result in upper mesospheric temperatures which are too warm. These heating rates are calculated using the opaque lid condition momentum and energy flux profiles. Physically this formulation represents depositing all of the direct heating from gravity waves which would propagate and break above the model lid into the model domain. While this is desirable for gravity wave momentum forcing under the downward control principle, it is most likely to produce direct heating rates which are too large. Thus a second heating rate experiment was designed to address this issue. In this experiment, which we shall refer to *HeatTrans*, the momentum and energy flux profiles used to calculate the direct heating using Equation 6.41 have non-zero values at the model lid, i.e. the transparent lid condition momentum and energy flux profiles. Physically the Heat Trans formulation allows the direct heating rates to be calculated without including the contribution of gravity waves which would otherwise propagate and break above the model lid. It should be noted that while transparent profile values are used in calculating the direct heating rates of the Heat Trans run, the gravity wave momentum forcing is still obtained using the opaque lid condition. As for the Heat experiment we limit our attention to the winter and summer seasons of both hemispheres.

From Figure 8.6 we find that the gravity wave heating rates of the Heat Trans run also have maximum magnitudes in the extra-tropics of the upper mesosphere, however with smaller magnitudes than those of the Heat run, as expected from the physical argument presented above. Here the maximum magnitudes, and reductions compared to the Heat run shown in parentheses, of the Heat Trans direct heating rates are ~3.5→4.0 K day^{-1} (~0.5→1.0 K day^{-1}) in the SH winter and

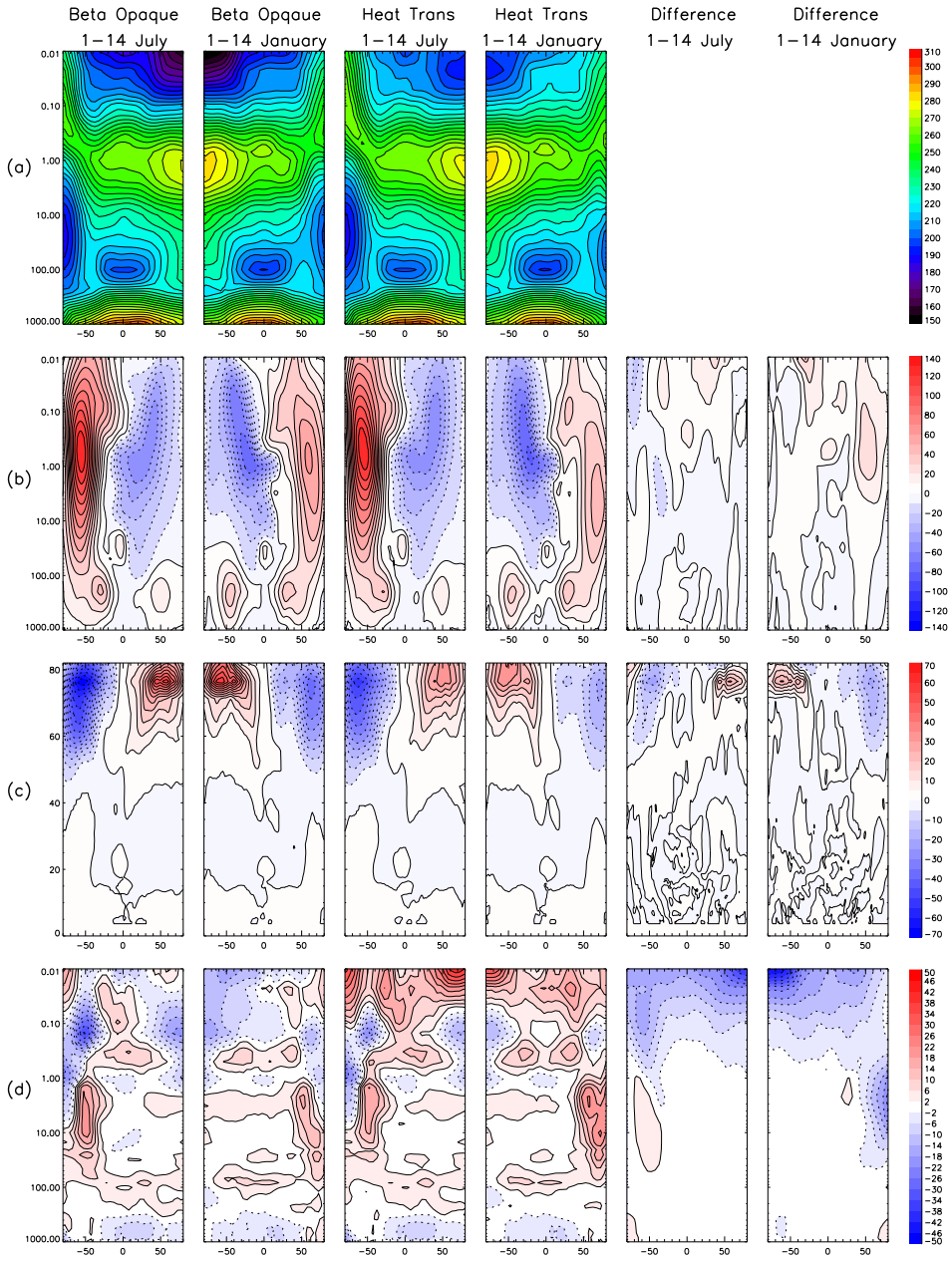

Figure 8.8: (a) Zonal mean temperature; (b) zonal mean wind; (c) zonal mean gravity wave drag and (d) zonal mean temperature error with respect to EOS MLS for β opaque and Heat Trans experiments averaged from the 1^{st} to 14^{th} of July 2008 and January 2009. Differences, β opaque minus Heat Trans, shown for (b), (c) and (d) fields. Contour interval is 0.5 K for gravity wave heating and 10 ms^{-1} for wind fields, all other contour intervals are identical to Figure 8.4.

~2.0→2.5 K day^{-1} (~1.0→1.5 K day^{-1}) in the NH winter, while for the summer season they are ~3.5→4.5 K day^{-1} (~1.0→2.0 K day^{-1}) in the SH and ~3.0→3.5 K day^{-1} (~0.5→1.0 K day^{-1}) in the NH. Again, based on radiative timescale of the upper mesosphere, it is expected that the above heating rates would likely result in a temperature increase at extra-tropical latitudes of the upper mesosphere of ~15→20 K in the SH winter, ~10→15 K in the NH winter, ~20→25 K in the SH summer and ~15→20 K in the NH summer hemisphere of the Heat Trans experiment.

Figure 8.8 details the zonal mean temperature, zonal wind, zonal gravity wave drag and zonal temperature error for both the β opaque and Heat experiments averaged from 1^{st} to 14^{th} for July 2008 and January 2009. Also shown are the differences, β opaque minus Heat Trans fields, for the zonal mean zonal wind, gravity wave drag and zonal temperature bias. As for the Heat run the maximum temperature response for the Heat Trans experiment, compared to the β opaque run, are seen at polar latitudes of the upper mesosphere with larger increases in temperature than the expected values noted above seen in the summer and smaller increases in winter. However, while the direct heating rates of the Heat Trans run are lower than those of the Heat experiment, a comparison of Figures 8.7 (d) and 8.8 (d) reveals that temperatures in the polar upper mesosphere of the Heat Trans SH winter, SH summer and NH summer seasons have increased compared to the Heat experiment. Here difference magnitudes between the Heat Trans and Heat runs are ~4 K in the SH and NH summer seasons, and ~8 K in the SH winter season.

Temperature differences between the Heat and Heat Trans experiments would suggest that there is a strengthening of the Heat Trans upper mesospheric mean meridional circulation in the SH winter (increased adiabatic warming) and a weakening in the SH summer (decreased adiabatic cooling), as confirmed by the $\bar{v}^*$ and $\bar{w}^*$ distributions in Figure 8.3. Here the weakening of the upper meso-

spheric SH summer circulation in the Heat Trans experiment compared to the Heat run is most likely associated with the increases (~10→15 ms^{-1} day^{-1}) in resolved eastward forcing, as detailed in Figure 8.3 (a), since for this season the eastward gravity wave forcing of both seasons is qualitatively similar. The strengthening of the upper mesospheric SH winter circulation in the Heat Trans experiment compared to the Heat run is most likely associated with the ~5→10 ms^{-1} day^{-1} increase in westward forcing of this season, as the resolved forcing has qualitatively similar magnitudes.

The largest differences between the Heat and Heat Trans fields occur in the NH winter stratosphere, where Heat Trans temperatures undergo a rapid increase of ~25→30 K from the ~10^{th} to the ~14^{th} of July 2008 (not shown) in contrast to the Heat experiment where temperatures remain approximately constant throughout this period. Thus it appears that the Heat Trans experiment has prematurely simulated the SSW of the NH winter, which actually occurs in the later half of January 2009. Associated with the above stratospheric warming, compared to the Heat run the Heat Trans experiment has increased (decreased) EP-flux convergence (divergence) in the NH winter upper stratosphere and lower mesosphere, resulting in reduced eastward zonal winds for this region as found when comparing Figures 8.7 (c) and 8.8 (c). The reduction of eastward winds in the Heat Trans NH winter upper stratosphere and lower mesosphere compared to the Heat run is the most likely cause of the ~5→10 ms^{-1} day^{-1} decrease in gravity wave forcing seen in the upper mesosphere.

The results of the Heat Trans experiments highlight the sensitivity of upper mesospheric temperatures to changes in the meridional circulation, via changes in both small scale gravity wave and resolved forcing, induced from the addition of direct gravity wave heating rates. This is demonstrated by the fact that even though the direct heating rates of the Heat Trans experiment are less than those of the Heat run, changes in the meridional circulation primarily result

in an increased non-local temperature response at polar latitudes of the upper mesosphere. Thus as for the Heat experiment, the application of direct heating rates in the Heat Trans run has a detrimental impact on the temperature biases of the β opaque run in the upper mesosphere.

8.5.3 Statistical Significance of Experimental Results

To ascertain the statistical significance of the above experimental results, i.e. to what extent experimental changes may noticeably have an impact on future operational analyses, it is necessary to compare the experimental responses against the natural internal variability of the middle atmosphere. Figure 8.9 details the standard deviation of middle atmospheric temperatures for six years of EOS MLS temperature data averaged over six years from 2005 to 2010 for the first 15 days of January, April, July and October.

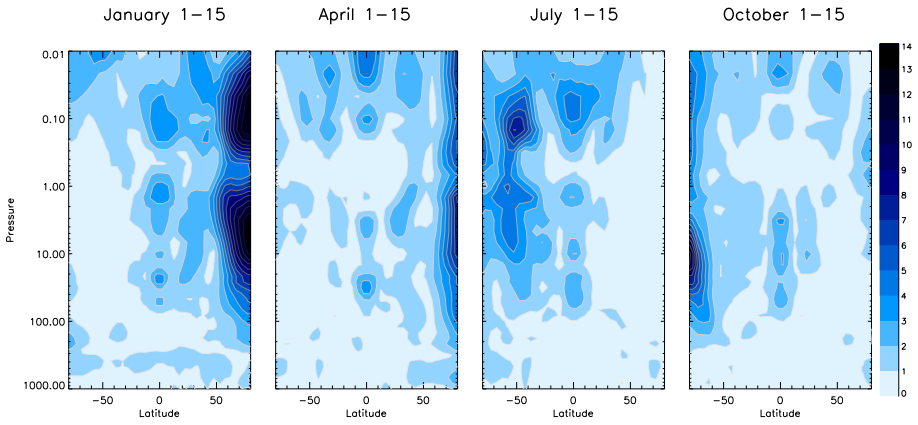

Figure 8.9: The natural inter-annual variability (standard deviation) of middle atmospheric temperatures diagnosed from EOS MLS data over six years from 2005-2010 for the first 15 days of January, April, July and October. Contour interval is 1 K.

From Figure 8.9 we find that the largest standard deviation of middle atmospheric temperatures occurs in the NH winter polar regions of the middle stratosphere and mesosphere. Here the large variability is most likely associated with the large variability of Rossby wave forcing in the stratosphere during this season

and it's subsequent impact on filtering of vertical propagating gravity waves, as discussed in Section 2.3.5. The standard deviation of temperatures in the NH polar winter upper mesosphere is ~6→10 K. Such variability in this region is comparable in magnitude to the temperature responses seen for all experiments detailed in Table VIII. Therefore, such results are not very significant, in a statistical sense, as the magnitude of the experimental responses are naturally seen in the atmospheric due to it's inherent variability. Similarly the temperature responses in the upper mesosphere for all solstice seasons of the β experiment and the SH winter of the control and β opaque experiments have magnitudes close to the natural variability, where temperature standard deviations are primarily below ~2→4 K. The opposite is true for the remainder of the experiment results, where the temperature responses in the polar regions of the upper mesosphere are noticeably larger than the natural variability, placing a higher degree of statistical significance in these results.

8.5.4 Impact of Turbulent Eddy Diffusion

It is clear from both the Heat and Heat Trans experiments that employing direct heating rates results in excessive temperatures in the upper mesosphere. It is plausible that such inaccuracies could partly be caused by such heating rates being too excessive, where as noted above the direct heating rates of the Heat and Heat Trans experiments are marginally larger than previous estimates. It is plausible that such excessive heating rates will result from the theoretical approach applied, where as noted in Section 6.2.2 all of the energy removed from the gravity wave field is available for dissipation to heat, while in reality a proportion of this energy will go into increasing the potential energy of the atmosphere [McIntyre, 1989]. However it is also plausible, and most likely, that the excessive temperatures seen in the upper mesosphere are attributed to the neglect of additional physical processes, i.e the impact of turbulent eddy diffusion.

As noted in Section 6.2.1, along with the dissipative heating associated with the turbulence generated from gravity wave breaking, there exists small scale mixing and transport of heat, mass and momentum commonly known as eddy diffusion. Eddy diffusion of momentum (Reynolds stress) is related to the eddy thermal diffusion (heat flux) via the Prandtl number, which is defined as the momentum eddy diffusion coefficient dived by the thermal eddy diffusion coefficient. Currently the Prandtl number associated with gravity wave breaking is poorly defined, however previous studies suggest that it is highly variable and in the mesosphere has a value greater than unity [Huang and Smith, 1991, Strobel et al., 1985]. The heat flux associated with the turbulence generated from gravity wave breaking can act to both cool and heat the atmosphere, depending on the background temperature and the vertical structure of the eddy heat diffusion coefficient [Huang and Smith, 1991].

The majority of previous modelling studies that have included direct heating rates form gravity wave breaking generally also include the effects of eddy diffusion, both thermal and momentum. The study of Fomichev et al. [2002] included the direct heating rates and eddy diffusion coefficients calculated from a corrected version of the Hines Doppler spread parameterization (DSP) gravity wave scheme [Hines, 1999]. The results of this study show that thermal eddy diffusion acts to heat (under ~0.5 K day^{-1}) the atmosphere in the extra-tropics of the lower mesosphere, while cooling (~1→5 K day^{-1}) occurs in the extra-tropics of the upper mesosphere. Here the net effect of eddy diffusion and direct heating results in heating and cooling of ~0.5→1.0 K day^{-1} respectively in the summer and winter seasons approaching 0.01 hPa. However as noted in Fomichev et al. [2002] a Prandtl number of 3 is required to produce cooling rates from eddy heat diffusion which result in realistic temperatures in the upper mesosphere.

While the above study of Fomichev et al. [2002] applied direct heating rates and eddy heat diffusion, the earlier study of Huang and Smith [1991] using the

Lindzen [1981] gravity wave scheme also included an additional thermal term involving the wave heat flux convergence, associated with the limitation of the gravity wave vertical amplitudes when undergoing breaking. The contribution from this wave heat flux convergence is always negative (cooling) and as shown by Schoeberl et al. [1983] has magnitudes which are approximately equal to half the eddy heat diffusion. Under solstice conditions Figure 9 of Huang and Smith [1991] shows that eddy heat diffusion produces heating of ~2 K day^{-1} above the stratopause and cooling of ~10→15 K day^{-1} approaching ~0.01 hPa. The cooling from wave heat flux convergence has maximum magnitudes of ~6→8 K day^{-1} approaching ~0.05 hPa. The net thermal flux from combining the eddy heat diffusion and wave heat flux convergence is always negative (cooling) [Coy and Fritts, 1988], and here dominates over the direct heating rates from turbulent dissipation which have magnitudes of ~2→4 K day^{-1} approaching ~0.01 hPa. However this dominant net cooling reverses to net heating for large Prandtl numbers, where both the eddy heat diffusion and wave heat flux convergence terms tend to zero. The study of Huang and Smith [1991] predicts the reversal of net cooling to net heating occurs at a Prandtl number of 10, while the previous study of Strobel et al. [1985] predicts a lower value of 6. It should be noted that the magnitudes of the above heating and cooling rates in both the Fomichev et al. [2002] and Huang and Smith [1991] studies have large uncertainties, since as in all gravity wave schemes they involve several free parameters which to some extent are arbitrarily set.

As for the Heat and Heat Trans experiments described above, including the direct thermal effects of gravity wave breaking in the study of Huang and Smith [1991] impacts the strength of the mean meridional circulation in the upper mesosphere. Here the net thermal cooling acts to increase the meridional velocities in the upper mesosphere. This result is consistent, however with reversed sign, with the results of the Heat and Heat Trans experiments where the net heating primarily acts to reduce the residual circulation. As noted in Huang

and Smith [1991] the impact on the mean meridional circulation is highly dependent on the structure of radiative and wave induced heating/cooling. Thus different input values of free parameters in the gravity wave scheme can lead to a strengthening or weakening of the circulation. Furthermore, they noted that the direct thermal impact of wave breaking on the circulation is negligible when Prandtl numbers are larger than 10. The results of Huang and Smith [1991] highlight the importance of understanding gravity wave processes in understanding the mesospheric circulation, where depending on the choice of Prandtl number (which is still open to debate) gravity wave breaking maybe the dominant contribution to the thermal as well as the momentum budget. Thus it is apparent when employing direct heating associated with gravity wave breaking, to accurately represent the upper mesospheric circulation it is necessary (dependent on Prandtl number) to include the impacts of eddy diffusion.

8.6 Summary

This chapter has detailed the results from numerous UM experiments designed to improve the operational L70 analyses, specifically the temperature biases of the extra-tropical upper mesosphere associated with an insufficient mean meridional circulation, which is primarily driven by small scale gravity wave forcing. The 85 level configuration of the UM was chosen as the increased vertical resolution has been shown to reduce stratospheric biases for free running integrations over ~10→15 days. The results from all experiments were averaged over a two week period following one month of spin up integration.

Following the spin up time of one month the control run temperature biases had evolved from those seen in the initial ECMWF conditions, and for the upper stratosphere and lower mesosphere were found to be qualitatively similar to those of the L70 analyses for all seasons. This agreement would place additional confidence in the hypothesis that biases in these regions are due to the radiative deficiencies detailed in Chapter 4. However under solstice conditions in the

Experiment	Gravity Wave Forcing (ms^{-1} day^{-1})				Temperature Bias (K)			
	SH (W)	SH (S)	NH (W)	NH (S)	SH (W)	SH (S)	NH (W)	NH (S)
Control	-60→-65	30→35	-25→-30	30→35	10→14	2→2	6→6	18→22
β	-70→-80	35→45	-40→-50	35→45	14→18	2→6	10→2	14→18
Control Opaque	-70→-80	35→40	-40→-50	35→40	14→18	8→12	10→2	10→14
β Opaque	-75→-85	45→55	-30→40	45→50	14→18	16→20	10→2	4→8
Heat	-50→-55	25→35	-30→-35	25→35	16→24	24→28	6→10	28→36
Heat Trans	-55→-65	25→35	-20→25	25→35	24→28	28→32	10→14	28→36

Table VIII: Summary of all UM experiments, see text for details.

extra-tropical upper mesosphere there were noticeable differences between the control run and L70 analyses temperature biases. Here temperature biases in control run NH summer, NH winter, SH summer and SH winter, were respectively ~40 K colder, ~6 K colder, ~50 K colder and ~25 K warmer than those seen in the L70 analyses. In the SH winter the differences in temperature biases are most likely associated with an increase of the seasonal westward gravity wave forcing due to unconstrained stratospheric fields. In the NH winter the differences in temperature biases are most likely due to the large inter-annual variability of this season, due to strong Rossby wave activity. Here it is possible that the difference in vertical resolutions between the L70 and L85 configurations may have an impact. For both summer hemispheres the differences in temperature biases are most likely attributed to the inter-hemispheric coupling between the upper mesospheric winter and summer hemispheres, where increased forc-

ing seen in the control run winter hemisphere could possibly result in cooling in the opposing summer hemisphere, and potential differences in radiative errors approaching the model lids due to differences in vertical resolution.

Table VIII details the maximum values of gravity wave forcing and temperature biases seen in the extra-tropical upper mesosphere for all experiments discussed above. The results are shown for both the winter and summer seasons of each hemisphere. Here we limit are attention to the solstice months since they show the largest differences in gravity wave forcing and temperature response.

As expected by the off-line sensitivity studies in Section 7.5, an increase of the free parameter β in the USSP scheme (β run) results in increased gravity wave forcing and hence increased (decreased) adiabatic warming (cooling) in the winter (summer) polar upper mesosphere, through a strengthened mean meridional circulation. For the control opaque experiment the increases in gravity wave forcing are similar to those of the β run, however larger increases occur closer to the model lid. This results in an increased temperature response, compared to the β run, most noticeable in the summer hemisphere. This increased response highlights the importance and impact of applying the opaque lid condition and conserving momentum throughout the model domain. The largest increase in gravity wave forcing is seen in the β opaque experiment, resulting in the largest temperature response. This response is most prominent in the summer hemispheres where decreases of ~14→22 K occur in the SH and ~10→18 K in the NH. In should be noted that for all experiments using an opaque lid there were no issues of numerical instability due to depositing the remaining momentum flux at the model lid. In general the β, control opaque and β opaque experiments have a positive impact on the warm temperature biases in both summer hemispheres of the control run, thus it is expected a similar response would also be seen in the summer hemispheres of the operational L70 analyses. In the winter hemisphere the above experiments have a small negative impact (increase) of the

warm biases seen in the control, however a positive impact would be expected in the L70 analyses which unlike the control run show cold biases for this season in the polar upper mesosphere.

While the above experiments show a positive impact, the results of the Heat and Heat Trans experiment show a large negative impact. Here the turbulent dissipation direct heating rates have magnitudes which are comparable to previous estimates. However their introduction results in excessive temperatures at polar latitudes of all seasons in the upper mesosphere compared to the β opaque run. While it is expected that this increase of temperature would have a small positive effect in both polar winter hemispheres of the L70 analyses, there would also be a large negative impact in both summer polar mesospheres which as detailed in Section 4.6.3 already contains significant warm biases. This negative impact in the polar upper mesosphere is also strengthened by the alteration in the strength of the mean meridional circulation, seen when the direct heat rates are applied. Such an alteration is consistent with previous studies and highlights that cooling in the upper mesosphere from eddy diffusion should be accounted for to obtain an accurate representation of the mesospheric circulation and temperature structure.

Chapter 9

Conclusions

9.1 Conclusions

This thesis has been concerned with the evaluation and improvement of biases in the daily analyses of the stratospheric assimilated data set. These analyses are created using the global middle atmospheric configuration of the Met Office UM. The initial evaluation of temperature biases was performed by direct validation against independent satellite temperature profiles from the UARS MLS, EOS MLS and SABER experiments. Due to insufficient independent wind observations in the upper stratosphere and mesosphere, biases in the analyses wind fields were evaluated indirectly via comparison with the SPARC climatology. Based on the results of the temperature validation, individual systematic biases were attributed to specific aspects of the operational system. The first daily temperature and wind analyses from the stratospheric assimilated data set were produced in October 1991, since then numerous changes have been made to the operational system. Recently the most significant change was the vertical extension of the UM from 0.1 hPa to 0.01 hPa, in November 2009, resulting in an increase of the number of model vertical levels. The temperature validation results were separated into three distinct groups decided by the importance of changes to the operational system and satellite data coverage. Each group was classified via the number of vertical levels used in the UM for the selected peri-

ods; L42 for analyses from January 1992 to December 1994, L50 for analyses from January 2005 to October 2009 and L70 for analyses November 2009 to September 2010. The validation results for all periods were presented as monthly zonal means.

9.1.1 L50 Validation

The validation results for the L50 analyses against EOS MLS and SABER data showed that above ~1.0 hPa, the difference between the validation results and the known biases of each data set are generally smaller than the analyses biases. Thus we can confidently state that the analyses biases are predominantly due to inaccuracies of the analyses fields, and not biases in the observational data. Note that the differences between the EOS MLS and SABER temperatures were consistent with the previous results of Schwartz et al. [2007]. For the L50 analyses discussion of the temperature biases is further separated into periods pre and post March 2006. This comparison was made due to the large number of changes made to the system at this date including an increase of horizontal resolution, introduction of 4D-var assimilation and correction to the bugged AIRS assimilation code.

The largest biases seen in the L50 analyses occur at the upper most levels above ~0.3 hPa. Here there is a large cold bias of ~30→40 K seen in the winter polar regions of both hemispheres for months post March 2006, while slightly larger magnitudes are seen during winter months pre March 2006. Such large cold biases are not apparent at polar latitudes of the equinox seasons, where cold biases have magnitudes of ~15 K. The seasonal evolution of the above cold biases would strongly suggest that the mean meridional circulation, primarily driven by small scale gravity wave forcing for this region, is underestimated in the L50 analyses resulting in insufficient adiabatic warming at the winter pole.

With large cold biases in the polar winter lower mesosphere, one would ex-

pect by mass conservation to see larger warm biases in the opposing summer polar regions, however, as detailed in Section 4.3.1 there are small magnitude cold biases in the polar summer lower mesosphere of the L50 analyses. Thus, suggesting that the biases in this region result from misrepresenting the mean meridional circulation combined with an additional source of error. Based on the previous studies of Jackson et al. [2008], Mathison et al. [2007] this additional source of error is most likely attributed to the currently operational ozone climatology of Li and Shine [1995]. Here the ozone climatology has been shown to produce insufficient radiative heating in the lower mesosphere, consistent with the latitude independent cold bias signal seen above ~0.3 hPa for all months of the L50 analyses.

Accounting for the cold biases arising from the ozone climatology would hypothetically result in a warm bias in the summer polar upper mesosphere, thus is consistent with the evidence presented for insufficient gravity wave forcing in the L50 analyses. The deficiencies in gravity wave forcing of the L50 model are likely to be intensified by using the transparent lid condition in the USSP gravity scheme, where by allowing the remaining momentum flux to propagate out of the model domain where through the downward control principle the mean meridional circulation at the model lid has no knowledge of the wave forcing above. Comparison with the SPARC climatology reveals that the zonal mean zonal winds of the L50 analyses under solstice conditions are too strong in the lower mesosphere and have insufficient vertical shear when approaching the model lid. This result is consistent with the temperature validation results, where the excessive wind magnitudes would suggest gravity wave forcing, which acts to reduce the zonal flow at these altitudes, is too weak in the L50 analyses. Similar arguments and results for the solstice months also apply for temperature and wind biases seen in the polar latitudes of L50 equinox months above ~0.3 hPa. For tropical latitudes of the L50 analyses above ~0.3 hPa there is also a consistent cold bias, however magnitudes of ~22→30 K seen for months pre March 2006

are considerably larger than magnitudes of ~14 K seen for months post March 2006. The decrease of this tropical bias has been attributed to the correction of the bugged AIRS assimilation code.

The cold biases of the L50 analyses are accompanied by a warm bias directly below from ~2.0→0.3 hPa over various latitudes. This warm bias signal is stronger for months post March 2006 with maximum magnitudes of ~14 K and a latitude independent bias signal of ~2→10 K. The latitude independent nature of this warm bias would strongly suggest that it is associated with inaccuracies of the radiation scheme, and is most likely associated with the spectral files used by the operational Edwards-Slingo radiation scheme. Based on previous studies Jackson et al. [2001], Thelen [2006] it is apparent that the spectral files used in the operational model produce heating rates between ~2→0.4 hPa which are too large by ~1→2 K day^{-1}. Primarily, reasons for this bias are most likely the solar irradiance values of the Kurucz and Bell [1995] solar spectrum producing excessive ozone heating and an underestimation of the radiative cooling by H_2O. With the radiative time scale of the atmosphere between ~2→0.4 hPa being ~5→7 days, the above heating rate errors would produce the observed warm bias in the L50 analyses. Note that approaching ~0.4 hPa warm biases decrease in magnitude, most likely due to the impact of the cold bias associated with the Li and Shine ozone climatology noted above.

Here we have addressed thesis questions (1) and (2) from Chapter 1 for the L50 analyses, namely, **are there are any systematic biases in the temperature fields of the stratospheric assimilated data set?**, and,**through their magnitude, sign, location and seasonal evolution can we attribute these systematic biases to specific aspects of the operational system?**

9.1.2 L70 Validation

The validation results for the L70 analyses showed, as for the L50 analyses, the L70 analyses also have a latitude independent warm bias signal in the upper stratosphere and lower mesosphere accompanied by a latitude independent cold bias signal directly above. The magnitude and location of these L70 biases are similar to those seen in the L50 analyses, placing addition confidence that such biases are caused by the spectral files and ozone climatology which remain the same for both configurations.

The major difference between the L70 and L50 analyses occurs in the lower mesosphere polar winter, where they is a large decrease in the cold bias of this region. Such a reduction is expected by the increased vertical domain of the L70 model better representing the mean meridional circulation of the lower mesosphere, through increased knowledge of the wave forcing above under the downward control argument of Haynes et al. [1991]. The reduction of this lower mesosphere winter polar bias shows the positive impact of increasing the model lid. However, the results of Figures 4.10 and 4.11 show that strong biases now exist in the upper mesosphere approaching the L70 model lid. Here there is a large warm bias of ~60→70 K in the upper mesosphere summer polar latitudes. As for the cold biases of the L50 lower mesosphere winter polar regions, this strong warm bias shows a clear seasonal cycle which is not as apparent for equinox months. This seasonal cycle in upper mesospheric summer polar temperature biases would suggest that the gravity wave forcing in the L70 analyses is underestimated.

This insufficient gravity wave forcing is again likely to be compounded by the choice of a transparent lid for the USSP gravity wave scheme, and the non-physical boundary condition of w (set to zero) at the L70 model lid. Again though mass conservation one would expect to see large magnitude cold biases in the opposing upper mesosphere winter polar regions of the L70 analyses,

however only cold biases of ~15 K occur. Thus it is likely that biases approaching the model L70 model lid are due to a combination of insufficient gravity wave forcing and a latitude independent warm bias, most likely attributed to inaccuracies of the radiation scheme (since the impact of ozone heating has a relatively small contribution to the thermodynamic budget approaching 0.01 hPa). Likely reasons for this warm bias approaching the L70 lid are the inaccuracy of the spectral files at this altitude, which were originally optimised for tropospheric conditions and the fact that the operational radiation scheme does not account for the breakdown of LTE, which as shown by the modelling results of [Coy, 2010] can lead to a warm bias of ~15→25 K across all latitudes.

Here we have addressed thesis questions (1) and (2) from Chapter 1 for the L70 analyses, and also question (3), namely, **what is the impact of extending the model lid on the temperature biases, and what do the similarities/differences between the two models tell us?**

9.1.3 Off-line Studies

As model fields of gravity wave drag are not archived for the stratospheric assimilated data set, estimates of gravity wave drag in the operational model were obtained by applying the USSP scheme, full details given in Chapter 6, "off-line" to the temperature and wind fields of both the L50 and L70 analyses. The largest source of error for these off-line fields was expected to arise from the difference in altitudes of the initial launch spectrum, as detailed in Section 7.1. Here the magnitude of such error for both the L50 and L70 off-line results was shown to be generally less than 10%, with the L50 results most likely underestimating (overestimating) westward (eastward) forcing and the L70 results most likely overestimating (underestimating) eastward (westward) forcing.

Gravity wave forcing approaching the model lid of the L50 analyses was shown to have maximum magnitudes of ~12.5→17.5 ms^{-1} day^{-1} for both summer hemi-

sphere seasons. In the NH winter season maximum forcing magnitudes are also ~12.5→17.5 ms^{-1} day^{-1}, while the SH winter season has larger magnitudes of ~17.5→22.5 ms^{-1} day^{-1}. Such magnitudes of forcing are weaker, or at the lower limit, of previous estimates discussed in Section 7.3.3.

Gravity wave forcing approaching the model lid of the L70 analyses was shown to have maximum magnitudes of ~35→40 ms^{-1} day^{-1} for both summer hemisphere seasons. The SH winter season has larger maximum forcing magnitudes of ~50→55 ms^{-1} day^{-1} approaching the model lid, while the NH winter season has larger magnitudes of ~60→65 ms^{-1} day^{-1}. However, here the forcing magnitudes of the NH winter are likely to be unrealistic of the L70 analyses due to the uncharacteristic cold polar vortex in the lower mesosphere of this season, as discussed in Section 4.5. Such magnitudes of forcing are weaker, or at the lower limit, of previous estimates discussed in Section 7.3.3. Thus the off-line studies using the L50 and L70 analyses both support the argument that gravity wave forcing of the mesosphere is insufficient.

Off-line studies were also used to asses the sensitivity of the gravity wave forcing produced by the USSP when perturbing free parameters of the scheme. Based on observational and theoretical constraints the parameters β and s were chosen for such sensitivity studies. It was found, compared to standard values detailed in Table 6.1.7, that using β=0.14 increased the forcing by ~12→16 ms^{-1} day^{-1} in the winter upper mesosphere and ~8→12 ms^{-1} day^{-1} in the summer upper mesosphere. The sensitivity to perturbed values of s was smaller than the above values for β, where using a large perturbation of s=0.6 increased the forcing by ~6→8 ms^{-1} day^{-1} in both the summer and winter seasons.

In addition to the above momentum forcing studies, off-line studies of the direct heating rates due to gravity wave breaking were also performed using standard values of the free parameters in the USSP scheme. Full details of the formulation

and theoretical approach used in calculation of such heating rates was given in Section 6.2. When applied to the L70 analyses fields for June 2010 maximum direct heating rates of ~4→4.5 K day^{-1} occurred in the SH winter and ~7.5→8.0 K day^{-1} in the NH summer approaching the model lid. While these winter season heating rates are consistent with previous studies and observational estimates [Fomichev et al., 2002, Huang and Smith, 1991, Becker, 2004], summer value are noticeably larger by ~2~2.5 K day^{-1}. This difference is most likely attributed to the incorrect vertical wind shear of the summer season zonal mean winds, which act to increase the heating rate magnitudes through the residual work term of equation (6.41), thus highlighting the influence that biased wind fields have on the magnitude of heating rates obtained using the USSP formulation. As for the momentum forcing, the sensitivity of direct heating rates to perturbed free parameters in the USSP scheme was also investigated. Here the largest sensitivity was again found when perturbing the free parameter β, where using β=0.14 results in an increase of ~1.2 K day^{-1} in both winter and summer hemispheres.

9.1.4 UM Experiments

The remainder of this thesis concentrated on reducing the temperature biases associated with the insufficient mean meridional circulation of the mesosphere, via improvement of the currently operational USSP gravity wave scheme. Improvements to the USSP scheme were investigated via tuning of the free parameters in the scheme and the inclusion of additional physical processes associated with gravity wave breaking, i.e. the direct heating due to turbulent dissipation. Also investigated were the impacts on the gravity wave forcing from conserving momentum throughout the model domain, via imposing the opaque lid condition of the USSP scheme. To asses the impact of these three changes to the system, experimental integrations were performed on a free running 85 level configuration of the UM.

Concerning the USSP tuning experiment (referred to as the β run), based on

previous modelling studies it was decided that an increase of ~10 ms^{-1} day^{-1} would provide a significant strengthening of the mean meridional circulation, while providing gravity wave forcing magnitudes which were physically reasonable. From the results of the sensitivity experiments the parameter β was chosen as it demonstrated the required increase in forcing for the smallest perturbation from the standard USSP values, with a value of β=0.14 well within observational constraints. The opposite was true of the parameter s which required a larger perturbation to produce the required forcing, and is the least constrained of the USSP schemes free parameters. Experimental results for the tuning experiment were averaged over a two week period, following a spin up time of one month, for all four seasons and then compared to a control case using the standard parameter values of the USSP scheme. It was found that the largest increases in forcing occurred for solstice seasons, as expected since larger drag values are seen in these months and the parameter β scales the amount of drag produced by the USSP scheme. For the β run the increased forcing results in increased temperatures of the winter polar upper mesosphere by ~2→6 K, and decreased temperatures of the summer polar upper mesosphere by ~2→10 K. From these results it is expected that increasing the value of β will thus have a positive impact on the temperature biases of L70 analyses in the winter and summer polar upper mesosphere.

Both the control and β experiments were repeated using the opaque, instead of transparent, lid condition in the USSP scheme (referred to as control opaque and β opaque). Here we restrict our attention to the solstice seasons. For both runs applying the opaque lid produced more gravity wave forcing than the experiment without, as expected by Figure 7.4 which shows that there is significant amounts of momentum flux remaining at the L70 model lid. For the control opaque run the increased forcing is comparable in magnitude to that seen for the β run, however since the additional forcing is located closer to the model lid resulting in an increased temperature response at polar latitudes, most notice-

able in the summer hemisphere. Here the control opaque run has increased the temperatures of the winter polar upper mesosphere by ~2→6 K, and decreased the temperatures of the summer polar upper mesosphere by ~6→14 K. Applying the opaque lid condition in the β run results in the largest increase in gravity wave forcing, compared to the control case, and thus the largest temperature response at polar latitudes. Here the β opaque run has increased temperatures of the winter polar upper mesosphere by ~2→6 K, and decreased temperatures of the summer polar upper mesosphere by ~10→22 K. The fact the temperature response of the control opaque run is larger than that seen for the β run highlights the importance of conserving momentum throughout the model domain. Thus, it is expected that applying the opaque lid condition in the operational model will have a positive impact on the temperature biases seen in the upper mesosphere of the L70 analyses. It should also be noted that employing the opaque lid condition in the above experiments did not result in any numerical instability as noted in the previous study of Scaife et al. [2002] which used a configuration of the UM with a lower lid of 0.1 hPa.

In addition to the experiments detailed above, the impact of applying direct heating rates due to turbulent dissipation was also investigated. In the free running UM experiments two formulations of the direct heating rates were used. The first (referred to as Heat) used the opaque values of momentum flux in the heating rate calculations, while the second (referred to as Heat Trans) used the transparent momentum flux values. For both experiments the value of β=0.14 is used and the opaque lid condition is used for the momentum forcing from gravity wave breaking. As for the control opaque and β opaque experiments we restrict our attention to the solstice seasons. Compared to the β opaque experiment, both heating rate experiments show a large increase in upper mesospheric temperatures. However this increase in temperature is primarily seen at polar latitudes of the upper mesosphere and is noticeably larger for the summer hemisphere compared to the winter. This difference in increased temperatures

is larger than would be expected by the difference between summer and winter heating rates shown in Figure 8.6, and is a direct consequence of the change in strength of the mean meridional circulation seen when the direct heating rates are employed. It is therefore expected that applying the direct heating rates in either of the formulations presented above would have a negative impact on the L70 analyses, which as stated above already show strong warm biases approaching the model lid. Here the inclusion of the turbulent dissipation rates would most likely have a negative impact since they imply that gravity wave breaking always results in heating of the upper mesosphere. However, previous studies [Fomichev et al., 2002, Huang and Smith, 1991] have shown that the impact of eddy diffusion primarily results in a cooling of the upper mesosphere which dominates over the magnitudes of direct heating due to turbulent dissipation. Thus the neglect of this dominant eddy diffusion term is the most probable reason why applying just the contribution from turbulent dissipation would result in a negative impact on the upper mesospheric temperature biases of the L70 analyses. The negative impact of applying just the turbulent dissipation heating rate term, i.e. not including eddy diffusion, is compounded by the resulting change in the mean meridional circulation. Where as shown by the Heat and Heat Trans experiments, heating of the upper mesosphere primarily results in a weakening of the mean meridional circulation, compared to the β opaque experiment, and associated adiabatic warming (cooling) at the winter (summer) poles.

The results of the UM experiments, and those of the off-line studies detailed above, have addressed thesis questions (4) and (5) from Chapter 1, namely, **through the tuning of operational parameterizations is it possible to alleviate previously defined systematic temperature biases**, and, **what is the impact of applying previously neglected physical processes on the current temperature biases?**

9.2 Operational Considerations and Further Work

Based on the evidence presented in this thesis we recommend the following changes, and further work, to improve the representation of the upper stratosphere and mesosphere in the operational L70 analyses. Note that we also advise these improvements should be made sequentially in the order they are presented.

1. The radiative issues should be addressed. In the upper mesosphere this includes treatment of the breakdown of LTE and improvement of the spectral files used in the Edward-Slingo scheme. In the lower mesosphere this would again include improvement of the spectral files and replacing the currently operational Li and Shine ozone climatology, either through a more updated climatology (preferably one that includes a diurnal cycle) or through the real time assimilation of ozone data. Based on the relaxational nature of diabatic heating which acts like a spring to resist against dynamically forced adiabatic heating/cooling [Fels, 1985], it is logical to address these radiative issues first as by tuning the USSP we could just be compensating for diabatic heating biases. Furthermore addressing radiative biases in the upper stratosphere and lower mesosphere would result in reduced wind biases, through the adjustment of thermal wind balance, resulting in a more realistic filtering of the USSP gravity wave spectrum which again could be over compensated by incorrect tuning of the scheme.

2. The opaque lid condition in the USSP should be used, instead of the currently operational transparent condition. Both opaque lid experiments showed significant increases in gravity wave forcing and a positive temperature response at polar latitude in the upper mesosphere. Furthermore, conserving momentum is theoretically preferred [Shepherd and Shaw, 2004] and no issues of numerical instability were found for any of the experimental integrations.

3. Gravity wave direct heating rates due to turbulent dissipation should be included into the USSP scheme. However the contributions from eddy

diffusion should first be added to the scheme. Ideally this include contributions from both heat and momentum eddy diffusion, however the relative impact of these terms is dependent on the choice of Prandtl number which is still open to debate. If a relatively low Prandtl number is chosen then it is recommend that the impacts of eddy heat diffusion should be implemented first. Since the impacts of applying direct heating due to gravity wave breaking have been shown to strongly impact the strength of the mesospheric circulation, again it maybe counter productive to tune the gravity wave scheme until these operational changes have been made to the system.

4. Finally, once the above changes have been implement the gravity wave scheme could be tuned to improve any remaining biases. Ideally the tuning of the gravity wave scheme should be conducted within the fully operational system, i.e. with the inclusion of data assimilation, since, as shown by the free running UM experiments, the values of gravity wave forcing produced by the USSP scheme are sensitive to the biases in the zonal wind distribution which occur due to unconstrained stratospheric fields.

Here we have addressed thesis question number (6) from Chapter 1, **from what we have found above, what is the further logical approach to reduce operational temperature biases?**

Recent research involving using the middle atmospheric configuration of the UM [Shutts, 2010] has shown that the current operational model horizontal resolution, zonal and meridional grid spacing of ~40 km at mid-latitudes, is able to explicitly resolve a substantial proportion of the gravity wave spectrum in the middle atmosphere. However the currently operational vertical resolution is most likely insufficient to resolve gravity waves which have intrinsic frequencies close to the inertial frequency and vertical wavelengths less than ~2 km. For the currently operational time-step of 15 minutes used in the UM resolved gravity wave motion is strongly damped, via off-centring (time weights) used in the

semi-implicit time integration. The technique used by Shutts [2010] determines the characteristics of resolved gravity wave motion in the UM by subtracting forecasts produced with the operation time-step of 15 minutes from those with a shorter time-step of 2 minutes, where there is significant resolved gravity wave motion. Such studies could then be used to isolate gravity wave motions and determine wave characteristics to improve calibration of the free parameters of the USSP scheme. Also knowledge of the wavenumber-frequency spectra for selected regions allow the determination of non-orographic gravity forcing functions, which could be used to incorporate details of gravity wave sources into the USSP scheme. In addition this development of the USSP scheme, the above technique could also be used to ensure that the USSP scheme is not "double counting" the gravity wave forcing that has already been resolved in the operational model.

Evidence from the UM experiments also highlights the need for further research and understanding of the mean meridional circulation response to dynamical forcing of the mesosphere. The control run EP-Flux divergence distributions, shown in Figure 8.3, have been shown to be sensitive to changes in the gravity wave forcing applied in the β, control opaque and β opaque experiments. Further understanding of the conditions and mechanisms which generate planetary waves in the upper mesosphere and the large scale wave-wave interactions which exist between planetary (also tidal) and gravity waves is needed. Such interactions are highly sensitive to the relative phase speeds of gravity wave dissipation and large scale wind shear, and advances in this field would lead to a better understanding of the mean meridional circulation response to changes (coupled) in gravity and resolved forcing of the upper mesosphere. The response of the mean meridional circulation seen in the Heat and Heat Trans experiments also highlights the need to better understand the interplay, and competition, of the non-linear processes which control gravity wave dissipation (including turbulent generation) and their sensitivity to the atmospheric background state. Further advances in the above areas would hopefully result in parameterisa-

tions involving a more explicit treatment of the mechanism governing gravity wave propagation, dissipation and turbulent generation, thus improving the knowledge of the mean meridional circulation response to wave forcing and the non-local temperature response seen in modern GCM's such as the Met Office UM.

Appendix A

USSP Momentum Flux Profile

A.1 Derivation of momentum flux at z_k

Below is a derviation of the momentum flux profile at altitude z_k for each of the chop types detailed in Table VI. Details of how the individual chop types intersect the launch momentum flux spectrum are also given.

A.1.1 Chop Type 0

Here there are no intersect points and the launch momentum flux spectrum is conservatively propagated to z_k unaffected. Occurs when $A/B \geq 0$ and $B[1 + A/B]^{t-2} \geq 1$.

$$\rho_k \mathbf{F_{pktot}} = \rho_{k-1} \mathbf{F_{p(k-1)tot}} \tag{A.1}$$

A.1.2 Chop Type I

There is one intersect point m_{NlY} in the high wavenumber part of the spectrum. Occurs when $A/B < 0$ and $B[1+A/B]^{t-2} \geq 1$. Using (6.22) the total momentum flux at z_k in each ϕ_j direction is expressed as a fraction of the total launch momentum flux.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} \left[1 - A_0 \left(\int_{m_{lY}}^{\infty} \frac{m_*^{t-1}}{m^t} dm - \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \int_{m_{kY}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right) \right]$$

$$= \rho_l \mathbf{F_{pltot}} \left[1 - A_0 \left(\left[\frac{1}{1-t} \left(\frac{m_*}{m} \right)^{t-1} \right]_{m_{lY}}^{\infty} - B \left(\frac{N_k}{N_l} \right)^{t-1} \left[\frac{1}{1-t} \left(\frac{m_*}{m} \right)^{t-1} \right]_{m_{kY}}^{\infty} \right) \right]$$

$$= \rho_l \mathbf{F_{pltot}} \left[1 + \frac{A_0}{1-t} \left(\left(\frac{m_*}{m_{lY}} \right)^{t-1} - B \left(\frac{N_k\, m_*}{N_l\, m_{kY}} \right)^{t-1} \right) \right] \tag{A.2}$$

Noticing that,

$$\frac{N_k\, m_*}{N_l\, m_{kY}} = \frac{N_k\, m_{lY}}{N_l\, m_{kY}} \cdot \frac{m_*}{m_{lY}}$$

$$= \frac{N_l - m_{lY}(U_k - U_l)}{N_l} \cdot \frac{m_*}{m_{lY}}$$

$$= \left(1 + \frac{m_{lY}}{m_*} \cdot \frac{A}{B} \right) \frac{m_*}{m_{lY}} \tag{A.3}$$

Substitution of (A.3) into (A.2) results in the following expression.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} \left[1 + \frac{A_0}{1-t} \left(\frac{m_*}{m_{lY}} \right)^{t-1} \left(1 - B \left(1 + \frac{A}{B} m_{NlY} \right)^{t-1} \right) \right] \tag{A.4}$$

Rearranging (6.33) we have.

$$B \left[1 + \frac{A}{B} m_{NlY} \right]^{t-2} = 1$$

$$\left[1 + \frac{A}{B} m_{NlY} \right]^{t-1} \cdot \left[1 + \frac{A}{B} m_{NlY} \right]^{-1} = \frac{1}{B}$$

$$B \left[1 + \frac{A}{B} m_{NlY} \right]^{t-1} = \left[1 + \frac{A}{B} m_{NlY} \right]$$

Which after substitution into (A.4) leads to the following expression.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} \left[1 + \frac{A_0}{1-t} m_{NlY}^{1-t} \left(1 - \left(1 + \frac{A}{B} m_{NlY} \right) \right) \right]$$

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} \left[1 + \frac{A_0}{t-1} \frac{A}{B} m_{NlY}^{2-t} \right] \tag{A.5}$$

A.1.3 Chop Type II

There is one intersect m_{NlX} in the low wavenumber part of the spectrum. Occurs when $A/B \leq 0$ and $B[1 + A/B]^{t-2} < 1$. Using (6.22) we obtain the following expression for the total momentum flux at z_k in each ϕ_j direction.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} A_0 \left[\int_{m_{lmin}}^{m_{lX}} \frac{m^s}{m_*^{s+1}} dm + \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \int_{m_{kX}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{s+1} \left(\left(\frac{m_{lX}}{m_*} \right)^{s+1} - \left(\frac{m_{lmin}}{m_*} \right)^{s+1} \right) - \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \frac{1}{1-t} \left(\frac{m_*}{m_{kX}} \right)^{t-1} \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{s+1} \left(m_{NlX}^{s+1} - m_{Nlmin}^{s+1} \right) - \frac{B}{1-t} \left[\frac{N_k \, m_*}{N_l \, m_{kX}} \right]^{t-1} \right]. \tag{A.6}$$

Again noticing that,

$$\frac{N_k \, m_*}{N_l \, m_{kX}} = \left(1 + \frac{m_{lX}}{m_*} \cdot \frac{A}{B} \right) \frac{m_*}{m_{lX}}. \tag{A.7}$$

Substitution of (A.7) into (A.6) we obtain the following expression.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{s+1} \left(m_{NlX}^{s+1} - m_{Nlmin}^{s+1} \right) - \frac{1}{1-t} \cdot B \left(1 + \frac{A}{B} m_{NlX} \right)^{t-1} \left(\frac{m_*}{m_{lX}} \right)^{t-1} \right] \tag{A.8}$$

Rearranging (6.33) we have.

$$m_{NlX}^{s+t} - B \left[1 + \frac{A}{B} m_{NlX} \right]^{t-2} = 0$$

$$\left[1 + \frac{A}{B} m_{NlX} \right]^{t-1} \cdot \left[1 + \frac{A}{B} m_{NlX} \right]^{-1} = \frac{m_{NlX}^{s+t}}{B}$$

$$B \left[1 + \frac{A}{B} m_{NlX} \right]^{t-1} = m_{NlX}^{s+t} \left[1 + \frac{A}{B} m_{NlX} \right] \tag{A.9}$$

Which when substituted into (A.8) leads to the following expression.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{s+1} \left(m_{NlX}^{s+1} - m_{Nlmin}^{s+1} \right) - m_{NlX}^{1-t} \frac{1}{1-t} m_{NlX}^{s+t} \left(1 + \frac{A}{B} m_{NlX} \right) \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{s+1} \left(m_{NlX}^{s+1} - m_{Nlmin}^{s+1} \right) - \frac{1}{1-t} m_{NlX}^{s+1} \left(1 + \frac{A}{B} m_{NlX} \right) \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\left(\frac{1}{s+1} - \frac{1}{1-t} - \frac{1}{1-t} \frac{A}{B} m_{NlX} \right) m_{NLX}^{s+1} - \frac{1}{s+1} m_{Nlmin}^{s+1} \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\left(\frac{1}{s+1} + \frac{1}{t-1} + \frac{1}{t-1} \frac{A}{B} m_{NlX} \right) m_{NlX}^{s+1} - \frac{1}{s+1} m_{Nlmin}^{s+1} \right]$$

$$\rho_k \mathbf{F_{pktot}} = \frac{\rho_l \mathbf{F_{pltot}} A_0}{(s+1)(t-1)} \left[\left(s + t + (s+1) \frac{A}{B} m_{NlX} \right) m_{NlX}^{s+1} - (t-1) m_{Nlmin}^{s+1} \right]. \quad \text{(A.10)}$$

There are three different cases of chop type II, where (A.10) is the most common and is chop type IIa described above. The second case, known as IIb, occurs when $m_{min} \geq m_{lX}$ and chopping is so severe that only the tail from the quasi-saturated spectrum at z_k is left. For case IIb the momentum flux at z_k is given by.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \int_{m_{kmin}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{t-1} \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \left(\frac{m_*}{m_{kmin}} \right)^{t-1} \right]$$

$$= \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{1}{t-1} B \left(1 + \frac{A}{B} m_{Nlmin} \right)^{t-1} \left(\frac{m_*}{m_{Nlmin}} \right)^{t-1} \right]$$

$$\rho_k \mathbf{F_{pktot}} = \frac{\rho_l \mathbf{F_{pltot}} A_0}{(t-1)} \frac{B}{m_{Nlmin}^{t-1}} \left(1 + \frac{A}{B} m_{Nlmin} \right)^{t-1} \tag{A.11}$$

The third case, known as chop type IIc, occurs when $A = 0$, which simplifies (A.10) to the following.

$$\rho_k \mathbf{F_{pktot}} = \frac{\rho_l \mathbf{F_{pltot}} A_0}{(s+1)(t-1)} \left[(s+t) m_{NlX}^{s+1} - (t-1) m_{Nlmin}^{s+1} \right] \tag{A.12}$$

A.1.4 Chop Type III

There are two intersect points, m_{NlX} and m_{NlY} respectively in the low and high wavenumber parts of the spectrum. Occurs when $A > 0$ and $B[1 + A/B]^{t-2} < 1$. Using a combination of the terms in (A.1.2) we obtain the following expression for the total momentum flux at z_k in each ϕ_j direction.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} A_0 * \tag{A.13}$$

$$\left[\int_{m_{min}}^{m_{lX}} \frac{m^s}{m_*^{s+1}} dm + \frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \left(\int_{m_{kX}}^{\infty} \frac{m_*^{t-1}}{m^t} dm - \int_{m_{kY}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right) + \int_{m_{lY}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right]$$

Using the corresponding terms in (A.5) and (A.10), (A.13) evaluates to.

$$\rho_k \mathbf{F_{pktot}} = \frac{\rho_l \mathbf{F_{pltot}}}{(s+1)(t-1)} \left[-(t-1) m_{Nlmin}^{s+1} + \left(s + t + (s+1) \frac{A}{B} m_{NlX} \right) m_{NlX}^{s+1} - (s+1) \frac{A}{B} m_{NlY}^{2-t} \right] \tag{A.14}$$

There are two cases of chop type III, where (A.14) is the most common and is known as chop type IIIa. The second, known as chop type IIIb, occurs when

$m_{min} \geq m_{lX}$. For case IIIb the momentum flux at z_k in each ϕ_j direction is given by.

$$\rho_k \mathbf{F_{pktot}} = \rho_l \mathbf{F_{pltot}} A_0 \left[\frac{\rho_k}{\rho_l} \frac{C_{Sk}}{C_{l0}} \left(\int_{m_{kX}}^{\infty} \frac{m_*^{t-1}}{m^t} dm - \int_{m_{kY}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right) + \int_{m_{lY}}^{\infty} \frac{m_*^{t-1}}{m^t} dm \right] \quad \text{(A.15)}$$

Again using the corresponding terms from (A.5) and (A.10), (A.15) evaluates to.

$$\rho_k \mathbf{F_{pktot}} = \frac{\rho_l \mathbf{F_{pltot}} A_0}{(t-1)} \left[\frac{B}{m_{Nlmin}^{t-1}} \left(1 + \frac{A}{B} m_{Nlmin} \right)^{t-1} - \frac{A}{B} m_{NlY}^{2-t} \right] \quad \text{(A.16)}$$

A.1.5 Chop Type IV

Here the entire launch spectrum has been Doppler shifted to ∞ at z_k. Occurs when $A/B < 0$ and $|B/A| \leq m_{Nlmin}$. Therefore no flux can propagate higher resulting in.

$$\rho_k \mathbf{F_{pktot}} = 0$$

Bibliography

R.F. Adler. Mean meridional circulation in the Southern Hemisphere stratosphere based on satellite information. *Journal of the Atmospheric Sciences*, 32(5):893–898, 1975.

R.A. Akmaev. Diagnostics and simulation of an annual cycle in the middle atmosphere. *Journal of Geophysical Research*, 99(D9):18933, 1994.

S.J. Allen and R.A. Vincent. Gravity wave activity in the lower atmosphere: Seasonal and latitudinal variations. *Journal of Geophysical Research-Atmospheres*, 100(D1), 1995.

Ø. Andreassen, P.E.R.Ø. Hvidsten, D.C. Fritts, and S. Arendt. Vorticity dynamics in a breaking internal gravity wave. Part 1. Initial instability evolution. *Journal of Fluid Mechanics*, 367:27–46, 1998.

D.G. Andrews and M.E. McIntyre. Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration. *Journal of the Atmospheric Sciences*, 33(11):2031–2048, 1976.

D.G. Andrews, J.R. Holton, and C.B. Leovy. *Middle atmosphere dynamics*. Academic press, 1987.

A. Arakawa and V. Lamb. 1977. Computational design of the basic dynamical processes of the UCLA general circulation model. *Meths. Comput. Phys*, 17: 173–265, 1977.

H.H. Aumann, M.T. Chahine, C. Gautier, M.D. Goldberg, E. Kalnay, L.M. McMillin, H. Revercomb, P.W. Rosenkranz, W.L. Smith, D.H. Staelin, et al.

AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(2):253–264, 2003.

M.P. Baldwin and T.J. Dunkerton. Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research*, 104(D24): 30937, 1999.

J.J. Barnett. The mean meridional temperature behaviour of the stratosphere from November 1970 to November 1971 derived from measurements by the Selective Chopper Radiometer on Nimbus IV. *Quarterly Journal of the Royal Meteorological Society*, 100(426), 1974.

S.R. Beagley, J. De Grandpre, J.N. Koshy, N.A. McFarlane, and T.G. Shepherd. Radiative-Dynamical Climatology of the First-Generation Canadian Middle Atmoshpere Model. *Atmosphere Ocean*, 35:293–331, 1997.

E. Becker. Direct heating rates associated with gravity wave saturation. *Journal of atmospheric and solar-terrestrial physics*, 66(6-9):683–696, 2004.

E. Becker and D.C. Fritts. Enhanced gravity-wave activity and interhemispheric coupling during the MaCWAVE/MIDAS northern summer program 2002. *Annales Geophysicae*, 24(4):1175–1188, 2006.

E. Becker and G. Schmitz. Energy deposition and turbulent dissipation owing to gravity waves in the mesosphere. *Journal of the Atmospheric Sciences*, 59:54–68, 2002.

P. Bergthorsson and B. Doos. Numerical weather map analysis. *Tellus*, 7(3): 329–340, 1955.

P.F. Bernath, C.T. McElroy, M.C. Abrams, C.D. Boone, M. Butler, C. Camy-Peyret, M. Carleer, C. Clerbaux, P.F. Coheur, R. Colin, et al. Atmospheric chemistry experiment(ACE): mission overview. In *Proceedings of SPIE*, volume 5542, pages 146–156, 2004.

B.A. Boville. The Influence of the Polar Night Jet on the Tropospheric Circulation in a GCM. *Journal of Atmospheric Sciences*, 41:1132–1142, 1984.

B.A. Boville. Middle atmosphere version of CCM2 (MACCM2): Annual cycle and interannual variability. *Journal of Geophysical Research-Atmospheres*, 100 (D5), 1995.

F.P. Bretherton. The propagation of groups of internal gravity waves in a shear flow. *Quarterly Journal of the Royal Meteorological Society*, 92(394):466–480, 1966.

A.W. Brewer. Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere. *Quarterly Journal of the Royal Meteorological Society*, 75(326), 1949.

A.C. Bushell. Preliminary Test of Wenyi Zhong Radiation Changes. *UK-MOMA Model Development Report*, 1:1–13, 2005.

A.C. Bushell. personal communication, 2010.

A.C. Bushell, A.A Scaife, and C.D. Warner. UMDP no 34: Non-orographic (Spectral) Gravity Wave Parametrization. Technical report, Met Office, 2007.

N. Butchart and J. Austin. Middle Atmosphere Climatologies from the Troposphere–Stratosphere Configuration of the UKMOs Unified Model. *Journal of the Atmospheric Sciences*, 55(17):2782–2809, 1998.

S. Chapman, R.S. Lindzen, VM Canuto, and S. Chapman. *Atmospheric tides*. Reidel Dordrecht, 1970.

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

M.D. Chou and M.J. Suarez. A solar radiation parameterization for atmospheric studies. *NASA Tech. Memo*, 104606:40, 1999.

M.D. Chou, M.J. Suarez, X.Z. Liang, and M.M.H. Yan. A thermal infrared radiation parameterization for atmospheric studies. *NASA Tech. Memo*, 2001.

B. Christiansen. Downward propagation of zonal mean zonal wind anomalies from the stratosphere to the troposphere: Model and reanalysis. *Journal of Geophysical Research*, 106(D21):27307, 2001.

C. Cot. Equatorial mesoscale wind and temperature fluctuations in the lower atmosphere. *Journal of Geophysical Research-Atmospheres*, 106(D2), 2001.

L. Coy. personal communication, 2010.

L. Coy and D.C. Fritts. Gravity wave heat fluxes: A Lagrangian approach. *Journal of the atmospheric sciences*, 45(12):1770–1780, 1988.

L. Coy, S. Eckermann, and K. Hoppel. Planetary Wave Breaking and Tropospheric Forcing as Seen in the Stratospheric Sudden Warming of 2006. *Journal of the atmospheric sciences*, 66(2):495–507, 2009.

A.P. Cracknell. *The advanced very high resolution radiometer (AVHRR)*. CRC Press, 1997.

M.J.P. Cullen. The unified forecast/climate model. *Meteorological Magazine*, 122 (1449):81–94, 1993.

M.J.P. Cullen and T. Davies. A conservative split-explicit integration scheme with fourth-order horizontal advection. *Quarterly Journal of the Royal Meteorological Society*, 117(501):993–1002, 1991.

T. Davies, M.J.P. Cullen, A.J. Malcolm, M.H. Mawson, A. Staniforth, A.A. White, and N. Wood. A new dynamical core for the Met Offices global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131:1759–1782, 2005.

G.M.B. Dobson. Origin and distribution of the polyatomic molecules in the atmosphere. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 187–193, 1956.

A. Dowdy, R.A. Vincent, K. Igarashi, Y. Murayama, and D.J. Murphy. A comparison of mean winds and gravity wave activity in the northern and southern polar MLT. *Geophysical Research Letters*, 28(8):1475–1478, 2001.

T. Dunkerton. On the mean meridional mass motions of the stratosphere and mesosphere. *Journal of the Atmospheric Sciences*, 35(12):2325–2333, 1978.

T.J. Dunkerton. Nonlinear Hadley circulation driven by asymmetric differential heating. *Journal of the Atmospheric Sciences*, 46(7):956–974, 1989.

I. Durre, R.S. Vose, and D.B. Wuertz. Overview of the integrated global radiosonde archive. *Journal of Climate*, 19(1), 2006.

H.J Edmon Jr, B.J. Hoskins, and M.E. McIntyre. Eliassen-Plam cross sections for the troposphere. *Journal of Atmospheric Sciences*, 37:2600–2616, 1995.

J.M. Edwards and A. Slingo. Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. *Quarterly Journal of the Royal Meteorological Society*, 122(531):689–719, 1996.

J.M. Edwards and J-C. Thelen. UMDP no 23a: Spectral files for the radiation scheme. Technical report, Met Office, 2003.

J.M. Edwards, J-C. Thelen, and W.J Ingram. UMDP no 23: The radiation code. Technical report, Met Office, 2004.

A. Eliassen. Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophisica Norvegica*, 5, 1951.

A. Eliassen and E. Palm. *On the transfer of energy in stationary mountain waves*. Aschehoug i komm., 1961.

J. Eluszkiewicz, D. Crisp, R. Zurek, L. Elson, E. Fishbein, L. Froidevaux, J. Waters, RG Grainger, A. Lambert, R. Harwood, et al. Residual circulation in the stratosphere and lower mesosphere as diagnosed from Microwave Limb Sounder data. *Journal of the Atmospheric Sciences*, 53(2):217–240, 1996.

M. Ern, P. Preusse, M.J. Alexander, and C.D. Warner. Absolute values of gravity wave momentum flux derived from satellite data. *J. Geophys. Res*, 109(1), 2004.

M. Ern, P. Preusse, and C.D. Warner. Some experimental constraints for spectral parameters used in the Warner and McIntyre gravity wave parameterization scheme. *Atmos. Chem. Phys*, 6:4361–4381, 2006.

M. Ern, P. Preusse, M. Krebsbach, M.G. Mlynczak, et al. Equatorial wave analysis from SABER and ECMWF temperatures. *Atmospheric Chemistry and Physics Discussions*, 7(4):11685–11723, 2007.

S.B. Fels. Radiative-dynamical interactions in the middle atmosphere. *Advances in geophysics.*, 28:277–300, 1985.

H. Fischer, M. Birk, C. Blom, B. Carli, M. Carlotti, T. Von Clarmann, L. Delbouille, A. Dudhia, D. Ehhalt, M. Endemann, et al. MIPAS: an instrument for atmospheric and climate research. *Atmospheric Chemistry and Physics Discussions*, 7 (3):8795–8893, 2007.

E.F. Fishbien, R.E. Cofield, L. Froidevaux, R.F. Jarnot, T. Lungu, W.G. Read, Z. Shippony, J.W. Waters, I.S. Mcdermid, T.J. Mcgee, et al. Validation of UARS Microwave Limb sounder temperature and pressure measurements. *Journal of geophysical research*, 101(D 6):9983–10016, 1996.

E.L. Fleming, Scientific, Technical Information Branch, National Aeronautics, Space Administration, and United States. *Monthly mean global climatology of temperature, wind, geopotential height, and pressure for 0-120 km*. National Aeronautics and Space Administration, Scientific and Technical Information Branch, 1988.

E.L. Fleming, S. Chandra, J.J. Barnett, and M. Corney. Zonal mean temperature, pressure, zonal wind, and geopotential height as functions of latitude, COSPAR International Reference Atmosphere: 1986, Part II: Middle Atmosphere Models. *Adv. Space Res*, 10(12):11–59, 1990.

VI Fomichev and GM Shved. Net radiative heating in the middle atmosphere. *Journal of Atmospheric and Terrestrial Physics*, 50(8):671–688, 1988.

V.I. Fomichev, W.E. Ward, S.R. Beagley, C. McLandress, J.C. McConnell, N.A. McFarlane, and T.G. Shepherd. Extended Canadian middle atmosphere model: zonal-mean climatology and physical parameterizations. *J. Geophys. Res*, 107 (D10):4087, 2002.

V.I. Fomichev, V.P. Ogibalov, and S.R. Beagley. Solar heating by the near-IR CO2 bands in the mesosphere. *Geophysical Research Letters*, 31(21):L21102, 2004. ISSN 0094-8276.

Y. Fouquart and B. Bonnel. Computations of solar heating of the earth's atmosphere- A new parameterization. *Beitrage zur Physik der Atmosphare*, 53: 35–62, 1980.

D.C. Fritts and M.J. Alexander. Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics-Richmond*, 41(1):1003, 2003.

D.C. Fritts and W. Lu. Spectral estimates of gravity wave energy and momentum fluxes. Part II: Parameterization of wave forcing and variability. *Journal of the Atmospheric Sciences*, 50(22):3695–3713, 1993.

D.C. Fritts and T.E. Vanzandt. Spectral estimates of gravity wave energy and momentum fluxes. Part I: Energy dissipation, acceleration, and constraints. *Journal of the Atmospheric Sciences*, 50(22):3685–3694, 1993.

D.C. Fritts and R.A. Vincent. Mesospheric momentum flux studies at Adelaide, Australia: Observations and a gravity wave-tidal interaction model. *Journal of the Atmospheric Sciences*, 44(3):605–619, 1987.

D.C. Fritts and J.A. Werne. Turbulence dynamics and mixing due to gravity waves in the lower and middle atmosphere. *Geophysical monograph*, 123:143–159, 2000.

D.C. Fritts and L. Yuan. Measurement of momentum fluxes near the summer mesopause at Poker Flat, Alaska. *Journal of the Atmospheric Sciences*, 46(16): 2569–2579, 1989.

D.C. Fritts, S. Arendt, and Ø. Andreassen. Vorticity dynamics in a breaking internal gravity wave. Part 2. Vortex interactions and transition to turbulence. *Journal of Fluid Mechanics*, 367:47–65, 1998.

D.C. Fritts, C. Bizon, J.A. Werne, and C.K. Meyer. Layering accompanying turbulence generation due to shear instability and gravity-wave breaking. *J. Geophys. Res*, 108(20):1–13, 2003.

A.J. Gadd. The 15-level weather prediction model. *Meteorological Magazine*, 114 (1357):222–226, 1985.

R.R. Garcia and B.A. Boville. Downward Control of the Mean Meridional Circulation and Temperature Distribution of the Polar Winter Stratosphere. *Journal of the Atmospheric Sciences*, 51(15):2238–2245, 1994.

R.R. Garcia and S. Solomon. The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere. *Journal of Geophysical Research*, 90(D2):3850–3868, 1985.

R.R. Garcia, R. Lieberman, J.M. Russell III, and M.G. Mlynczak. Large-scale waves in the mesosphere and lower thermosphere observed by SABER. *Journal of the Atmospheric Sciences*, 62(12):4384–4399, 2005.

A.E. Gill. *Atmosphere-ocean dynamics*. Academic press, 1982.

J.C. Gille and L.V. Lyjak. Radiative heating and cooling rates in the middle atmosphere. *Journal of the Atmospheric Sciences*, 43(20):2215–2229, 1986.

J.C. Gille and J.M. Russell. The Limb Infrared Monitor of the Stratosphere- Experiment description, performance, and results. *Journal of Geophysical Research*, 89:5125–5140, 1984.

J.C. Gille, L.V. Lyjak, and A.K. Smith. The global residual mean circulation in the middle atmosphere for the northern winter period. *Journal of the Atmospheric Sciences*, 44(10):1437–1454, 1987.

A. Gobiet, U. Foelsche, AK Steiner, M. Borsche, G. Kirchengast, and J. Wickert. Climatological validation of stratospheric temperatures in ECMWF operational analyses with CHAMP radio occultation data. *Geophysical Research Letters*, 32(12):L12806, 2005.

C. Gordon, C. Cooper, C.A. Senior, H. Banks, J.M. Gregory, T.C. Johns, J.F.B. Mitchell, and R.A. Wood. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, 16(2):147–168, 2000.

D. Gregory, G.J. Shutts, and J.R. Mitchell. A new gravity-wave-drag scheme incorporating anisotropic orography and low-level wave breaking: Impact upon the climate of the UK Meteorological Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, 124(546):463–493, 1998.

K.U. Grossmann, O. Gusev, M. Kaufmann, A. Kutepov, and P. Knieling. A review of the scientific results from the CRISTA missions. *Advances in Space Research*, 34(8):1715–1721, 2004.

G.V. Groves. Hough components of ozone heating. *Journal of Atmospheric and Terrestrial Physics*, 44(2):111–121, 1982.

M.E. Hagan and J.M. Forbes. Migrating and nonmigrating diurnal tides in the middle and upper atmosphere excited by tropospheric latent heat release. *Journal of geophysical research*, 107(D24):4754, 2002.

M.E. Hagan and J.M. Roble. Modeling diurnal tidal variability with the NCAR TIME-GCM. *Journal of geophysical research*, 106:24869–24882, 2001.

C.D. Hall, R.A. Stratton, and M.L. Gallani. Climate simulations with the Unified Model: AMIP runs. Technical report, Met Office, 1995.

K. Hamilton. Diagnostic study of the momentum balance in the northern hemisphere winter stratosphere. *Monthly Weather Review*, 111(7):1434–1441, 1983.

K. Hamilton. Interannual variability in the Northern Hemisphere winter middle atmosphere in control and perturbed experiments with the GFDL SKYHI general circulation model. *Journal of the Atmospheric Sciences*, 52(1):44–66, 1995.

K. Hamilton and J.D. Mahlman. General circulation model simulation of the semiannual oscillation of the tropical middle atmosphere. *Journal of the Atmospheric Sciences*, 45(21):3212–3235, 1988.

K. Hamilton, R.J. Wilson, and R.S. Hemler. 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. *Journal of the Atmospheric Sciences*, 56(22):3829–3846, 1999.

S.C. Hardiman, N. Butchart, S.M. Osprey, L.J. Gray, A.C. Bushell, and T.J. Hinton. The Climatology of the Middle Atmosphere in a Vertically Extended Version of the Met Office's Climate Model. Part I: Mean State. *Journal of the Atmospheric Sciences*, 67(5):1509–1525, 2010.

D.L. Hartmann. The dynamical climatology of the stratosphere in the Southern Hemisphere during late winter 1973. *Journal of the Atmospheric Sciences*, 33(9): 1789–1802, 1976.

P.H. Haynes. Stratospheric Dynamics. *Annual Review of Fluid Mechanics*, 37: 263–293, 2005.

P.H. Haynes, M.E. McIntyre, T.G. Shepherd, C.J. Marks, and K.P. Shine. On the downward control of extratropical diabatic circulations by eddy-induced mean zonal forces. *Journal of the Atmospheric Sciences*, 48(4):651–678, 1991.

P.B. Hays, V.J. Abreu, M.E. Dobbs, D.A. Gell, H.J. Grassl, and W.R. Skinner. The high-resolution Doppler imager on the Upper Atmosphere Research Satellite. *Journal of Geophysical Research (ISSN 0148-0227)*, 98(D6), 1993.

A.E. Hedin, E.L. Fleming, A.H. Manson, F.J. Schmidlin, S.K. Avery, R.R. Clark, S.J. Franke, G.J. Fraser, T. Tsuda, F. Vial, et al. Empirical wind model for the upper, middle and lower atmosphere. *Journal of atmospheric and terrestrial physics*, 58(13):1421–1447, 1996.

A. Hertzog and F. Vial. A study of the dynamics of the equatorial lower stratosphere by use of ultra-long-duration balloons. 2. Gravity waves. *J. Geophys. Res*, 106(22):745–22, 2001.

A. Hertzog, C. Souprayen, and A. Hauchecorne. Measurements of gravity wave activity in the lower stratosphere by Doppler lidar. *Journal of Geophysical Research-Atmospheres*, 106(D8), 2001.

C.O. Hines. Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 1: Basic formulation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 59(4):371–386, 1997.

C.O. Hines. Correction to Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 1: Basic formulation.. *J. Atmos. Sol.–Terr. Phys*, 61:941, 1999.

C.O. Hines and C.A. Reddy. On the propagation of atmospheric gravity waves through regions of wind shear. *Journal of Geophysical Research*, 72(3):1015–1034, 1967.

M.H. Hitchman and C.B. Leovy. Estimation of the Kelvin wave contribution to the semiannual oscillation. *Journal of the Atmospheric Sciences*, 45(9):1462–1475, 1988.

M.H. Hitchman, J.C. Gille, C.D. Rodgers, and G. Brasseur. The separated polar winter stratopause: A gravity wave driven climatological feature. *Journal of the Atmospheric Sciences*, 46(3):410–422, 1989.

R.R. Hodges. Eddy diffusion coefficients due to instabilities in internal gravity waves. *Journal of Geophysical Research*, 74:4087–4090, 1969.

J.R. Holton. The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. *Journal of the Atmospheric Sciences*, 39(4): 791–799, 1982.

J.R. Holton. The influence of gravity wave breaking on the general circulation of the middle atmosphere. *Journal of the Atmospheric Sciences*, 40(10):2497–2507, 1983.

J.R. Holton. *An introduction to dynamic meteorology*. Academic press, 2004.

J.R. Holton and M.J. Alexander. The role of waves in the transport circulation of the middle atmosphere. *Geophysical monograph*, 123:21–35, 2000.

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

M.M. Huaman and B.B. Balsley. Differences in near-mesopause summer winds, temperatures, and water vapor at northern and southern latitudes as possible causal factors for inter-hemispheric PMSE differences. *Geophysical Research Letters*, 26(11):1529–1532, 1999.

T.Y.W. Huang and A.K. Smith. The mesospheric diabatic circulation and the parameterized thermal effect of gravity wave breaking on the circulation. *Journal of the Atmospheric Sciences*, 48(8):1093–1111, 1991.

W.J. Ingram. Radiation, Meteorological Office Unified Model Doc Paper no.23. Technical report, Met Office, 1990.

D.R. Jackson. Assimilation of EOS MLS ozone observations in the Met Office data-assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 133(628):1771–1788, 2007.

D.R. Jackson. personal communication, 2010.

D.R. Jackson and L.J. Gray. Simulation of the semi-annual oscillation of the equatorial middle atmosphere using the Extended UGAMP General Circu-

lation Model. *Quarterly Journal of the Royal Meteorological Society*, 120(520): 1559–1588, 1994.

D.R. Jackson, J. Austin, and N. Butchart. An Updated Climatology of the Troposphere–Stratosphere Configuration of the Met Office's Unified Model. *Journal of the Atmospheric Sciences*, 58(14):2000–2008, 2001.

D.R. Jackson, M. Keil, and B.J. Devenish. Use of Canadian Quick covariances in the Met Office data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 134(635):1567–1582, 2008.

J.H. Jiang, S.D. Eckermann, D.L. Wu, and D.Y. Wang. Inter-annual variation of gravity waves in the Arctic and Antarctic winter middle atmosphere. *Advances in Space Research*, 38(11):2418–2423, 2006.

M.N. Juckes and M.E. McIntyre. A high-resolution one-layer model of breaking planetary waves in the stratosphere. *Nature*, 328:590–596, 1987.

H. Kanzawa. Warm stratopause in the Antarctic winter. *Journal of the Atmospheric Sciences*, 46(3):435–438, 1989.

B. Karlsson, H. Kornich, and J. Gumbel. Evidence for interhemispheric stratosphere-mesosphere coupling derived from noctilucent cloud properties. *Geophysical Research Letters*, 34(16):L16806, 2007.

B. Karlsson, C. McLandress, and T.G. Shepherd. Inter-hemispheric mesospheric coupling in a comprehensive middle atmosphere model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 71(3-4):518–530, 2009.

A. Kasahara, T. Sasamori, and W.M. Washington. Simulation Experiments with a 12-Layer Stratospheric Global Circulation Model. I. Dynamical Effect of the Earth's Orography and Thermal Influence of Continentality. *Journal of Atmospheric Sciences*, 30:1229–1251, 1973.

J.T. Kiehl and S. Solomon. On the radiative balance of the stratosphere. *Journal of the Atmospheric Sciences*, 43(14):1525–1534, 1986.

J.T. Kiehl, J.J. Hack, G.B. Bonan, and B. Boville. Description of the NCAR Community Climate Model(CCM 3). *NASA*, 1996.

E.W. Koenig. Performance of the HIRS-2 instrument on TIROS-N(High Resolution Infrared Radiation Sounder). *Remote sensing of atmospheres and oceans*, page 1980, 1980.

H. Kornich and E. Becker. A simple model for the inter-hemispheric coupling of the middle atmosphere circulation. *Advances in Space Research*, 45:661–668, 2010.

D.B. Kunkee, G.A. Poe, D.J. Boucher, S.D. Swadley, Y. Hong, J.E. Wessel, and E.A. Uliana. Design and Evaluation of the First Special Sensor Microwave Imager/Sounder. *IEEE Transcations on geoscience and remote sensing*, 46(4 Part 1):863–883, 2008.

R.L. Kurucz and B. Bell. Kurucz CD-Rom 23. *Atomic Line Data (Cambridge: SAO)*, 1995.

K. Labitzke. Temperature Changes in the Mesosphere and Stratosphere Connected with Circulation Changes in Winter. *Journal of Atmospheric Sciences*, 29: 756–766, 1972.

D. Labs and H. Neckel. Transformation of the absolute solar radiation data into the International Practical Temperature Scale of 1968. *Solar Physics*, 15(1): 79–87, 1970.

U. Langematz and S. Pawson. The Berlin troposphere-stratosphere-mesosphere GCM: Climatology and forcing mechanisms. *Quarterly Journal of the Royal Meteorological Society*, 123(540):1075–1096, 1997.

D. Li and K.P. Shine. A 4-dimensional ozone climatology for UGAMP models. *UGAMP Internal Rep*, 35, 1995.

J. Lighthill. *Waves in fluids*. Cambridge University Press, 1978.

R.S. Lindzen. Wave-mean flow interactions in the upper atmosphere. *Boundary-Layer Meteorology*, 4(1):327–343, 1973.

R.S. Lindzen. Turbulence and stress owing to gravity wave and tidal breakdown. *Journal of Geophysical Research*, 86:9707–9714, 1981.

R.S. Lindzen and J.R. Holton. A Theory of the Quasi-Biennial Oscillation. *Journal of Atmospheric Sciences*, 25:1095–1107, 1968.

J.-L. Lions, R. Teman, and S. Wang. New formulations of the primitive equations of atmosphere and applications. *Nonlinearity*, 5:237–288, 1992.

N.J. Livesey and W.G. Read. Direct retrieval of line-of-sight atmospheric structure from limb sounding observations. *Geophysical Research Letters*, 27(6):891–894, 2000.

N.J. Livesey, W.G. Read, L. Froidevaux, J.W. Waters, M.L. Santee, H.C. Pumphrey, D.L. Wu, Z. Shippony, and R.F. Jarnot. The UARS Microwave Limb Sounder version 5 dataset: Theory, characterization and validation. *J. Geophys. Res*, 108 (D13):4378, 2003.

N.J. Livesey, W. Van Snyder, W.G. Read, and P.A. Wagner. Retrieval algorithms for the EOS Microwave limb sounder (MLS). *Geoscience and Remote Sensing, IEEE Transactions on*, 44(5):1144–1155, 2006.

M. Loiselet, N. Stricker, Y. Menard, and JP Luntama. GRAS-Metop's GPS-based atmospheric sounder. *ESA bulletin*, 102:38–44, 2000.

J. London. Radiative energy sources and sinks in the stratosphere and mesosphere. In *Atmospheric Ozone and its Variation and Human Influences*, page 703, 1980.

M. López-Puertas, G. Zaragoza, and M. Lopez-Valverde. Non local thermodynamic equilibrium(LTE) atmospheric limb emission at 4. 6 microns. II- An analysis of the daytime wideband radiances as measured by UARS improved

stratospheric and mesospheric sounder. *Journal of Geophysical Research*, 103: 8515–8530, 1998.

A.C. Lorenc, R.S. Bell, and B. Macpherson. The Meteorological Office analysis correction data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 117(497):59–89, 1991.

A.C. Lorenc, S.P. Ballard, R.S. Bell, N.B. Ingleby, P.L.F. Andrews, D.M. Barker, J.R. Bray, A.M. Clayton, T. Dalby, D. Li, et al. The Met. Office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000.

F.J. Lübken. On the extraction of turbulent parameters from atmospheric density fluctuations. *Journal of Geophysical Research*, 97(D18):20385, 1992.

F.J. Lübken. Seasonal variation of turbulent energy dissipation rates at high latitudes as determined by in situ measurements of neutral density fluctuations. *Journal of geophysical research*, 102(D12):13441, 1997.

F.J. Lübken. Thermal structure of the Arctic summer mesosphere. *Journal of geophysical research*, 104(D 8):9135–9149, 1999.

S. Mahmood. personal communication, 2010.

S. Manabe and B.G. Hunt. Experiments with a stratospheric general circulation model. *Monthly Weather Review*, 96(8):477–502, 1968.

S. Manabe and J.D. Mahlman. Simulation of seasonal and interhemispheric variations in the stratospheric circulation. *Journal of the Atmospheric Sciences*, 33(11):2185–2217, 1976.

J. Manners. personal communication, 2010.

C.J. Marks. Some features of the climatology of the middle atmosphere revealed by Nimbus 5 and 6. *Journal of the Atmospheric Sciences*, 46(16):2485–2508, 1989.

C. Mathison, D.R. Jackson, and Keil M. NWP Technical Report No. 502: 'Methods of improving the representation of ozone in the Met Office model'. Technical report, Met Office, 2007.

K. Matthes, Y. Kuroda, K. Kodera, and U. Langematz. Transfer of the solar signal from the stratosphere to the troposphere: Northern winter. *Journal of Geophysical Research*, 111(D6):D06108, 2006.

N.A. McFarlane. The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *Journal of the Atmospheric Sciences*, 44(14):1775–1800, 1987.

M. McIntyre. On dynamics and transport near the polar mesopause in summer. *Journal of Geophysical Research*, 94:14617–14628, 1989.

M.E. McIntyre and T.N. Palmer. Breaking planetary waves in the stratosphere. *Nature(London)*, 305(5935):593–600, 1983.

C. McLandress and W.E. Ward. Tidal and gravity wave interactions and their influence on the large-scale dynamics of the middle atmosphere: Model results. *Journal of Geophysical Research*, 99(D4):8139–8155, 1994.

C. McLandress, W.E. Ward, V.I. Fomichev, K. Semeniuk, S.R. Beagley, N.A. McFarlane, and T.G. Shepherd. Large-scale dynamics of the mesosphere and lower thermosphere: An analysis using the extended Canadian Middle Atmosphere Model. *Journal of Geophysical Research*, 111(D17):D17111, 2006.

R.D. McPherson, K.H. Bergman, R.E. Kistler, G.E. Rasch, and D.S. Gordon. The NMC operational global data assimilation system. *Monthly Weather Review*, 107(11):1445–1461, 1979.

C.J. Mertens. Retrieval of mesospheric and lower thermospheric kinetic temperature from measurements of carbon dioxide 15 micrometer earth limb emission under non-LTE conditions. *Geophysical Research Letters*, 28(7), 2001.

S. Miyahara and D.H. Wu. Effects of solar tides on the zonal mean circulation in the lower thermosphere: solstice condition. *Journal of atmospheric and terrestrial physics*, 51(7-8):635–647, 1989.

M.G. Mlynczak and S. Solomon. A detailed evaluation of the heating efficiency in the middle atmosphere. *Journal of Geophysical Research*, 98(D6):10517, 1993.

J.J. Morcrette, L. Smith, and Y. Fouquart. Pressure and temperature dependence of the absorption in longwave radiation parametrizations. *Beiträge zur Physik der Atmosphäre*, 59(4):455–469, 1986.

T. Nakamura, T. Tsuda, S. Fukao, A.H. Manson, C.E. Meek, R.A. Vincent, and I.M. Reid. Mesospheric gravity waves at Saskatoon(52 N), Kyoto(35 N), and Adelaide(35 S). *Journal of geophysical research*, 101(D 3):7005–7012, 1996.

W.A. Norton and J. Thuburn. Sensitivity of mesospheric mean flow, planetary waves, and tides to strength of gravity wave drag. *Journal of Geophysical Research*, 104(D24):30897, 1999.

A. Orr, P. Bechtold, J. Scinocca, M. Ern, and M. Janiskova. Technical Memorandum No. 625: Improved middle atmosphere climate and forecasts in the ECMWF model through a non-orographic gravity wave drag parametrization. Technical report, ECMWF, 2010.

T.N. Palmer, G.J. Shutts, and R. Swinbank. Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parmetrization. *Quarterly Journal of the Royal Meteorological Society*, 112(474):1001–1039, 1986.

M. Park, W.J. Randel, D.E. Kinnison, R.R. Garcia, and W. Choi. Seasonal variation of methane, water vapor, and nitrogen oxides near the tropopause: Satellite observations and model simulations. *J. Geophys. Res*, 109, 2004.

S. Pawson, K. Kodera, K. Hamilton, T.G. Shepherd, S.R. Beagley, B.A. Boville, J.D. Farrara, T.D.A. Fairlie, A. Kitoh, W.A. Lahoz, et al. The GCM–Reality In-

tercomparison Project for SPARC (GRIPS): Scientific Issues and Initial Results. *Bulletin of the American Meteorological Society*, 81(4):781–796, 2000.

D. Pendlebury, T.G. Shepherd, M. Pritchard, and C. McLandress. Normal mode Rossby waves and their effects on chemical composition in the late summer stratosphere. *Atmospheric Chemistry and Physics*, 8(7):1925–1935, 2008.

R.A. Plumb. Baroclinic Instability of the Summer Mesosphere: A Mechanism for the Quasi-Two-Day Wave?. *Journal of Atmospheric Sciences*, 40:262, 1983.

V.D. Pope, M.L. Gallani, P.R. Rowntree, and R.A. Stratton. The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics*, 16(2):123–146, 2000.

M. Pulido and J. Thuburn. Gravity wave drag estimation from global analyses using variational data assimilation principles. I: Theory and implementation. *Quarterly Journal of the Royal Meteorological Society*, 131(609):1821–1840, 2005.

M. Pulido and J. Thuburn. Gravity-wave drag estimation from global analyses using variational data assimilation principles. II: Case study. *Quarterly Journal of the Royal Meteorological Society*, 132(618):1527–1543, 2006.

M. Pulido and J. Thuburn. The Seasonal Cycle of Gravity Wave Drag in the Middle Atmosphere. *Journal of Climate*, 21(18):4664–4679, 2008.

V. Ramaswamy, M.L. Chanin, J. Angell, J. Barnett, D. Gaffen, M. Gelman, P. Keckhut, Y. Koshelkov, K. Labitzke, J.J.R. Lin, et al. Stratospheric temperature trends: Observations and model simulations. *Rev. Geophys*, 39(1):71–122, 2001.

W. Randel, P. Udelhofen, E. Fleming, M. Geller, M. Gelman, K. Hamilton, D. Karoly, D. Ortland, S. Pawson, R. Swinbank, et al. The SPARC intercomparison of middle-atmosphere climatologies. *Journal of Climate*, 17(5):986–1003, 2004.

W.J. Randel and F. Wu. A stratospheric ozone trends data set for global modeling studies. *Geophysical Research Letters*, 26(20):3089–3092, 1999.

F. Rawlins, S.P. Ballard, K.J. Bovis, A.M. Clayton, D. L.I., G.W. Inverarity, A.C. Lorenc, and T.J. Payne. The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 133(623 B):347–362, 2007.

A. Reale, F. Tilley, M. Ferguson, and A. Allegrino. NOAA operational sounding products for advanced TOVS. *International Journal of Remote Sensing*, 29(16): 4615–4651, 2008.

R.J. Reed. Zonal wind behavior in the equatorial stratosphere and lower mesosphere. *Journal of Geophysical Research*, 71(18):4223–4233, 1966.

I.M. Reid and R.A. Vincent. Measurements of mesospheric gravity wave momentum fluxes and mean flow accelerations at Adelaide, Australia. *Journal of atmospheric and terrestrial physics*, 49(5):443–460, 1987.

M.M. Reinecker et al. The GEOS-5 data assimilation system: A documentation of GEOS-5.0. *NASA Tech. Rep. 104606 V*, 2007.

E.E. Remsberg, L.L. Gordley, B.T. Marshall, R.E. Thompson, J. Burton, P. Bhatt, V.L. Harvey, G. Lingenfelser, and M. Natarajan. The Nimbus 7 LIMS version 6 radiance conditioning and temperature retrieval methods and results. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 86(4):395–424, 2004.

E.E. Remsberg, B.T. Marshall, M. Garcia-Comas, D. Krueger, G.S. Lingenfelser, F.J. Martin-Torres, M.G. Mlynczak, J.M. Russell III, A.K. Smith, Y. Zhao, et al. Assessment of the quality of the Version 1.07 temperature versus pressure profiles of the middle atmosphere from TIMED/SABER. *J. Geophys. Res*, 113, 2008.

P.J.W. Roberts and D.R. Webster. Turbulent diffusion. *Environmental Fluid Mechanics-Theories and Application*, 2002.

C. Rocken, K. Ying-Hwa, W.S. Schreiner, D. Hunt, S. Sokolovskiy, and C. McCormick. COSMIC system description. *Terrestrial Atmospheric and Oceanic Sciences*, 11(1):21–52, 2000.

C. Rodgers. Characterization and error analysis of profiles retrieved from remote sounding measurements. *Journal of Geophysical Research*, 95(D5):5587–5595, 1990.

C.D. Rodgers. Inverse methods for atmospheric sounding–Theory and practise, edited by: Taylor, F.W., 2000.

C.D. Rodgers, R.J. Wells, R.G. Grainger, and F.W. Taylor. Improved stratospheric and mesospheric sounder validation: General approach and in-flight radiometric calibration. *Journal of Geophysical Research*, 101(D6), 1996.

K.H. Rosenlof. Seasonal cycle of the residual mean meridional circulation in the stratosphere. *J. Geophys. Res*, 100(5):173–5, 1995.

K.H. Rosenlof. Summer hemisphere differences in temperature and transport in the lower stratosphere. *Journal of Geophysical Research-Atmospheres*, 101(D14), 1996.

J.M. Russell et al. The Halogen Occultation Experiment. *J. Geophys. Res.*, 98(10): 777–797, 1993.

J.M. Russell III, M.G. Mlynczak, L.L. Gordley, J.J. Tansock Jr, and R.W. Esplin. Overview of the SABER experiment and preliminary calibration results. In *Proceedings of SPIE*, volume 3756, page 277. SPIE, 1999.

S.I. Sargoytchev, S. Brown, B.H. Solheim, Y.M. Cho, G.G. Shepherd, and M.J. López-González. Spectral airglow temperature imager (SATI): a ground-based instrument for the monitoring of mesosphere temperature. *Applied optics*, 43 (30):5712–5721, 2004.

F. Sassi, B.A. Boville, D. Kinnison, and R.R. Garcia. The effects of interactive ozone chemistry on simulations of the middle atmosphere. *atmosphere*, 32 (L07811), 2004.

R. Saunders, P. Brunel, F. Chevallier, G. Deblonde, S.J. English, M. Matricardi, and P. Rayer. RTTOV-7 science and validation report. *NWP Forecasting Research Tech. Rep*, 387, 2002a.

R.W. Saunders, M. Matricardi, P. Brunel, S.J. English, and G. Deblonde. RTTOV-7: a satellite radiance simulator for the new millennium. *Proceedings of ITSC-XII, Lorne, Australia*, 28, 2002b.

A.A. Scaife, N. Butchart, C.D. Warner, D. Stainforth, W. Norton, and J. Austin. Realistic quasi-biennial oscillations in a simulation of the global climate. *Geophysical Research Letters*, 27(21):3481–3484, 2000.

A.A. Scaife, N. Butchart, C.D. Warner, and R. Swinbank. Impact of a Spectral Gravity Wave Parameterization on the Stratosphere in the Met Office Unified Model. *Journal of the Atmospheric Sciences*, 59(9):1473–1489, 2002.

M.R. Schoeberl, D.F. Strobel, and J.P. Apruzese. A numerical model of gravity wave breaking and stress in the mesosphere. *Journal of Geophysical Research*, 88 (C9):5249–5259, 1983.

M. J. Schwartz et al. Polarized radiative transfer for the Zeeman-split oxygen lines in the EOS MLS forward model. *IEEE Transcations on geoscience and remote sensing*, 44(5):1181–1190, 2006.

M.J. Schwartz et al. Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements. *J. Geophys. Res*, 2007.

J.F. Scinocca. An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. *Journal of the Atmospheric Sciences*, 60(4): 667–682, 2003.

T.G. Shepherd. The middle atmosphere. *Journal of Atmospheric and Solar-Terrestrial Physics*, 62(17-18):1587–1601, 2000.

T.G. Shepherd and T. Shaw. The angular momentum constraint on climate sensitivity and downward influence in the middle atmosphere. *Journal of the Atmospheric Sciences*, 61(23):2899–2908, 2004.

K. Shine. Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation. *Quarterly Journal of the Royal Meteorological Society*, 115(486), 1989.

K.P. Shine. The middle atmosphere in the absence of dynamical heat fluxes. *Quarterly Journal of the Royal Meteorological Society*, 113(476), 1987.

G.J. Shutts. personal communication, 2010.

D. Siméoni, C. Singer, and G. Chalon. Infrared atmospheric sounding interferometer. *Acta Astronautica*, 40(2-8):113–118, 1997.

A. Simmons, M. Hortal, G. Kelly, A. McNally, A. Untch, and S. Uppala. ECMWF Analyses and Forecasts of Stratospheric Winter Polar Vortex Breakup: September 2002 in the Southern Hemisphere and Related Events. *Journal of the Atmospheric Sciences*, 62(3):668–689, 2005.

A. Simmons, S. Uppala, D. Dee, and S. Kobayashi. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter*, 110:25–35, 2007.

A.J. Simmons and D.M. Burridge. An Energy and Angular-Momentum Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinates. *Monthly Weather Review*, 109(4):758–766, 1981.

A.K. Smith. Stationary planetary waves in upper mesospheric winds. *Journal of the Atmospheric Sciences*, 54(16):2129–2145, 1997.

A.K. Smith. The origin of stationary planetary waves in the upper mesosphere. *Journal of the Atmospheric Sciences*, 60(24):3033–3041, 2003.

A.K. Smith and L.V. Lyjak. An observational estimate of gravity wave drag from the momentum balance in the middle atmosphere. *Journal of Geophysical Research*, 90:2233–2242, 1985.

S.A. Smith, D.C. Fritts, and T.E. Vanzandt. Evidence for a saturated spectrum of atmospheric gravity waves. *Journal of the Atmospheric Sciences*, 44(10):1404–1410, 1987.

W.L. Smith, H.M. Woolf, C.M. Hayden, D.Q. Wark, and L.M. McMillin. The TIROS-N operational vertical sounder. *Bull. Amer. Meteor. Soc*, 60(10):1177–1187, 1979.

S. Solomon, J.T. Kiehl, R.R. Garcia, and W. Grose. Tracer transport by the diabatic circulation deduced from satellite observations. *Journal of the Atmospheric Sciences*, 43(15):1603–1617, 1986.

A. Staniforth, A. White, N. Wood, J. Thurburn, M. Zerroukat, and E. Cordero. UMDP no 15: Joy of U.M. 6.0 Model Formulation. Technical report, Met Office, 2004.

G.G. Stokes. On the theory of oscillating waves. *Trans. Camb. Phil. Soc*, 8:441, 1847.

R.A. Stratton. A high resolution AMIP integration using the Hadley Centre model HadAM2b. *Climate Dynamics*, 15(1):9–28, 1999.

D.F. Strobel, J.P. Apruzese, and M.R. Schoeberl. Energy balance constraints on gravity wave induced eddy diffusion in the mesosphere and lower thermosphere. *Journal of Geophysical Research*, 90(D7):13067, 1985.

Z. Sun and L. Rikus. Improved application of exponential sum fitting transmissions to inhomogeneous atmosphere. *Journal of Geophysical Research*, 104(D6): 6291–6303, 1999.

R. Swinbank and D.A. Ortland. Compilation of wind data for the Upper Atmosphere Research Satellite (UARS) reference atmosphere project. *Journal of Geophysical Research-Atmospheres*, 108(D19):4615, 2003.

R. Swinbank, W.A. Lahoz, A. O'neill, C.S. Douglas, A. Heaps, and D. Podd. Middle atmosphere variability in the UK Meteorological Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, 124(549):1485–1525, 1998.

B.D. Tapley and C. Reigber. The GRACE mission: status and future plans. In *EGS XXVII General Assembly, Nice, 21-26 April 2002*, volume 27, page 4819, 2002.

J-C. Thelen. Informal Met Office Report: 'Radiative aspects relating to the middle atmosphere model'. Technical report, Met Office, 2006.

Y. Tomikawa, K. Sato, S. Watanabe, Y. Kawatani, K. Miyazaki, and M. Takahashi. Wintertime temperature maximum at the subtropical stratopause in a T213L256 GCM. *J. Geophys. Res*, 113, 2008.

T. Tsuda and K. Hocke. Vertical Wave Number Spectrum of Temperature Fluctuations in the Stratosphere using GPS Occultation Data. *J Meteorol*, 80(4B): 925–938, 2002.

T. Tsuda, Y. Murayama, M. Yamamoto, S. Kato, and S. Fukao. Seasonal variation of momentum flux in the mesosphere observed with the MU radar. *Geophysical Research Letters*, 17(6):725–728, 1990.

T. Tsuda, M. Nishida, C. Rocken, and Ware R.H. A global morphology of gravity wave activity in the stratosphere revealed by the GPS occulation data (GPS/MET). *J. Geophys. Res*, 105(D6):7257–7273, 2000.

T. Tsuda, M.V. Ratnam, P.T. May, M.J. Alexander, R.A. Vincent, and A. MacKinnon. Characteristics of gravity waves with short vertical wavelengths observed with radiosonde and GPS occultation during DAWEX (Darwin Area Wave Experiment). *J. Geophys. Res*, 109, 2004.

V.M. Tunbridge and N.J. Mitchell. The two-day wave in the Antarctic and Arctic mesosphere and lower thermosphere. *Atmos. Chem. Phys*, 9:6377–6388, 2009.

TE VanZandt. A universal spectrum of buoyancy waves in the atmosphere. *Geophysical Research Letters*, 9(5), 1982.

R.A. Vincent and D.C. Fritts. A Climatology of Gravity Wave Motions in the Mesopause Region at Adelaide, Australia. *Journal of the Atmospheric Sciences*, 44(4):748–760, 1987.

C.D. Warner and M.E. McIntyre. On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere. *Journal of the Atmospheric Sciences*, 53(22):3213–3235, 1996.

C.D. Warner and M.E. McIntyre. Toward an ultra-simple spectral gravity wave parameterization for general circulation models. *Earth Planets and Space*, 51 (7/8):475–484, 1999.

C.D. Warner and M.E. McIntyre. An Ultrasimple Spectral Parameterization for Nonorographic Gravity Waves. *Journal of the Atmospheric Sciences*, 58(14):1837–1857, 2001.

S. Watanabe, Y. Kawatani, Y. Tomikawa, K. Miyazaki, M. Takahashi, and K. Sato. General aspects of a T213L256 middle atmosphere general circulation model. *J. Geophys. Res*, 113, 2008.

S. Watanabe, Y. Tomikawa, K. Sato, Y. Kawatani, K. Miyazaki, and M. Takahashi. Simulation of the eastward 4-day wave in the Antarctic winter mesosphere using a gravity wave resolving general circulation model. *J. Geophys. Res*, 114, 2009.

J.W. Waters, K.F. Kunzi, R.L. Pettyjohn, R.K.L. Poon, and D.H. Staelin. Remote sensing of atmospheric temperature profiles with the Nimbus 5 microwave spectrometer. *Journal of the Atmospheric Sciences*, 32(10):1953–1969, 1975.

J. Weinstock. Energy dissipation rates of turbulence in the stable free atmosphere. *Journal of the Atmospheric Sciences*, 38(4):880–883, 1981.

A. White, B.J. Hoskins, I. Roulstone, and A. Staniforth. Consistent approximate models of the global atmosphere: shallow, deep, hydrostatic, quasi-hydrostatic and non-hydrostatic. *Quarterly Journal of the Royal Meteorological Society*, 131 (609):2081–2107, 2005.

A.A. White and R.A. Bromley. A new set of dynamical equations for use in numerical weather prediction and global climate models. *Meteorological Office, Met O 13 Branch Memo*, 1988.

J.A. Whiteway, T.J. Duck, D.P. Donovan, J.C. Bird, S.R. Pal, and A.I. Carswell. Measurements of gravity wave activity within and around the Arctic stratospheric vortex. *Geophysical Research Letters*, 24(11):1387–1390, 1997.

J. Wickert, T. Schmidt, G. Beyerle, R. König, C. Reigber, and N. Jakowski. The Radio Occultation Experiment aboard CHAMP: Operational Data Analysis and Validation of Vertical Atmospheric Profiles. *Journal of the Meteorological Society of Japan*, 82(1B):381–395, 2004.

C.A. Wilson and R. Swinbank. Gravity Wave Drag, Meteorological Office Unified Model Doc Paper no.21. Technical report, Met Office, 1990.

1986. WMO. WMO Global Ozone Research and Monitoring Project. *World Meteorological Organisation, Geneva*, 1986.

R.F. Woodman, B.B. Balsley, F. Aquino, L. Flores, E. Vazquez, M. Sarango, M.M. Huaman, and H. Soldi. First observations of polar mesosphere summer echoes in Antarctica. *Journal of geophysical research*, 104(A10):22577, 1999.

D.L. Wu, W.G. Read, Z. Shippony, T. Leblanc, T.J. Duck, D.A. Ortland, R.J. Sica, P.S. Argall, J. Oberheide, A. Hauchecorne, et al. Mesospheric temperature from UARS MLS: retrieval and validation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 65(2):245–267, 2003.

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

X. Zhang, J.M. Forbes, M.E. Hagan, J.M. Russell III, S.E. Palo, C.J. Mertens, and M.G. Mlynczak. Monthly tidal temperatures 20–120 km from TIMED/SABER. *J. Geophys. Res*, 111, 2006.

W. Zhong, S.M. Osprey, L.J. Gray, and J.D Haigh. Influence of the prescribed solar spectrum on calculations of atmospheric temperature. *Geophysical Research Letters*, 35(22):L22813, 2008.