Dynamics of the Tropical Tropopause and Lower Stratosphere

Julia Claire Tindall

A thesis submitted for the degree of Doctor of Philosophy

November 2003
'Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged'

Julia Tindall
Abstract

Dynamical variability in the tropical upper troposphere/lower stratosphere (UTLS) is important for troposphere-stratosphere exchange and the driving of the quasi-biennial oscillation (QBO). However the relative importance of motions with different timescales is not fully understood. In this thesis dynamical variability in the tropical UTLS is investigated using the ECMWF 15 year reanalysis (ERA15) dataset.

The first long term, tropical, climatology of idealised, linear, equatorial waves, with period greater than 1 day, is produced using a novel method. The method calculates the seasonally averaged activity of several wave types. It is based on wavenumber-frequency spectral analysis and linear wave theory. At 70hPa and 100hPa there is usually an annual cycle in the activity of each wave. At 50hPa the activity of each wave is correlated with the QBO in a way qualitatively consistent with the QBO’s modulation of each wave’s vertical group velocity. The observed wave climatologies are insufficient to drive the QBO or explain the tropical UTLS temperature variance on subseasonal timescales. This suggests that much of the subseasonal variability in the tropical UTLS is inconsistent with idealised linear waves of period greater than 1 day.

Finally the annual temperature cycle in the tropical UTLS is realistically modelled using the thermodynamic equation. At 50hPa the model results compare better than ERA15 data with radiosonde observations, although the model is less accurate at 100hPa. At 70hPa the model produces a good representation of the ERA15 temperature and its annual cycle. Approximately half of the annual cycle can be explained by the residual upwelling alone; annual variations in the radiative timescale and the radiative equilibrium temperature are needed to explain the rest. Interannual variations in 70hPa temperature can be explained by interannual variations in the poleward mass flux between 70hPa and 10hPa.
Acknowledgements

Thanks to my supervisors, John Thuburn and Ellie Highwood, who have spent many hours helping me to understand the finer points of my work and reading thesis chapters. The project has been so much easier because of their knowledge, guidance, and enthusiasm. Thanks to members of my thesis committee and tropopause club for useful discussions. Thanks also to everyone at the Department of Meteorology, Reading, for contributing to such a pleasant working environment.

Thanks to my parents and my sister, Angela, for all their encouragement and for providing many necessary distractions (particularly in the final year Angela). Thanks to John H. for being there when times were tough, for understanding, and for dragging me away from the computer for regular breaks.

I couldn’t have done it without you all.
# Contents

1 Introduction                                      1

1.1 The tropical tropopause and lower stratosphere 1

1.2 Transport in the region of the tropical tropopause 4

1.3 Theories of stratospheric dehydration 5

1.4 Equatorial waves in the lower stratosphere 7

1.4.1 Kelvin waves 8

1.4.2 Rossby-gravity waves 11

1.4.3 Inertio-gravity waves 13

1.4.4 Equatorial Rossby waves 15

1.5 The Quasi-Biennial Oscillation (QBO) 17

1.5.1 Effects of the QBO 17

1.5.2 The QBO and equatorial waves 19

1.6 The annual temperature cycle in the tropical lower stratosphere 20

1.7 Thesis questions 23

2 Detecting equatorial wave activity in re-analysis datasets 27

2.1 Introduction 27

2.2 The ECMWF-ERA dataset 28
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1</td>
<td>Waves in the ECMWF-ERA dataset</td>
<td>29</td>
</tr>
<tr>
<td>2.3</td>
<td>Background to wave detection method</td>
<td>35</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Linear wave theory</td>
<td>35</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Wavenumber-frequency spectral analysis</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>Results from wavenumber-frequency spectral analysis</td>
<td>43</td>
</tr>
<tr>
<td>2.5</td>
<td>The wave detection method</td>
<td>46</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Reducing computation time</td>
<td>47</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Estimates of waves in a wavenumber-frequency bin</td>
<td>48</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Constraints on method</td>
<td>50</td>
</tr>
<tr>
<td>2.6</td>
<td>Discussion of the accuracy of the method</td>
<td>51</td>
</tr>
<tr>
<td>2.6.1</td>
<td>A mathematical issue and its effect on the wave detection method</td>
<td>53</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Interaction between symmetric and antisymmetric waves</td>
<td>55</td>
</tr>
<tr>
<td>2.7</td>
<td>Summary</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>A climatology of equatorial waves in the lower stratosphere</td>
<td>59</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>59</td>
</tr>
<tr>
<td>3.2</td>
<td>Wave climatologies</td>
<td>60</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Kelvin waves</td>
<td>61</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Rossby-gravity waves</td>
<td>65</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Inertio-gravity waves</td>
<td>68</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Equatorial Rossby waves</td>
<td>75</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>3.3</td>
<td>Explaining $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$</td>
<td>81</td>
</tr>
<tr>
<td>3.4</td>
<td>Variance in Dynamical Fields</td>
<td>84</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Converting between $[u' \frac{\partial T'}{\partial x}], [v' \frac{\partial T'}{\partial y}]$ and $[\partial T^2], [u'^2], [v'^2]$</td>
<td>85</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Results</td>
<td>89</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Additional variance</td>
<td>92</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Considering $n = 1$ waves</td>
<td>98</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
<td>102</td>
</tr>
<tr>
<td>4</td>
<td>The interaction between equatorial waves and the QBO</td>
<td>104</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>104</td>
</tr>
<tr>
<td>4.2</td>
<td>Variability of the waves in different QBO phases</td>
<td>105</td>
</tr>
<tr>
<td>4.3</td>
<td>Forcing of the QBO by equatorial waves</td>
<td>109</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Background</td>
<td>109</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Calculating momentum flux and zonal acceleration</td>
<td>112</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Momentum Flux Results</td>
<td>115</td>
</tr>
<tr>
<td>4.3.4</td>
<td>QBO Results</td>
<td>121</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary</td>
<td>125</td>
</tr>
<tr>
<td>5</td>
<td>The annual temperature cycle near the tropical tropopause</td>
<td>127</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>127</td>
</tr>
<tr>
<td>5.2</td>
<td>The thermodynamic model</td>
<td>130</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Inputs common to all the models</td>
<td>132</td>
</tr>
</tbody>
</table>
5.3 Model [A] .............................................................. 137
  5.3.1 Description ..................................................... 137
  5.3.2 Mean results at 70hPa from model [A] ......................... 139
5.4 Model [B] .............................................................. 145
  5.4.1 Description ..................................................... 145
  5.4.2 Mean results at 70hPa from model [B] ......................... 147
5.5 Model [C] .............................................................. 150
  5.5.1 Description ..................................................... 150
  5.5.2 Mean results at 70hPa from model [C] ......................... 152
5.6 Considering the radiative timescale .............................. 157
5.7 How should the temperature be modelled? ....................... 163
5.8 Using the TEM thermodynamic equation (Model [C]) ........ 164
  5.8.1 Modelling individual years at 70hPa ......................... 164
5.9 Modelling the temperature at other levels ...................... 168
  5.9.1 Radiative parameters .......................................... 168
  5.9.2 The ‘average’ annual cycle at 100hPa ....................... 170
  5.9.3 Interannual variations in temperature at 100hPa .......... 173
  5.9.4 The ‘average’ annual cycle at 50hPa ....................... 175
5.10 Summary ............................................................. 179

6 Conclusions ........................................................... 183
6.1 The equatorial wave climatology ........................................... 184
6.2 Equatorial waves and the QBO .............................................. 186
6.3 Modelling the annual temperature cycle in the tropical lower stratosphere ... 188
6.4 Future work ................................................................. 190
6.4.1 Improvements in linear wave theory ................................. 190
6.4.2 Detecting individual equatorial waves .............................. 191
6.4.3 Enhancements to the wave detection method ...................... 192
6.4.4 Improvements to observations ....................................... 193
6.4.5 Enhancements to the prediction of the annual temperature cycle ....... 194
CHAPTER 1

Introduction

1.1 The tropical tropopause and lower stratosphere

The troposphere is the layer of the atmosphere that extends from the earth’s surface to an altitude of between 8 and 16km. It is characterised by convective overturning which is driven by surface heating. In the troposphere temperature generally decreases with altitude with a global mean lapse rate of approximately 6.5K km$^{-1}$. Above the troposphere, ozone heating causes the temperature to increase with altitude in the atmospheric layer referred to as the stratosphere. The upper boundary of the stratosphere is approximately 50km above the surface of the earth. In the stratosphere there are only weak vertical motions, and this layer is dominated by radiative processes. Chemically the stratosphere is characterised by much higher ozone and much lower water vapour concentrations than the troposphere. The interface between the troposphere and the stratosphere is referred to as the tropopause. The tropopause is highest in the tropics ($\approx 16$km) and lowest in polar regions ($\approx 8$km); however this thesis will only be concerned with the tropical tropopause.

A location of the tropical tropopause can be found by considering where a chosen atmospheric component changes from its tropospheric values to its stratospheric values. There are several definitions of the tropical tropopause, which are discussed by Highwood and Hoskins (1998). One of the most useful definitions is the cold point tropopause as this is thought to be important in determining how much water vapour enters the stratosphere (e.g. Selkirk, 1993). The cold point tropopause is located at the temperature minimum that necessarily exists between the troposphere (where temperature is decreasing with altitude) and the stratosphere (where temperature is increasing with altitude). Another thermal tropopause definition is the lapse rate tropopause. This is defined as the lowest level at which the lapse rate decreases to 2K/km or less provided that the average lapse rate between this level and all higher levels within 2km does not exceed
2K/km (WMO, 1957). However the lapse rate tropopause has little physical relevance in the tropics (Highwood and Hoskins, 1998). It has been suggested that a meaningful location for the tropical tropopause is at the top of convective heating (e.g. Forster et al., 1997); this is because the troposphere is convectively driven while the stratosphere is dominated by radiative processes.

The cold point tropopause, the lapse rate tropopause and the top of convection are some of the locations that have been suggested to divide the stratosphere from the troposphere. However these locations do not all occur at the same altitude. Thuburn and Craig (1997) examined the zonal-mean tropical tropopause height in a GCM and found that the lapse rate tropopause was around 16.5km (100hPa), the cold point tropopause was around 18km (80hPa) and the peak of the winter branch of the Hadley circulation was around 13.5km (160hPa). It is therefore thought to be more meaningful to think of the tropical tropopause as a transition zone rather than a two-dimensional surface (e.g. Atticks and Robinson 1983, Highwood and Hoskins 1998, Folkins et al. 1999). In the tropopause transition zone the coverage by deep convection decreases nearly exponentially with increasing altitude (Gettelman et al., 2002b). This transition zone is often termed the tropical tropopause layer, but its vertical limits have not been universally agreed. For example Gettelman and Forster (2002) defined the tropical tropopause layer as extending from the level of maximum temperature lapse rate (approx 12km (190hPa)) to the cold point tropopause (approx 17km (90hPa)) while Sherwood and Dessler (2001) defined the tropical tropopause layer as the region bounded by the 150hPa and 50hPa surfaces. The tropical tropopause layer (defined as the region between the main convective outflow and the cold point) has also been referred to as the tropical substratosphere (Thuburn and Craig, 2002). This is because it is more similar to the stratosphere than the troposphere in the fact that the radiative timescale is shorter than the convective timescale so that radiation dominates in setting the temperature structure. Figure 1.1 shows a schematic of all the different levels relevant to the tropical tropopause that have been discussed in this section.

This chapter will review some important processes in the tropical tropopause layer and the tropical lower stratosphere. Section 1.2 will discuss transport in the tropical tropopause layer in general and section 1.3 will focus on the transport of water vapour between the troposphere and
Figure 1.1  Black line is temperature profile for January, obtained from radiosonde data at Truk (7.5N,151.9E), averaged over 1951-1997. The blue, green and purple horizontal lines relate to the temperature profile and show the cold point tropopause, the lapse rate tropopause and the level of maximum lapse rate respectively. The red line shows the level of the peak of the winter branch of the Hadley circulation as modelled by Thuburn and Craig (1997). The region marked ‘TTL GF02’ shows the tropical tropopause layer corresponding to this temperature profile as defined by Gettelman et al. (2002b). The region marked ‘TTL SD01’ shows the tropical tropopause layer as used by Sherwood and Dessler (2001).

the stratosphere. Sections 1.4, 1.5 and 1.6 will discuss dynamical variability near the tropical tropopause and will focus on equatorial waves, the Quasi-Biennial Oscillation (QBO) and the annual temperature cycle respectively. Section 1.7 will highlight the important questions that this thesis will consider, in order to increase the understanding of the tropical tropopause and tropical lower stratospheric region.
1.2 Transport in the region of the tropical tropopause

The tropopause is not impermeable. Air crosses the tropopause and mass is exchanged between the troposphere and the stratosphere. Tropospheric mass usually enters the stratosphere through upwelling in the tropics and stratospheric air is returned to the troposphere at higher latitudes. This characterises a stratospheric circulation, which was first suggested by Brewer (1949), and though the existence of the stratospheric circulation is now in little doubt some questions remain as to its exact details.

The tropical upwelling branch of the stratospheric circulation transports natural and anthropogenic chemical constituents and water vapour, through the tropical tropopause layer, which can alter the radiative and chemical balance of the stratosphere. The water vapour in the stratosphere is thought to depend crucially on the properties of the tropical tropopause layer (e.g. Gettelman et al., 2002a) particularly the temperature at the cold point tropopause. Indeed it was the observations of stratospheric water vapour over Southern England which initially led Brewer (1949) to postulate a stratospheric circulation. The water vapour mixing ratio in the stratosphere was so low that Brewer hypothesised that the stratospheric air must have passed through the tropical tropopause as only there were temperatures cold enough to ‘freeze dry’ the air to a stratospheric water vapour mixing ratio consistent with observations. The tropical tropopause is around 25K colder than the mid latitude tropopause, and so air that had entered the stratosphere through the tropical tropopause would contain less water vapour than air that had entered the stratosphere through the mid latitude tropopause. However it is still not clear whether stratospheric water vapour concentrations are consistent with the saturation mixing ratio at the tropical cold point tropopause. Newell and Gould-Stewart (1981) found that the tropical cold point was too warm to explain the observed stratospheric water vapour mixing ratio, whilst Dessler (1998) found that stratospheric water vapour was consistent with the tropical cold point temperature.

On long timescales, it appears that a direct relationship between tropical tropopause temperature and water vapour entering the stratosphere does not hold. For example in recent years strato-
spheric water vapour has been increasing (Oltmans et al., 2000) which is opposite to what is expected given that the tropical tropopause has been cooling (Zhou et al., 2001). It appears that the water vapour entering the stratosphere may have a more complex dependence on the processes in the tropical tropopause layer, than simply freeze drying at the cold point. Forster and Shine (1999) suggested that recent increases in stratospheric water vapour may be capable of explaining as much of the observed stratospheric cooling as ozone loss. However the physical processes which determine the water vapour content of air entering the stratosphere are not fully understood. They are discussed in section 1.3.

1.3 Theories of stratospheric dehydration

Recently much research has been concerned with understanding the physical processes responsible for setting the values of stratospheric water vapour. Following their observations that the stratosphere was drier than the mean saturation mixing ratio at the cold point tropopause, Newell and Gould-Stewart (1981) hypothesised that air would only cross the tropopause at certain times and locations. These times and locations were those where the tropopause was cold enough to ‘freeze dry’ the air to the observed water vapour content, and included regions such as the western Pacific in the northern hemisphere winter. However Newell and Gould-Stewart (1981) did not explain why their regional ascent would not carry ice particles into the stratosphere that would subsequently evaporate and hydrate the air.

An alternative theory, (Holton and Gettelman, 2001) is that air is dehydrated as it passes horizontally through regions of coldest temperatures. This theory is similar to what was suggested by Newell and Gould-Stewart (1981) in that certain regions are responsible for dehydrating the air, but it does not impose the strict constraints on the locations of cross tropopause transport. The typical ratio of horizontal velocity to vertical velocity in the tropical tropopause region is about $10^4$, and so it is likely that most air to enter the stratosphere will have been subject to quasi-horizontal transport through a region which is colder than the average cold point tropopause.

Stratospheric dehydration was first attributed to convective processes by Danielsen (1982). His
hypothesis involved processes operating near the tropopause on the huge anvils of cumulonimbus clouds. Radiative heating at the base of the cumulonimbus would lead to subsaturated air while radiative cooling at the top of the anvil would lead to supersaturated air. The supersaturated air at the top of the anvil would lead to a rapid growth of ice crystals, which could then precipitate, dehydrating the anvil region. A slightly different mechanism of convective dehydration was suggested by Teitelbaum et al. (2000). Here air was assumed to be dried to low mixing ratios as it ascended within deep convective towers. During periods of intense convective phenomena, there would be an uplift of the tropopause which would lead to a reduction in tropopause temperature. If most of the air entered the stratosphere during these periods of intense convection, when the tropopause was particularly cold and high, then the stratosphere would be drier than the saturation mixing ratio at the average cold point.

Potter and Holton (1995) considered the effects of convectively generated gravity waves on stratospheric water vapour. Their model results show that wave clouds would form in the cold phase of stratospheric gravity waves and that ice fall out from these wave clouds would remove water from the stratosphere. In contrast the results of Jensen et al. (1996) show that gravity waves, alone, can have very little impact on the water vapour concentrations near the tropopause. This is because these waves have such short periods that ice particles in the cold phase of the wave do not have time to grow large enough to fall out before they evaporate. Instead Jensen et al. (1996) suggested that dehydration would occur as air is slowly lifted by large scale motions as ice crystals would have more time to grow large enough to fall out of the stratosphere. Jensen et al. (1996) suggested that large scale ascent could be caused by air flowing over large scale convective systems or large scale wave motions.

The role of Kelvin waves in stratospheric dehydration has been discussed by Fujiwara et al. (2001) and Fujiwara and Takahashi (2001). They noted that water-vapour would preferentially enter the stratosphere during the upward displacement phase of a Kelvin wave, but that this upward motion would cause adiabatic cooling, which would further freeze dry the air. In support of this Boehm and Verlinde (2000) found that subvisible cirrus occurred exclusively at cold phases of a Kelvin wave. Because Kelvin waves occur over much longer timescales than gravity
waves, the fall out of ice crystals from Kelvin wave induced subvisible cirrus is more likely.

Recently a combination of mechanisms have been used to model stratospheric dehydration. Jensen et al. (2001) found that while large scale ascent could dehydrate the air, the dehydration was more effective if temperature oscillations associated with gravity wave motions were included. Gettelman et al. (2002a) used a simple model to reproduce many of the features of the tropical tropopause layer. They found that horizontal advection of air through the cold trap region could account for some of the dehydration, but they found that temperature variations due to gravity waves were necessary to reduce the modelled water vapour further. Vömel et al. (2002) considered the relative roles of convective dehydration, slow ascent dehydration and large-scale wave-driven dehydration. Different combinations of these mechanisms were found to be important in different regions of the tropics and at different times.

Although modelling studies have used a variety of mechanisms to dehydrate the stratosphere, a lack of observations means that the relative roles of dehydration mechanisms in the real atmosphere cannot be quantified. As noted by Jensen et al. (1996) “... the relative importance of motions with different time scales must ultimately depend on what scales dominate the variability in this region”. With this in mind, we now review some of the processes which lead to variability near the tropical tropopause. Since this section has shown that wave motions are likely to be important in determining values of water vapour to enter the stratosphere, the following section will review the current level of understanding about equatorial waves.

1.4 Equatorial waves in the lower stratosphere

In addition to their likely role in stratospheric dehydration, equatorial waves are responsible for driving the Quasi-Biennial Oscillation (QBO) (see section 1.5) and are thought to be the dominant short timescale variability in the lower stratosphere. Equatorial waves are predicted by linear wave theory (Matsuno, 1966) which is reviewed in chapter 2. In the lower stratosphere there are a broad spectrum of equatorially trapped waves. Previous observations of these waves and their
likely forcing mechanisms are considered in this section. Kelvin waves, Rossby-gravity waves, inertio-gravity waves and equatorial Rossby waves will be discussed in turn.

### 1.4.1 Kelvin waves

The simplest type of equatorial wave is the eastward propagating Kelvin wave. This wave can modulate temperature, geopotential height, vertical velocity and zonal velocity. It does not affect meridional velocity. Kelvin wave signals in all fields do not change sign with latitude, and so, at a given height, time and longitude, meridional structure is dependent on distance from the equator only. Kelvin waves show downward phase propagation (consistent with upward flux of energy). It is thought that these waves originate in the troposphere and are associated with latent heat release through deep convection (e.g. Holton, 1972); however these waves may also be forced from midlatitudes (Hoskins and Yang 2000, Straub and Kiladis 2003).

Kelvin waves were first observed in the equatorial lower stratosphere by Wallace and Kausky (1968a). They used six months of radiosonde data from three tropical stations and found fluctuations in the zonal wind with an average period of around 15 days. Corresponding fluctuations in the temperature field were also found which led the zonal wind fluctuations by 1/4 of a cycle. These wave motions showed downwards phase propagation and did not appear to involve the meridional wind. All of these characteristics resembled the Kelvin waves predicted by the wave theory to be described in chapter 2. These waves were found to have a typical phase speed of 25m/s.

Disturbances consistent with Kelvin waves have also been found at higher altitudes. Hirota (1978) observed Kelvin waves, with a phase speed of around 70m/s, in the middle and upper stratosphere, while Salby et al. (1984) observed Kelvin waves, with a phase speed of around 120m/s, in the upper stratosphere. The Kelvin waves of Hirota (1978) and Salby et al. (1984) have been referred to as ‘fast Kelvin waves’ and ‘ultra fast Kelvin waves’ respectively. For clarity the Kelvin waves of Wallace and Kausky have been referred to as ‘slow Kelvin waves’. It has also been suggested (Parker, 1973) that there are Kelvin waves confined to the region
between 150hPa and 70hPa which have a period of between 25 and 40 days. In keeping with the traditional naming standards these waves are now known as ‘ultra slow Kelvin waves’.

Ultra-slow Kelvin waves have not been the subject of much interest since their initial discovery. Perhaps this is because Parker’s initial study contained some unexpected results. Firstly these waves were detected very infrequently (they were detectable at 4 time periods totalling only 20 months in 10 years of radiosonde data), and secondly observations of their vertical wavelengths (8-27km) disagreed with the value predicted from wave theory (3.7km). More recently Hamilton (1997) observed a disturbance with the characteristics of an ultra-slow Kelvin wave in radiosonde data from TOGA-COARE (Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response Experiment). This disturbance had period 30-40 days and vertical wavelength 3.8km, which are consistent with the theory of ultra-slow Kelvin waves. However a wave of such short wavelength would be difficult to detect without high resolution data such as Hamilton used. Despite seeing a clear indication of ultra-slow Kelvin waves, Hamilton did point out that the limited time (120 days), and spatial extent (7 radiosonde stations), of his study meant there was no conclusive evidence that this variation was either a Kelvin wave or a global phenomenon. Further evidence of ultra-slow Kelvin waves was found by Canziani (1999) in satellite data from the Cryogenic Limb Array Etalon Spectrometer (CLAES) instrument on the Upper Air Research Satellite (UARS). Again this was a fairly short study (16 months), and there were some gaps in the data. In his initial study, Parker (1973) did not find that ultra-slow Kelvin waves followed any well defined pattern of occurrence, and there is insufficient data from subsequent studies to be able to confirm or deny this. It certainly seems as though further study of the ultra slow Kelvin wave on a longer timescale would be useful.

Slow Kelvin waves have been studied in much more detail than ultra slow Kelvin waves, and since they have been found in the tropical lower stratosphere the rest of this review will be concerned with the properties of slow Kelvin waves only. Because slow Kelvin waves affect the temperature structure near the tropopause, they can modulate the temperature and height of the cold point. Tsuda et al. (1994a) found a 20 day period Kelvin wave in one month of radiosonde data at Indonesia. This wave caused the cold point tropopause to vary sinusoidally in the range
186-194K. It also caused a slow downward progression of the cold point tropopause height from 17km to 15.5km over approximately 15 days, followed by a 1.5km sudden height increase. This downward motion of the tropopause followed by a discontinuous jump was also seen in the atmospheric stability parameter ($N^2$). Similar results to those of Tsuda et al. (1994a) were found in five months of radiosonde data at Indonesia by Shimizu and Tsuda (1997).

Day to day variations in the cold point tropopause temperature were attributed to a mixture of Kelvin waves and tropospheric convection by Shimizu and Tsuda (2000). They also noted that day to day variations in the cold point tropopause temperature seemed to have larger amplitude than seasonal or interannual variations. Such large day to day fluctuations in tropopause temperature are likely to be very important in troposphere-stratosphere exchange and highlight the importance of Kelvin waves in this region.

The effect of Kelvin waves on stratosphere-troposphere exchange is perhaps the most important issue regarding the presence of Kelvin waves around the tropopause. The potential role of Kelvin waves as a mechanism for dehydrating the stratosphere has already been discussed, but Kelvin waves are also thought to be able to transport stratospheric ozone into the troposphere. Fujiwara et al. (1998) investigated an enhancement of ozone in the tropical upper troposphere where the 40nmol/mol isoline of ozone moved 5.0km downwards to 12.8km while the tropopause height varied by only 1.6km. They attributed this ozone transport to the downward motions associated with the Kelvin wave, and the Madden-Julian Oscillation, that were observed at this time. Fujiwara et al. (1998) also suggested that air mixing due to Kelvin wave breaking at the tropopause would lead to additional stratosphere-troposphere exchange.

Evidence of slow Kelvin wave activity has also been found in the temperature data obtained with the CLAES instrument on the UARS by Shiotani et al. (1997) and Canziani (1999). These studies were able to show that Kelvin waves were not longitudinally confined to certain regions. Kelvin wave signals have also been found in Outgoing Longwave Radiation (OLR; e.g Wheeler and Kiladis 1999), stratospheric water vapour (Mote et al., 1998), total column ozone (Ziemke and Stanford, 1994a), and the stratospheric trace species $O_3$, CFC-11, HNO$_3$, N$_2$O, and CH$_4$ (Smith
et al., 2002).

1.4.2 Rossby-gravity waves

Matsuno (1966) considered quasi horizontal wave motions in the equatorial area. It was found that wave motions with non-zero meridional velocity could exist for $n \geq 0$, where $n$ represents the number of zeros of the wave solution in meridional velocity at a fixed height and longitude. The $n = 0$ wave solution was found to have an eastward propagating component similar to inertio-gravity waves (section 1.4.3) and a westward propagating component similar to equatorial Rossby waves (section 1.4.4). Following the notation of Wheeler and Kiladis (1999), we refer to the westward propagating component of this wave solution as the Rossby-gravity (RG) wave and the eastward propagating component of this wave solution as the eastwards inertio-gravity $n = 0$ (EIG0) wave. The EIG0 wave has received relatively little attention in the literature and is considered along with higher order inertio-gravity waves in section 1.4.3. The Rossby-gravity wave has been the subject of more intense study, due to a large number of observations and its potential role in driving the easterly phase of the QBO (section 1.5). It is reviewed in this section.

The Rossby-gravity wave affects all dynamical fields. In zonal wind and temperature it appears as the lowest order antisymmetric mode. In meridional wind it appears as the lowest order symmetric mode. Like the Kelvin wave, observations of the Rossby-gravity wave show downward phase propagation (consistent with upward flux of energy).

The Rossby-gravity wave was first observed by Yanai and Marayanna (1966), where it was found to have period 4-5 days, zonal wavenumber 4, vertical wavelength 4-8km and phase speed 23m/s. More recent observations of Rossby-gravity waves have found its characteristics to be consistent with those detected by Yanai and Marayanna (1966). Disturbances have been attributed to Rossby-gravity waves that have been found in rawinsonde data (Maruyama 1991, Dunkerton 1991a, Wikle et al. 1997), OLR data (Hendon and Liebmann 1991, Wheeler and Kiladis 1999), data from the Indian Mesosphere-Stratosphere-Troposphere (MST) radar (Sasi and Deepa, 2001) and model data (Boville and Randel 1992, Hayashi and Golder 1994). In the upper stratosphere
Rossby-gravity waves have been found to have quite different characteristics (Randel et al., 1990) with vertical wavelengths 13-16km, period 2-3 days and zonal phase velocities around 200m/s.

The downward phase propagation of observed Rossby-gravity waves suggests that they are excited in the troposphere. However, Dunkerton (1993) found that, in general, stratospheric Rossby-gravity waves were incoherent with tropospheric oscillations. He suggested that this could be explained by selective transmission of higher frequency waves.

Initially there were two main hypotheses to explain the forcing of Rossby-gravity waves. The first is that they are due to forcing from midlatitudes. This was proposed by Mak (1969) who was able to model a tropical Rossby-gravity wave by applying stochastic forcing to the lateral boundaries (30N and 30S) of a two layer model. However as noted by Itoh and Ghil (1988), “it appears unlikely that the zonal wavenumber or frequency of the equatorially trapped wave are directly influenced by lateral forcing away from the equator where the waves are exponentially evanescent”. The second hypothesis is called the wave-CISK theory (Hayashi, 1970). Here free Rossby-gravity modes are generated by the unstable interaction between waves and cumulus convection. Hayashi found these unstable waves showed realistic horizontal structure and vertical wavelength, but this theory failed because it would cause gravity waves with large wavenumbers and high frequencies to grow faster than Kelvin or Rossby-gravity waves in contrast to what is observed.

Though there are issues concerning both of the above hypotheses there is also evidence in support of each. In support of the forcing from midlatitudes hypothesis, Zangvil and Yanai (1980) found that Rossby-gravity waves were closely related to the energy flux of zonal wavenumbers 3-6 from midlatitudes. In support of the wave-CISK hypothesis, it is noted that many observations (e.g. Hendon and Liebmann 1991, Takayabu 1994a) have found that tropical cloud cover has spectral peaks corresponding to Rossby-gravity waves, suggesting convective forcing. Itoh and Ghil (1988) combined the two hypotheses to force realistic Rossby-gravity waves in a model. They found that the basic mechanism for the generation of Rossby-gravity waves is nonlinear.
wave CISK, but that antisymmetric forcing from midlatitudes is required for the Rossby-gravity waves to become noticeable against the dominant symmetric modes.

More recently Magaña and Yanai (1995) used forecast winds from the European Centre for Medium-Range Weather Forecasts (ECMWF), and found that Rossby-gravity waves can be caused by midlatitude disturbances propagating directly into the tropics. This could only happen if equatorial winds were westerly or weak easterly (Charney, 1969), and if there were extratropical disturbances with spatial and temporal scales similar to Rossby-gravity waves. The conditions most suitable for Rossby-gravity waves to be forced in this way are found in the northern hemisphere (NH) summer Eastern Pacific, in agreement with the temporal and spatial distribution of Rossby-gravity waves.

Hoskins and Yang (2000) found that a Gill type model could generate equatorial waves from high latitude forcing even in the case of strong easterly winds. It therefore appears that high latitude forcing can be directly projected onto equatorial wave modes, and the equatorial response does not rely on the propagation of Rossby waves into the tropics. Equatorial waves were attained using both vorticity forcing and heating in higher latitudes. Though the Kelvin wave response to eastward forcing was more dramatic, the Rossby-gravity wave response to westward forcing was still significant.

Though most of the observations of Rossby-gravity waves have used the meridional wind field, this wave has also been found in stratospheric ozone (Stanford and Ziemke, 1993).

**1.4.3 Inertio-gravity waves**

The previous subsection dealt with the Rossby Gravity wave which is the westward propagating equatorial wave solution with $n = 0$ (see chapter 2). The eastward propagating waves with $n = 0$ are more similar to inertio-gravity waves than Rossby waves (Matsuno, 1966), and will be considered here. Also considered are westward and eastward propagating inertio-gravity waves with $n \geq 1$, where, again, $n$ represents the number of zeros of the wave solution in the meridional
wind field at a fixed height, time and longitude. For all inertio-gravity waves the number of zeros at a fixed height, time and longitude in other dynamical fields is $n + 1$. For example the eastward propagating inertio-gravity $n = 0$ (EIG0) wave will be characterised by symmetric exponential decay from the equator in the meridional wind field. In other fields the EIG0 wave will be antisymmetric and will change sign at the equator.

Observations of inertio-gravity waves are limited, despite the fact that they are expected to provide a significant contribution to the driving of the QBO (Takahashi et al. 1997, Dunkerton 1997, Alexander and Holton 1997). The first study to attribute disturbances to inertio-gravity waves was by Cadet and Teitelbaum (1979), who analysed twenty-nine wind profile soundings (13-16 July 1974) when the QBO was westerly; however there was insufficient data in their study to determine the direction of wave propagation (Sato et al., 1994). Eastward propagating inertio-gravity waves have been found for the easterly regime of the QBO in radiosonde data (Tsuda et al. 1994b, Shimizu and Tsuda 1997) and radar data at Jicamarca, Peru (11.95S, 76.87W; Riggin et al. 1995). The first study to look at signals of inertio-gravity waves over several stations (Wada et al., 1999) also occurred in the easterly QBO regime; this study was able to attribute disturbances with a period of $\approx 2$ days to eastward propagating inertio-gravity (EIG) waves with $n = 0$ and $n = 1$.

Other observations of inertio-gravity waves in the lower stratosphere have often been made at single locations where it is impossible to determine the direction of propagation of the wave (e.g. Karoly et al., 1996). However it was found that there is a net upward flux of eastward momentum associated with waves of an approximately 2 day period (Vincent and Alexander 2000, Maruyama 1994) implying that eastward propagating waves are more active than westward propagating waves at inertio-gravity wave frequencies. However there is a significant cancellation between the momentum flux attributed to eastward and westward propagating waves (Sato and Dunkerton, 1997).

Although there is a lack of observations of westward propagating inertio-gravity (WIG) waves in the lower stratosphere, evidence of $n = 1$ and $n = 2$ WIG waves have been found in OLR data.
(Wheeler and Kiladis 1999, Wheeler et al. 2000) and cloud fields (Takayabu 1994a, Takayabu 1994b). These studies suggest that there are WIG waves coupled to convection, but they have not found any evidence of convectively coupled EIG waves for \( n > 0 \).

It is interesting that only westward propagating inertio-gravity waves have been found in convection, while only eastward propagating inertio-gravity waves have been observed in the stratosphere. If these studies are fully representative of inertio-gravity wave activity, then it implies that the troposphere and stratosphere are independent as regards these wave types. In reality this is unlikely to be the case because the stratospheric inertio-gravity waves are likely to have been forced in the troposphere. It is expected that a more complete set of observations would lead to greater coherence between stratospheric and tropospheric inertio-gravity waves. For example all of the eastward propagating stratospheric inertio-gravity waves have been found in the easterly phase of the QBO. No direction of wave propagation has been determined for inertio-gravity waves in the westerly phase of the QBO. To increase the understanding of these waves, the first step would be to find disturbances consistent with them over a longer time period and larger spatial area than has previously been considered. This would suggest avenues that should be explored in order to determine the role of this wave type and their forcing mechanisms.

1.4.4 Equatorial Rossby waves

The equatorial Rossby (ER) wave is a strictly westward propagating low frequency wave. Like the inertio-gravity wave, it affects all dynamical fields. Its latitudinal structure in meridional wind can have any number, \( n \ (n \geq 1) \), of zeros providing that the number of zeros in other dynamical fields is \( n + 1 \). Therefore this wave can be symmetric or antisymmetric about the equator.

There have been relatively few studies that have found evidence of ER waves in the lower stratosphere. Perhaps this is due to a difficulty in separating signals that are due to the ER wave from signals that are due to the Rossby-gravity wave. For example many studies (e.g. Maruyama 1991, Dunkerton 1991a) have attributed the meridional wind disturbances with a 3-6 day period,
that have been found in rawinsonde data, to Rossby-gravity waves. However, some of these
disturbances near the equator could be due to ER waves with symmetric meridional velocity, and
some of these disturbances which are further away from the equator could be due to ER waves
with antisymmetric meridional velocity. However ER waves are likely to be of lower frequency
than the 3-6 day disturbances found in these studies.

Although there has been a lack of observations of equatorial Rossby waves in ra-
diosonde/rawinsonde data, convectively coupled equatorial Rossby waves have been found
in other datasets. For example, Yang et al. (2003) used 8 days of ECMWF Re-Analysis 15
(ERA15) data and Cloud Archive User Service (CLAUS) data for July 1992. They found
evidence of ER waves, with $n = 1$ and $n = 2$, over the Eastern Pacific and evidence of ER
$n = 1$ waves over the Atlantic. The equatorial Rossby waves were detected at 850hPa and
200hPa, although the signal was much weaker at 200hPa. Evidence of convectively coupled ER
$n = 1$ waves was also found by Wheeler et al. (2000), who used satellite observed OLR and
National Centers for Environmental Prediction (NCEP) reanalysis dynamical fields. There was
no evidence that this data contained ER $n = 2$ waves. The ER $n = 1$ waves found by Wheeler
et al. typically had zonal wavenumber 4-5 and were largely confined to the troposphere but could
be detected to a certain extent in the lower stratosphere. A study of ER waves appropriate to the
stratosphere is that of Ziemke and Stanford (1994b). They considered disturbances of westward
wavenumber 3-6 in total column ozone between the years 1979 and 1991, and found that some
of these disturbances were consistent with ER waves of various meridional structure.

The likely forcing mechanism of ER waves will not be described in this section. However it will
be noted that much of the discussion about the forcing of the Rossby-gravity wave (section 1.4.2)
is also appropriate to the ER wave.

All of the waves discussed in this section interact strongly with the Quasi-Biennial Oscillation
(QBO), which is discussed in the next section.
Chapter 1 Introduction

1.5 The Quasi-Biennial Oscillation (QBO)

On timescales longer than a year the dominant variability in the tropical lower stratosphere is the Quasi-Biennial Oscillation (often referred to as the QBO). The QBO is an oscillation of downward propagating easterly and westerly wind regimes with a ‘quasi-biennial’ period averaging approximately 28 months. This oscillation is shown in figure 1.2 which is a time-height section of monthly mean zonal wind at Singapore. The QBO dominates the variability of the equatorial stratosphere, but can influence the global stratosphere, the mesosphere and even the troposphere. Baldwin et al. (2001) and references therein, provide details of the global effects of the QBO but they are summarised here for completeness.

1.5.1 Effects of the QBO

There is a QBO signal in temperature which is in thermal wind balance with the QBO signal in zonal winds. The QBO signal in temperature leads to a meridional circulation associated with the QBO, which was first discussed by Reed (1964). The QBO induced meridional circulation is such that the westerly winds are associated with equatorward mass flux while easterly winds are associated with poleward mass flux. This circulation is largely confined to the tropics but it can have a significant effect on transport of long lived trace species in the stratosphere (see Baldwin et al., 2001).

The QBO is seen in the winds and temperatures of the equatorial lower stratosphere. However it also has non local effects because it can modulate the tropical and extratropical waves that are propagating through the atmosphere. Midlatitude planetary (Rossby) waves generally propagate upwards and equatorwards, however they cannot propagate through easterly winds. This means that when the QBO is westerly these waves can penetrate the tropics, but when the QBO is easterly they cannot. As a result, during the easterly phase of the QBO these planetary waves are trapped in the winter hemisphere extratropics and wave activity in this region is stronger. This was observed by Holton and Tan (1980) who found that in November-December the amplitude of planetary wavenumber 1 was nearly 40% stronger in the easterly phase of the QBO than it was in the westerly phase of the QBO. The additional extratropical wave activity that occurs during
the easterly phase of the QBO can have important consequences. For example, the increased midlatitude planetary waves in the easterly QBO phase will lead to a weaker than normal polar night jet and warmer than normal polar temperature (see Dunkerton and Baldwin, 1991, and references therein), thus increasing the possibility of a sudden stratospheric warming.

The idea that the surface responds to the QBO can be traced back to Holton and Tan (1980). They showed that in January, the northern hemisphere 1000hPa geopotential, at high latitudes, was largest when the QBO was easterly. This means that the polar vortex at the surface would be weaker in the easterly QBO phase. The extratropical surface response to the QBO is discussed in more detail by Baldwin and Dunkerton (1999), who show that statistically the QBO can act to modulate the Arctic Oscillation from the middle stratosphere to the surface in December and January.

The role of the QBO on the tropical troposphere is less clear. Many authors (e.g. Gray, 1984) have noted a connection between the stratospheric QBO and the interannual variability of Atlantic hurricane activity. This connection is such that hurricanes occur more often in seasons when the 50hPa QBO is westerly. However a convincing physical explanation for why this should be the case has not been given. The stratospheric QBO can also affect the mesosphere, through selective filtering of tropical waves as they propagate upwards through the stratosphere.

---

**Figure 1.2** Time-height section of monthly mean zonal wind at Singapore 1993-98. Westerly winds are shaded, contour interval is 10m/s. (Obtained from http://tao.atmos.washington.edu/data_sets/qbo/)
1.5.2 The QBO and equatorial waves

Equatorial waves (section 1.4) and the QBO are mutually dependent. The phase of the QBO determines which waves propagate through the stratosphere, while the momentum fluxes associated with the waves can accelerate the mean zonal flow and drive the QBO. Equatorial waves can only propagate vertically through certain background conditions. Here we consider the background conditions necessary for the vertical propagation of waves.

Andrews et al. (1987) show that a Kelvin wave can only exist in regions where $u(z) < c$, where $c$ is the speed of the Kelvin wave and $u(z)$ is the background zonal wind at height $z$. Now suppose there is some vertical wind shear (which is characteristic of the QBO - figure 1.2) and suppose that there is some height, $z_{\text{crit}}$, such that $u(z_{\text{crit}}) = c$. The Kelvin wave will be able to propagate below $z_{\text{crit}}$, if $u(z) < c$, but, as the Kelvin wave approaches $z_{\text{crit}}$, its vertical wavelength and group velocity will become very small and the wave will become more and more susceptible to dissipation. $z_{\text{crit}}$ is called a critical level and the wave will have fully dissipated before this critical level is reached. A critical level can exist for any wave, because all waves will be able to propagate through certain background conditions only. Generally a westward propagating linear wave can only exist in regions where $u(z) > c_w$. An eastward propagating wave can only exist in regions where $u(z) < c_e$. $c_w$ and $c_e$ denote the intrinsic speeds of a westward and eastward propagating wave respectively. The critical levels, through which a wave cannot propagate, are therefore dependent on the type and speed of the wave considered.

As each wave type approaches its critical level, it will dissipate and its momentum will be absorbed by the mean zonal flow. This momentum will accelerate the mean zonal flow in the same direction as $c - u(z)$. For example, as a Kelvin wave approaches its critical level, it will dissipate and will accelerate the flow eastwards, because a kelvin wave must have $c - \pi > 0$. This will lead to a descent of the eastward wind regime. As a Rossby-gravity wave approaches its critical level, it will dissipate and will accelerate the flow westwards, because Rossby-gravity waves have $c - \pi < 0$. This will lead to a descent of the westward wind regime. The fact that a wave’s momentum will be absorbed just below its critical level can qualitatively explain the descending
pattern of alternating easterlies and westerlies that characterise the QBO (Lindzen and Holton 1968, Holton and Lindzen 1972).

1.6 The annual temperature cycle in the tropical lower stratosphere

So far this chapter has considered some processes which affect variability in the tropical lower stratosphere on a day to day basis and on an interannual basis. In addition there is significant variability caused by an annual cycle in tropical lower stratospheric temperatures. This will now be discussed.

The annual temperature cycle in the tropical lower stratosphere consists of maximum temperatures in July-August and minimum temperatures in January-February. The phase of the annual temperature cycle is similar throughout the tropics but the amplitude is slightly larger in the Northern Hemisphere. The amplitude of the annual cycle is largest at the 80hPa level (the level of the winter cold point tropopause) and diminishes rapidly below and more slowly above (Reed and Vleck, 1969). Radiosonde data considered by Reed and Vleck (1969) suggest that the amplitude of the annual temperature cycle is approximately 2.5K at 100hPa and approximately 4.5K at 80hPa.

The annual cycle in lower stratospheric temperatures leads to an annual cycle in the temperature, height and pressure of the cold point tropopause. Since the amount of water vapour entering the stratosphere may be determined by the coldest temperature an air parcel had experienced, the annual cycle in tropopause temperatures implies an annual cycle in water vapour entering the stratosphere. This implication was investigated quantitatively by Mote et al. (1996). They found that air in the stratosphere was indeed ‘marked’ by the temperature of the cold point tropopause through which the air had passed. This meant that air which had crossed the cold point in the NH summer contained more water vapour than air which had crossed the cold point in the NH winter. This imprint of tropical tropopause temperature on stratospheric water vapour was termed an ‘atmospheric tape recorder’. Thus it appears that the annual temperature cycle near the tropical tropopause is not just interesting for its own sake but has a direct impact on the amount of water
vapour to enter the stratosphere.

Reed and Vleck (1969) were the first to suggest that the annual temperature cycle in the lower stratosphere was due to an annual cycle in adiabatic cooling, caused by an annual cycle in tropical upwelling. They attributed the tropical upwelling to the Hadley circulation penetrating into the stratosphere; however tropical upwelling is now thought to be due to some other mechanism. More recently, the annual cycle in tropical upwelling (and hence tropical lower stratospheric temperatures) has been attributed to an annual cycle in the strength of the stratospheric circulation. Yulaeva et al. (1994) showed that when the tropical lower stratosphere was coldest the extratropical lower stratosphere was warmest and when the tropical lower stratosphere was warmest the extratropical lower stratosphere was coldest. In fact there was practically no annual cycle in lower stratospheric temperatures averaged over the globe. This was explained by the tropical upwelling branch of the stratospheric circulation causing tropical temperatures to fall below their radiative values, whilst descending air in the extratropics increased extratropical temperatures above their radiative values. So it appears that the processes that determine the strength of the stratospheric circulation also determine the temperature near the tropical tropopause.

Throughout most of the stratosphere, contours of angular momentum vertically span extratropical regions because of the dominance of the earth’s rotation (see Haynes et al., 1991, figure 1). This means that the extratropics are inaccessible to tropical air unless there is some force which can reduce angular momentum and hence drive air polewards. Therefore the stratospheric circulation can clearly not be driven by tropical convection. Instead, the required angular momentum reducing force is provided by the upward propagation and breaking of Rossby and gravity waves in midlatitudes (e.g. Haynes et al. 1991, Dunkerton 1991a). The fact that the stratospheric circulation is driven by these waves is now in little doubt (Holton et al., 1995) and because of this the stratospheric circulation is often referred to as the ‘wave driven circulation’.

The wave driven circulation was investigated by Haynes et al. (1991), who, based on the results of Dickinson (1968), showed that on long timescales the extratropical mass flux across a
given isentropic surface was controlled solely by the momentum forcing distribution above that surface. In other words, the forcing acts to produce the circulation via ‘downwards control’. The downward control principle was used to quantitatively calculate the residual mean meridional circulation for the lower stratosphere by Rosenlof and Holton (1993). Their calculation involved solving the Transformed Eulerian Mean (TEM) (Andrews and McIntyre, 1976) steady state zonal momentum and continuity equations, with momentum forcing provided by the Eliassen-Palm flux divergence. Rosenlof and Holton (1993) found that this method was able to realistically calculate the residual circulation for solstice seasons, but because the steady state assumption was included, it was inadequate for equinox seasons. Seol and Yamazaki (1999) calculated the residual stratospheric circulation in a similar way to Rosenlof and Holton (1993), but did not assume a steady state form of the momentum equation. They were able to obtain a more reasonable residual circulation for equinox seasons. The results of Rosenlof and Holton (1993) and Seol and Yamazaki (1999) were able to show that the downward control principle and the wave-driven circulation could produce a quantitatively accurate circulation. However, because angular momentum contours do not vertically span tropical regions, their results could only be applied to the extratropics. Upwelling in the tropics could be calculated because it is equal to the downwelling at higher latitudes, but no information about the distribution of tropical upwelling could be obtained from these studies. Eluszkiewicz et al. (1996, 1997) calculated the residual circulation using diabatic heating rates, and so were able to obtain information about the circulation in the tropics. It was found that maximum upwelling occurred in the summer hemisphere near the equator. This observation cannot be explained by the wave-driven circulation because the dominant wave forcing will occur in the ‘surf zone’ (McIntyre and Palmer, 1983), which extends from the edge of the winter polar vortex to the winter subtropics. In the absence of any other processes the wave-driven circulation would consist of upwelling in the winter subtropics and downwelling at higher latitudes which is not as is observed.

Because there is only a weak background angular momentum gradient in the tropics, nonlinear effects can become important. Dunkerton (1989) was able to model a nonlinear stratospheric ‘Hadley’ regime driven by antisymmetric thermal equilibria in the absence of internal sources or sinks of angular momentum. This Hadley regime consisted of a single mean meridional cell in the tropics, equatorial easterlies and strong winter westerlies. Model results of Plumb and
Eluszkiewicz (1999), Semeniuk and Shepherd (2001) and Tung and Kinnersley (2001) suggest that there is a middle atmosphere Hadley cell which is driven by diabatic heating, but is only able to extend into midlatitudes because of the extratropical wave driving. Therefore the observed stratospheric circulation can be explained by a combination of tropical heating and extratropical wave driving.

1.7 Thesis questions

It has been seen that dynamical processes near the tropical tropopause and in the lower stratosphere have wide reaching implications for the climate system as a whole. The tropical tropopause is a gateway to the stratosphere for water vapour and anthropogenic ozone destroying chemicals, while the QBO in the tropical lower stratosphere can affect the global stratosphere, the mesosphere and even the troposphere.

Despite the importance of processes in the tropical tropopause and lower stratosphere, the exact mechanisms operating in this region are not fully understood. It is not yet clear which physical processes are responsible for dehydrating air at the tropopause or which waves are most important for driving the QBO. The relative roles of diabatic heating and extratropical wave driving, in the stratospheric circulation and the annual cycle in lower stratospheric temperatures, are only just starting to emerge.

In order to increase understanding of this region this thesis will consider three main questions. These questions are concerned with dynamical variability of the region on a variety of timescales and are detailed below.

Question 1: What is the climatology of idealised linear equatorial waves in the tropical tropopause/lower stratosphere region?
Some lower stratospheric equatorial waves (Kelvin, Rossby-gravity) have been extensively studied while others (inertio-gravity, equatorial Rossby) have received very little attention. Yet a long term, tropics wide, climatology does not exist for any type of equatorial wave. This does not mean that such a climatology would not be valuable. A climatology of equatorial waves would help understand the role of different waves in driving the QBO and would also lead to a more accurate understanding of how equatorial waves affect stratosphere-troposphere exchange. Perhaps the reason that no tropics wide climatology of equatorial waves exists is because there is no ideal dataset from which such a climatology could be produced. Such a dataset would ideally provide information at regularly distributed points throughout the tropics (to ensure that all locations were equally represented) and the horizontal, vertical and temporal resolution of the dataset must be sufficient for the waves to be detected.

In chapter 2 a method is derived that will produce a climatology of idealised linear equatorial waves in a dataset such as is described above. The method further requires that the dataset contains the fields of temperature, zonal wind and meridional wind. A dataset which fulfils all the requirements of the derived method is the ERA15 dataset (discussed in chapter 2). A climatology of idealised linear waves could be produced using ERA15, provided that one was willing to accept the models ‘first guess’ in data sparse regions and the relatively coarse vertical resolution near the tropical tropopause. Chapter 2 shows that it is reasonable to produce a climatology of idealised linear equatorial waves from ERA15 data, and that such a climatology could be cautiously applied to the real atmosphere.

The method of chapter 2 is used in chapter 3 to detect idealised linear wave activity in ERA15. Chapter 3 considers how the activity of each wave type varies seasonally and interannually. It also considers the wavenumbers and frequencies at which each wave type exists. Since equatorial waves are thought to be the dominant, short timescale, variability in the lower stratosphere, chapter 3 will also investigate the extent to which the linear wave activity found in ERA15 is able to explain the variance in temperature, zonal wind, and meridional wind that ERA15 contains.

**Question 2: How do these waves interact with the QBO?**
The global importance of the lower stratospheric QBO has already been discussed. The QBO exists because of an interaction between equatorial waves and the mean zonal flow (section 1.5.2) although the relative roles of different wave types in driving the QBO has not been fully determined. It is therefore logical to investigate the extent to which the wave climatologies of chapter 3 influence and are influenced by the QBO. This is the focus of chapter 4.

The extent to which each wave type should be affected by the vertical wind shear is first considered. This is compared with the wave climatology for each wave type to evaluate whether the wave climatologies are theoretically consistent with the background zonal winds. After this consistency check the relative importance of each wave type in driving the QBO is assessed. It is hoped that this will provide a useful reference for modelling studies which require an accurate representation of the QBO.

**Question 3: Can the annual cycle in lower stratospheric temperatures be quantitatively modelled?**

The dominant temperature variability in the tropical lower stratosphere occurs on an annual timescale, and is attributed to variations in the tropical upwelling (section 1.6). Chapter 5 will quantitatively assess the extent to which tropical upwelling can account for the annual temperature cycle averaged over $10^\circ N - 10^\circ S$. This will be done by modelling the temperature using a simple form of the thermodynamic equation which attributes temperature change to a combination of mean tropical upwelling and a simple estimate of radiative heating. Understanding, quantitatively, what determines the temperature near the tropical tropopause could be used in the future to determine how the temperature is likely to change in response to climate change, and could help predict future amounts of stratospheric water vapour.

The temperature near the tropical tropopause is relatively easy to measure, while the stratospheric circulation is not. Many estimates of the stratospheric circulation have been obtained which give similar but slightly different answers (e.g. Rosenlof and Holton 1993, Eluszkiewicz et al. 1996,
Seol and Yamazaki (1999). If the upwelling used in this study can reasonably explain the annual temperature cycle near the tropical tropopause, then confidence in the accuracy of the upwelling is increased. An accurate quantification of the mean tropical upwelling will be useful in studies of stratosphere-troposphere exchange.
CHAPTER 2

Detecting equatorial wave activity in re-analysis datasets

2.1 Introduction

Equatorial waves in the lower stratosphere are thought to drive the QBO (e.g Baldwin et al., 2001) and to be important in stratosphere-troposphere exchange (Fujiwara and Takahashi, 2001), yet a long term, tropics wide climatology of such waves does not exist. This chapter presents a new method of quantifying linear equatorial wave activity, which can be used with reanalysis datasets of the tropical lower stratosphere. The method is based on wavenumber-frequency spectral analysis but imposes new constraints based on linear wave theory to help quantify the waves.

The method, to be presented, differs from most previous ways of detecting waves, in that no individual wave events are found; instead wave activity is looked for in zonally and seasonally averaged fields. The method is designed to produce a long term climatology of equatorial wave activity in datasets that span the tropics and contain the dynamical fields of zonal and meridional winds and temperature.

The wave detection method described in this chapter will be used in chapter 3 to detect waves in ECMWF-ERA15 data. Hence, before the wave detection method is discussed, section 2.2 will describe the ERA15 data and will consider the extent to which any waves found in ERA15 are likely to represent waves in the real atmosphere. In section 2.3 the idealised linear wave theory and wavenumber-frequency spectral analysis that will be used to derive the wave detection method are presented. Section 2.4 shows results obtained when applying wavenumber-frequency spectral analysis to ERA15. In section 2.5 the new wave detection method (which will be used in chapter 3) is derived from the theory of section 2.3. The issues affecting the accuracy of the
wave detection method are discussed in section 2.6, and the ideas and results of this chapter are summarised in section 2.7.

2.2 The ECMWF-ERA dataset

The ECMWF-ERA15 dataset (described by Gibson et al., 1999) is used throughout this thesis. This dataset contains a variety of dynamical fields, including temperature, zonal wind and meridional wind. The data used was obtained from the British Atmospheric Data Centre (BADC) website (see BADC, 2002). It has horizontal resolution $2.5^\circ \times 2.5^\circ$, and vertical levels in the region of the tropopause at 150hPa, 100hPa, 70hPa, and 50hPa.

The ERA15 dataset was produced using a data assimilation scheme to optimally combine atmospheric observations with a forecast model. The observations that were assimilated include radiosonde data, satellite data, aircraft data and surface observations which were all acquired from the World Meteorological Organisation’s (WMO) Global Telecommunication System (GTS) (see WMO, 2003). Additional observations assimilated into ERA15 include satellite radiance data and cloud track winds (Gibson et al., 1999).

Although the ERA15 dataset contains observations from a variety of different sources, there are still regions where observational data are sparse. Information about the atmosphere in these data sparse regions is provided by a numerical forecast model. In the horizontal, the forecast model used a T106 spectral representation (corresponding to horizontal resolution of $1.7^\circ$) to represent upper air fields and compute horizontal derivatives. However non-linear adiabatic terms and the diabatic physical parameterisation were computed using a coarser, grid point representation. In the vertical the forecast model has 31 levels between the surface and 10hPa. In the region of the tropopause, model levels are located at 156hPa, 132hPa, 110hPa, 90hPa, 70hPa and 50hPa. It is noteworthy that the ERA15 data used in this thesis has a coarser horizontal and vertical resolution than is provided by the forecast model. Also the 100hPa level, that will be used, is between two model levels.
The data assimilation system, used to optimally combine the forecast model with the observations, is described on page 21 of Gibson et al. (1999). One aspect of the data assimilation scheme which is of particular relevance to this thesis is that it uses diabatic normal mode initialisation. This initialisation process (discussed by Wergen, 1987) adjusts the mass and wind fields in such a way as to suppress inertio-gravity waves in ERA15. This could have serious consequences for the climatology of equatorial waves that this thesis aims to produce. Fortunately, the diabatic normal mode initialisation used in ERA15 is only performed on the 5 deepest vertical modes of the ERA15 data. This means that the only waves to be suppressed by the normal mode initialisation are those with vertical wavelength greater than 12km, and most of these will be higher frequency than this thesis aims to detect.

The annual temperature cycle at 100hPa and 70hPa in ERA15 compares well with radiosonde data (Pawson and Fiorino, 1998a). The QBO below 30hPa is also well represented by ERA15, although wind extrema, vertical shear and temperature anomalies associated with the QBO are too small (Pawson and Fiorino, 1998b). Previous observations have found equatorial waves in ERA15 to be representative of the atmosphere. For example, Straub and Kiladis (2002) showed that a typical, Eastern Pacific, convectively coupled, Kelvin wave disturbance in ERA15 compared remarkably well with radiosonde data. Yang et al. (2003) found several types of convectively coupled equatorial wave (including Kelvin, Rossby-gravity, equatorial Rossby and eastwards inertio-gravity) in just 8 days of ERA15 data. The waves found by Yang et al. were consistent with cloud variability from the Cloud Archive User Service. However the representation of equatorial waves by ERA15 will be considered further in the next subsection.

### 2.2.1 Waves in the ECMWF-ERA dataset

It has already been mentioned that the normal mode initialisation in ERA15 is unlikely to have a large effect on the waves considered in this thesis; also waves previously detected in ERA15 have compared well with other datasets. Nonetheless waves in the tropical stratosphere may not be adequately represented by ERA15 because of the dataset’s relatively coarse vertical resolution and its reliance on the forecast model’s ‘first guess’ in data sparse regions. It is therefore desirable to further test the representation of waves by ERA15. To verify that the forecast model’s ‘first
Chapter 2 Detecting equatorial wave activity in re-analysis datasets

guess’ is providing a reasonable representation of waves we will first compare the ERA15
data close to radiosonde stations with ERA15 data which is spatially distant from a radiosonde
station. We will also compare ERA15 data with radiosonde data to verify that the assimilated
observations are accurate.

The fields that will be used to detect waves in ERA15 are zonal wind, $u$, meridional wind, $v$, and temperature, $T$. There are large variations in these fields over long timescales. For example
chapter 1 explains that the temperature field is dominated by the annual cycle near the tropopause
and the zonal wind field is dominated by the QBO near 50hPa. To obtain an estimate of the
variation that is due to equatorial waves, a 40 day running mean is subtracted from $u$, $v$ and $T$ at each latitude and longitude to give $u_w$, $v_w$ and $T_w$. The variance of $u_w$, $v_w$ and $T_w$ for each latitude-longitude gridbox at 70hPa for 1986 is shown in figures 2.1(a), 2.1(b) and 2.1(c) respectively. Figure 2.1(d) shows the locations of radiosonde stations which provided data that
was assimilated into ERA15.

Figure 2.1 shows that there is large tropical variability in the time variance of $u_w$, $v_w$ and $T_w$. However, it is noteworthy that there is no clear correlation between the variance in $u_w$, $v_w$ or $T_w$ and the distribution of radiosonde stations. This means that the disturbances of period less than 40 days (which could be due to equatorial waves) are not confined to areas close to a radiosonde
station, nor are they dominant in data sparse regions. While this analysis does not show that
disturbances in ERA15 are representative of waves in the atmosphere, it is important. It shows that the
amplitudes of disturbances in ERA15 are not dependent on distance from a radiosonde
station. The forecast models first guess therefore does not lead to a temporal variance in dynamical fields
that is inconsistent with the temporal variance in radiosonde data.

Figure 2.1 provides information about the variance due to short period disturbances in dynamical
fields. Yet further analysis is required to assess the validity of waves in the ERA15 data. This
is because the variance in figure 2.1 may not be fully due to equatorial waves, and also the
contribution of each wave type to the variance is unclear. For example, the variance at two
locations may be similar but it may be composed of disturbances of different frequencies.
Figure 2.1  a), b) and c) show the variance of temperature, zonal wind and meridional wind respectively at 70hPa for 1986. In each case a 40 day running mean was removed from each latitude-longitude before the variance was found. d) shows the distribution of radiosondes between 10N and 10S.
To extend this analysis to provide information about those disturbances that contribute to the variance, spectral power for 1986 at 70hPa is compared for three different data sources. These data sources are a) the radiosonde station at Truk (7.5N,151.9E), b) the ERA15 data for the gridbox containing Truk, and c) the ERA15 gridbox centred on 7.5N,120.0W. The ERA15 gridbox centred on 7.5N, 120.0W is at the same latitude as Truk but is spatially distant from a radiosonde station (see figure 2.1(d)).

A histogram showing the spectral power in each of the data sources for the fields temperature, zonal wind and meridional wind at 70hPa in 1986 is shown in figure 2.2. The red bars show the spectral power for the radiosonde data at Truk, the green bars show the spectral power for the ERA15 data at Truk and the blue bars show the spectral power for the ERA15 data at 7.5N, 120.0W. Each bar in the histograms shows the spectral power integrated over 0.05 cycles per day of frequency.

The spectral power of temperature in each frequency band is shown in figure 2.2(a). It is seen that the two ERA15 gridboxes have similar spectral power in each frequency band. This is despite the fact that ERA15 data near Truk should be heavily influenced by radiosonde observations, while ERA15 data at 7.5°N, 120.0°W may be dominated by the models ‘first guess’. The spectral power of the ERA15 temperature data is also similar to that of the radiosonde temperature data. However the ERA15 data is more ‘red’, with higher power at lower frequencies, while the radiosonde data is more ‘white’, with a more even spread of spectral power over the frequency range. It is hypothesised that this difference, between ERA15 temperature data and radiosonde temperature data, is due to the fact that the ERA15 data also uses observations from satellites, which may not resolve high frequency waves as well as the radiosonde data.

Figure 2.2(b) shows the spectral power in each frequency band for zonal wind. Here the three data types have remarkably similar spectral power in each frequency band. The only noticeable difference is that there is more spectral power in the 0-0.05 cycle per day frequency band in the ERA15 data at 7.5°N, 120.0°W than there is in either dataset at Truk. However this extra power only represents an increase of around 30% over the corresponding spectral power in radiosonde data at
Figure 2.2 Spectral power in 0.05 cycle per day frequency bands for 2.2(a) temperature, 2.2(b) zonal wind and 2.2(c) meridional wind in 1986 at 70hPa. The red bars show spectral power in radiosonde data at Truk, the green bars show spectral power in ERA15 data at Truk and the blue bars show spectral power in ERA15 data centred on 7.5N, 120.0W. The spectral harmonic corresponding to the annual cycle has been removed in each case for clarity.
Truk. Overall the differences, in the spectral power of the zonal wind field, between the ERA15 data at 7.5°N, 120°W and the data at Truk are small. This suggests that the model is able to represent atmospheric wave-like disturbances but does not generate large amounts of spurious waves.

The meridional wind field, contains the largest difference in spectral power between the data at Truk and the ERA15 data at 7.5°N, 120.0°W. At low to medium frequencies the spectral power at 7.5°N, 120.0°W is up to twice as large as the spectral power at Truk, although there is better agreement between the fields at higher frequencies. The high spectral power in meridional wind at low frequencies at 7.5°N, 120.0°W could be due to the ERA15 dataset generating spurious waves. However it could be an accurate representation of the atmosphere since most wave types are not evenly distributed over the tropics (Wheeler and Kiladis, 1999). It is difficult to say whether or not the extra power in meridional wind at 7.5°N, 120.0°W is accurate since we do not have observations at this longitude for comparison. It is seen from figure 2.1(c) that the variance in this field (or the sum of spectral power over all frequencies) is not unusually large at 7.5°N, 120.0°W, however it is larger than the variance at Truk, which would explain the extra spectral power. It is also seen that some locations with large variance in meridional wind are geographically close to a radiosonde station. This means that the spectral power in ERA15 data at 7.5°N, 120.0°W may be in better agreement with the spectral power at another radiosonde station.

The results of this section have not conclusively proved that waves in ERA15 are representative of waves in the atmosphere. Such a proof would require observations at every gridpoint and timestep for comparison, and these observations are unlikely to become available. Instead this section has considered the variance of dynamical fields throughout the tropical region, and has also considered the spectral power of dynamical fields at certain gridboxes. None of the results are contrary to what would occur if waves in ERA15 were accurate representations of waves in the real atmosphere.

Similar results to those shown here have been found for 100hPa and 50hPa and for other years in the ERA15 dataset. This suggests that looking for waves in the ERA15 data is at least a reasonable task. The properties of any waves that are detected can be compared and contrasted
Chapter 2  Detecting equatorial wave activity in re-analysis datasets

with waves found in previous studies, and further evidence about the validity of the waves in the
ERA15 data should emerge.

2.3 Background to wave detection method

The previous subsection showed it to be reasonable to derive a method to detect waves in ERA15.
Some background to the wave detection method is now explained. It has already been mentioned
that the method will combine linear wave theory with wavenumber-frequency spectral analysis.
Theory of these two areas, which is relevant to the wave detection method, is detailed in subsec-
tions 2.3.1 and 2.3.2 below.

2.3.1 Linear wave theory

There are many large-scale wave motions propagating through the tropical upper troposphere and
lower stratosphere. They were first observed by Yanai and Marayanna (1966) and Wallace and
Kausky (1968a), and are thought to be important in driving the QBO (see Baldwin et al., 2001, and
references therein) and stratosphere-troposphere exchange (Fujiwara and Takahashi, 2001). Since
these wave motions are largely confined to within around 15° of the equator they can be modelled
using the equatorial \( \beta \)-plane approximation. This involves substituting the coriolis parameter, \( f \)
(\( f = 2\Omega \sin \phi \), \( \Omega \) is the earth’s rotation rate, \( \phi \) is latitude) with \( \beta y \) (\( \beta = 2\Omega a^{-1} \), \( a \) is the earth’s
radius, \( y \) is distance northwards of the equator). The primitive equations are then linearised about
a basic flow, non conservative processes are neglected, and \( \overline{v} \) and \( \overline{\theta}_y \) are set to zero. This gives
the idealised linear wave equations that will be used in this thesis. These are:

\[
\frac{\partial u'}{\partial t} - \beta y v' + \frac{\partial \Phi'}{\partial x} = 0
\]  

(2.1a)
Equations (2.1a) and (2.1b) represent momentum balance in the zonal and meridional directions respectively, (2.1c) represents continuity of mass and (2.1d) is obtained by combining hydrostatic balance and the thermodynamic relation. Primes represent departures from the mean, $z$ is the log-pressure vertical co-ordinate (as used in Andrews et al., 1987), $u$, $v$ and $w$ are the zonal, meridional and vertical components of velocity, $t$ is time, $\Phi$ is geopotential height, $\rho_0$ is density and $N$ is the buoyancy frequency. These terms are as defined in more details in Andrews et al. (1987).

Solutions to (2.1) are periodic in $x$ and $t$, because (2.1) are linear and all coefficients are independent of $x$ and $t$. The solutions can then be written as:

$$
\begin{pmatrix}
  u' \\
  v' \\
  w' \\
  \Phi'
\end{pmatrix}
= e^{z/2H Re} \left\{ \begin{pmatrix}
  \hat{u}(y) \\
  \hat{v}(y) \\
  \hat{w}(y) \\
  \hat{\Phi}(y)
\end{pmatrix} \exp i(kx + mz - \nu t) \right\}
$$

(2.2)
Here $H$ is a mean scale height and is approximately 7km in the middle atmosphere, $2\pi/k$ is the zonal wavelength, $2\pi/m$ is the vertical wavelength and $\nu$ is the frequency. Throughout this thesis $\nu$ will always be positive and the sign of $k$ will determine whether the wave is eastward or westward propagating.

The hydrostatic balance equation written in the form

$$\frac{\partial \Phi}{\partial z} = H^{-1}RT \quad (2.3)$$

will be used to find waves in temperature ($T$). $R = 287 J K^{-1} kg^{-1}$ is the gas constant for dry air.

Many values of $\hat{u}(y)$, $\hat{v}(y)$, $\hat{w}(y)$ and $\hat{\Phi}(y)$ can be found such that (2.2) is an analytic solution of (2.1). These analytic solutions are now presented, firstly for the simplest case, where the wave has zero meridional velocity, and secondly for the cases where the wave solutions have non-zero meridional velocity.

### 2.3.1.1 Kelvin wave ($v = 0$) solutions and the concept of equivalent depth

The simplest wave type, the Kelvin wave, has $v = 0$ everywhere. Letting $v = 0$ and substituting (2.2) back into (2.1) gives Kelvin waves solutions:

$$u' = e^{z/2H} \hat{U}_m \exp \left( -\frac{\beta ky^2}{2\nu} \right) Re \{ \exp i(kx + mz - \nu t) \} \quad (2.4)$$

$$\Phi' = \frac{\nu}{k} e^{z/2H} \hat{U}_m \exp \left( -\frac{\beta ky^2}{2\nu} \right) Re \{ \exp i(kx + mz - \nu t) \} \quad (2.5)$$

and (2.3) gives
\[ T' = \frac{H \nu}{R k} e^{z/2H} \hat{U}_m \exp \left( -\frac{3k^2y}{2\nu} \right) \text{Re} \left[ \left( \frac{1}{2H} + im \right) \exp \left( ikx + mz - \nu t \right) \right] . \] (2.6)

\( \hat{U}_m \) is the amplitude of the wave in zonal wind on the equator at \( z=0 \) and all other terms have been previously defined.

For all linear waves there is a theoretical relationship between wavenumber and frequency; this is called a dispersion relation. Andrews et al. (1987) p202 show that the dispersion relation for an upward propagating Kelvin wave is

\[ \nu = -\frac{Nk}{m} . \] (2.7)

Only upward propagating Kelvin waves are considered in this thesis, \( m < 0 \) for an upward propagating Kelvin wave, and the speed of a Kelvin wave, \( c_e \), is equal to \( \nu/k \). Equation (2.7) can then be rewritten as

\[ c_e = \frac{N}{|m|} . \] (2.8)

Gill (1982) has shown that for waves in an homogeneous fluid, the speed of propagation of the wave, \( c \), can be written as \( c = \sqrt{gh} \) where \( h \) is the depth of the fluid. In a fluid such as the atmosphere the density is not constant and thus the wave solutions to (2.1) at different pressure levels are not the same. An equivalent depth, \( h_e \), is defined as being the depth of the fluid, \( h \), that would be used if the density of the fluid were constant. Throughout this work this concept of the equivalent depth as defined by \( c_e = \sqrt{gh_e} \) is used. Hence for a Kelvin wave the dispersion relationship can be written as \( \sqrt{gh_e} = \nu/k = N/|m| \).
2.3.1.2 Waves with non-zero meridional velocity

Modes with non-zero meridional velocity and $m^2ν^2 \neq N^2k^2$, (Andrews et al., 1987, section 4.7.2), have an infinite number of solutions for $\hat{u}(y)$, $\hat{v}(y)$ and $\hat{Φ}(y)$. These solutions have increasingly complex meridional structure, which depends on the Hermite polynomials, $H_n$, as follows:

\[
\hat{v}(y) = \hat{v}_0 e^{-(1/2)η^2} H_n(η) \tag{2.9}
\]

\[
\hat{u}(y) = i\hat{v}_0 (β|m|N)^{1/2} \left[ \frac{H_{n+1}(η)}{|m|ν - Nk} + \frac{nH_{n-1}(η)}{|m|ν + Nk} \right] e^{-(1/2)η^2} \tag{2.10}
\]

\[
T(y) = i\hat{v}_0 \frac{H_R}{R} \left( \frac{1}{2H} + im \right) \left( \frac{βN^3}{|m|} \right)^{1/2} \left[ \frac{H_{n+1}(η)}{|m|ν - Nk} - \frac{nH_{n-1}(η)}{|m|ν + Nk} \right] e^{-(1/2)η^2} \tag{2.11}
\]

Here $η = (\frac{β|m|}{N})^{1/2}y$, and the Hermite polynomials are defined as $H_0(η) = 1$, $H_1(η) = 2η$, $H_2(η) = 4η^2 - 2$, ..., $H_{n+1}(η) = 2nH_n(η) - 2H_{n-1}(η)$. Waves with non-zero meridional velocity exist for all $n \geq 0$ where $n$ is the number of sign changes in the meridional plane of a wave in $ν$.

The dispersion relation for waves of non-zero meridional velocity is dependent on $m$, $N$ and $n$ and is:

\[
\frac{m^2ν^2}{N^2} - k^2 - \frac{βk}{ν} = (2n + 1) \frac{β|m|}{N} \tag{2.12}
\]

In this study waves on a constant pressure surface are considered and so information about the vertical wavelength cannot be observed. It is therefore convenient to substitute for $c_e = N/|m|$, where $c_e$ is the equivalent speed of a Kelvin wave with these values of vertical wavenumber and
static stability. This substitution gives the dispersion relationship used in Gill (1982) which is:

\[
\frac{\nu^2}{c_e^2} - k^2 - \frac{\beta k}{\nu} = (2n + 1) \frac{\beta}{c_e}.
\]

(2.13)

It is noteworthy that substituting for \( n = -1 \) in (2.12) and (2.13) gives the dispersion relation for a Kelvin wave. For this reason a Kelvin wave is sometimes referred to as an \( n = -1 \) wave.

If \( n = 0 \) is substituted into (2.13) the dispersion relation becomes:

\[
\frac{\nu}{c_e} - k - \frac{\beta}{\nu} = 0
\]

(2.14)

of which there are two solutions for each value of \( c_e \). These correspond to the westward and eastward components of a Rossby-gravity (RG) wave. [See Matsuno (1966) for an excellent theoretical discussion of this wave.] To avoid confusion the eastward component of this wave type will be referred to as an eastward inertio-gravity wave of \( n=0 \) (or EIG0).

For \( n \geq 1 \), (2.13) has three solutions for each value of \( c_e \). Two of these solutions are relatively high frequency waves and are the eastward and westward propagating inertio-gravity wave for that \( n \) (EIGn, WIGn). The third solution is the low frequency and strictly westward propagating equatorial Rossby wave (ERn).

Wave solutions with different values of \( n \) have different symmetries with respect to the equator. For example, if \( n \) is odd then the wave solutions in \( \hat{u}(y) \) and \( \hat{T}(y) \) are symmetric about the equator and the wave solution in \( \hat{v}(y) \) is antisymmetric about the equator. If \( n \) is even then the wave solutions in \( \hat{u}(y) \) and \( \hat{T}(y) \) are antisymmetric about the equator and the wave solution in \( \hat{v}(y) \) is symmetric about the equator. Table 2.1 presents a summary of how the symmetries of
certain fields due to linear equatorial waves are influenced by \( n \).

<table>
<thead>
<tr>
<th>field</th>
<th>( n ) is odd</th>
<th>( n ) is even</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_n )</td>
<td>antisymmetric</td>
<td>symmetric</td>
</tr>
<tr>
<td>( u )</td>
<td>symmetric</td>
<td>antisymmetric</td>
</tr>
<tr>
<td>( v )</td>
<td>antisymmetric</td>
<td>symmetric</td>
</tr>
<tr>
<td>( T )</td>
<td>symmetric</td>
<td>antisymmetric</td>
</tr>
<tr>
<td>( \frac{\partial T}{\partial y} )</td>
<td>antisymmetric</td>
<td>symmetric</td>
</tr>
<tr>
<td>( u \frac{\partial T}{\partial x} )</td>
<td>symmetric</td>
<td>symmetric</td>
</tr>
<tr>
<td>( v \frac{\partial T}{\partial y} )</td>
<td>symmetric</td>
<td>symmetric</td>
</tr>
</tbody>
</table>

Table 2.1 symmetries of different fields for \( n \) odd or \( n \) even waves

When looking for waves, it is convenient to split the data into its symmetric and antisymmetric components with respect to the equator (Yanai and Murakami, 1970). When this is done, odd or even waves, in each field, need only be looked for in the symmetric or antisymmetric components of that field as appropriate.

The dispersion relations are often represented by their equivalent depth, \( h_e \), as this gives easy comparison between different wave types at different levels in the atmosphere. Wheeler and Kiladis (1999) followed this technique when analysing satellite OLR data. They found that the spectral peaks of OLR in wavenumber-frequency space corresponded well to the wave modes of shallow water theory of implied equivalent depth 12-50m. The dispersion relations for equivalent depths 12m and 100m split into odd and even values of \( n \) is shown in figure 2.3. On figure 2.3, \( s \) represents the zonal wavenumber, and is equal to \( k \times a \), where \( a \) is the radius of the earth. If we take \( N^2 = 5 \times 10^{-4} s^{-2} \), typical of the stratosphere, these equivalent depths correspond to vertical wavelengths of approximately 3000m and 9000m respectively.

To understand how atmospheric disturbances map onto the wave dispersion relations in wavenumber-frequency space (figure 2.3), wavenumber-frequency spectral analysis is used. This is discussed briefly in the next subsection.
Chapter 2 Detecting equatorial wave activity in re-analysis datasets

![Dispersion relations for odd and even waves. These dispersion relations are for waves of equivalent depth 12m and 100m with the lower curve being for 12m in each case.](image)

2.3.2 Wavenumber-frequency spectral analysis

Wavenumber-frequency spectral analysis (Kao 1968, Hayashi 1982) has been used by many authors including Zangvil and Yanai (1980), Wheeler and Kiladis (1999) and Wheeler et al. (2000). This involves transforming the data from longitude-time coordinates to the wavenumber-frequency coordinates of figure 2.3. Briefly, this is done by removing the time-longitude mean and linear trend from the data, performing complex Fast Fourier Transforms (FFT’s) in longitude on each field, followed by complex FFT’s in time on the result.

Once the data has been transformed into wavenumber-frequency co-ordinates, spectral peaks (large power at a certain wavenumber and frequency) imply that a particular wave type is contributing significantly to the variance of the data. Wavenumber-frequency spectral analysis is used in section 2.4 to investigate which waves are likely to be present in the ERA15 data. A method for quantifying these waves is explained in section 2.5.
2.4 Results from wavenumber-frequency spectral analysis

Before any wavenumber-frequency transforms were performed the ERA15 data was split into its symmetric/antisymmetric components about the equator. Subsection 2.3.1.2 showed this approach to be reasonable for theoretical dry, free, linear waves, and case studies of Straub and Kiladis (2002) and Yang et al. (2003) suggest it is reasonable for waves in dynamical fields in the real atmosphere.

Figure 2.4 shows the wavenumber-frequency spectra for the ERA15 data at 70hPa averaged over the latitudes $10^\circ$N - $10^\circ$S and the years 1980-1993, for temperature, zonal wind and meridional wind. From table 2.1 it is seen that waves with $n = \text{odd}$ should arise in symmetric temperature (figure 2.4(a)), symmetric zonal wind (figure 2.4(c)) and antisymmetric meridional wind (figure 2.4(f)). Waves with $n = \text{even}$ should arise in antisymmetric temperature (figure 2.4(b)), antisymmetric zonal wind (figure 2.4(d)) and symmetric meridional wind (figure 2.4(e)). The appropriate dispersion relations for each field are overplotted in figure 2.4. The colour bar on these figures shows the variance attributable to each wavenumber and frequency. It is shown on a log scale.

Figure 2.4 is similar to wavenumber-frequency spectra produced by other authors in satellite data (Gruber 1974, Bergman and Salby 1994, Takayabu 1994a, Wheeler and Kiladis 1999) and model data (Hayashi and Golder, 1994). It is consistent with forcing by diabatic heating with a red noise frequency distribution (Holton, 1973) leading to a red background spectrum on the wavenumber-frequency domain. On figure 2.4 spectral peaks of certain wavenumbers and frequencies are imposed upon this red background. Wheeler and Kiladis (1999) suggested that spectral peaks on a red background were due to equatorial wave activity.

Some of the strongest suggestions of spectral peaks in figure 2.4 occur in the temperature data. On figure 2.4(a) there are spectral peaks corresponding to Kelvin wave dispersion relations between zonal wavenumber 0 and 10 and between frequencies 0 and 0.3 cycles per day. On figure 2.4(b) there are spectral peaks corresponding to Rossby-gravity and EIG0 wave dispersion relations between zonal wavenumber -10 and 10 and between frequencies 0.1 and
Chapter 2  
Detecting equatorial wave activity in re-analysis datasets

Figure 2.4  
logarithmic wavenumber-frequency spectra for various fields at 70hPa averaged over 10°N – 10°S and 1980-1993. The solid lines overplotted are the dispersion relations for equivalent depths 12m and 100m as detailed in figure 2.3
0.4 cycles per day. Spectral peaks corresponding to other wave types are not as obvious, but are not necessarily absent, as Wheeler and Kiladis (1999) observed when they compared their wavenumber-frequency spectra with a measure of the red background.

The zonal wind fields (figures 2.4(c) and 2.4(d)) show spectral peaks corresponding to those observed in the temperature field. The spectral peaks are generally weaker in zonal wind, however there is a much stronger suggestion of $n \geq 1$ inertio-gravity wave activity in this field (see figure 2.4(c) at a frequency of approximately 0.6 cycles per day).

A large spectral peak also corresponds to the westward propagating Rossby-gravity wave in meridional wind. This is seen in figure 2.4(e) between zonal wavenumbers -10 and 0, and between frequencies 0.1 and 0.4 cycles per day. This is perhaps the largest signal relative to the background in all the plots, suggesting that this wave is prominent in the data. The dispersion relation corresponding to Kelvin waves in the meridional wind field contains noticeably lower power than surrounding wave numbers and frequencies. This is partly expected since meridional wind is zero in a Kelvin wave. However, this absence in power for this limited group of wavenumbers and frequencies does challenge the idea of a red background that is independent of the equatorial waves.

In addition to spectral peaks corresponding to the dispersion relations of linear wave theory, there is a spectral peak between the Rossby-gravity dispersion area and the ER2 dispersion area. This peak is seen on figures 2.4(b), 2.4(d) and 2.4(e), between zonal wavenumbers -10 and 0 and at a frequency of approximately 0.1 cycles per day. It is unclear what causes this spectral peak, but it is unlikely to be due to Doppler shifted equatorial waves as it occurs in both the easterly and westerly phases of the QBO (not shown). It will be interesting to see how features such as this are quantified in the wave detection method, for example it may appear as a Rossby-gravity wave of very low equivalent depth or an equatorial Rossby wave of very high equivalent depth. Alternatively it may not be consistent with any type of equatorial wave. The wave detection method, to be described in the next section, should prove very useful when trying to understand features, such as this, in wavenumber-frequency space, which have no obvious cause. It will also
quantify spectral peaks at those wavenumbers and frequencies where waves are expected.

2.5 The wave detection method

Here we derive a method to quantify wave activity in the fields of zonal velocity \(u\), meridional velocity \(v\), and temperature \(T\). The method will detect waves on a chosen pressure surface. Data is required at several latitudes and at every point in a longitude time domain which is large enough to contain the longest wavelength and the lowest frequency wave to be detected. No information about the vertical characteristics of a wave is used in the wave detection method; instead the vertical wavenumber is inferred, from the horizontal structure and frequency of the wave, according to the linear wave theory dispersion relationships which are described in section 2.3.1.

The ERA15 reanalysis data, (discussed in section 2.2) fulfils the requirements of a dataset to which the wave detection method can be applied. It contains data at every gridbox and timestep, with spatial resolution of \(2.5^\circ \times 2.5^\circ\) and temporal resolution of 6 hours. Any number of latitudes of data can be used. For maximum accuracy the method should use enough latitudes to cover the meridional scale of the waves. However, the more latitudes are used, the higher the computational cost of the wave detection method. When the wave detection method is used in chapter 3, 7 latitudes which lie between \(7.5^\circ\text{N}\) and \(7.5^\circ\text{S}\) will be used.

While the wavenumber-frequency diagrams of figure 2.4 provide evidence for the existence of equatorial waves in ERA15, the problem of quantifying the waves remains. It cannot be assumed that all power between the 12m and 100m equivalent depth dispersion relation is a wave of the suggested type or that all power outside this dispersion relation band is not. Wheeler and Kiladis (1999) used the concept of a red background spectrum to provide further evidence for the existence of the waves. They were able to show a high confidence for the existence of waves relative to this ‘background’ but were unable to quantify the background or hence the waves. Here we use some results of linear wave theory (section 2.3.1) to overcome this shortcoming and
will describe the method which will be used in chapter 3 to produce a quantitative climatology of waves in ERA15 data.

Firstly, for each latitude, the wavenumber-frequency domain\(^1\), which here consists of zonal wavenumbers between -20 and 20, and frequencies between 0 and 1 cycle per day (cpd), is split into bins containing small groups of wavenumbers and frequencies. The bins should be as small as possible to get the most accurate results, however, as is often the case, the time taken to perform the calculations is inversely proportional to the size of the bins. In chapter 3 wavenumber-frequency bins will be used which cover 3 wavenumbers and \(\approx 0.06\) cpd of frequency. The original data field \((u, v, T\text{ etc.})\) is then reconstructed just from the data within a single wavenumber-frequency bin. This will show variations in each field attributable to only a small group of wavenumbers and frequencies and waves in the reconstructed field should be easier to detect. Signals in the reconstructed field will be checked for consistency with idealised linear equatorial waves in two ways. Firstly, a signal must have meridional structure consistent with the linear wave theory of section 2.3.1, and secondly the interaction between different fields must be as linear wave theory of section 2.3.1 predicts.

### 2.5.1 Reducing computation time

Unfortunately, comparing \(u, v\) and \(T\) fields with idealised linear waves, for all longitudes, timesteps and wavenumber-frequency bins is extremely computer intensive. A shortcut is required, such that a measure of wave activity can be found in seasonally averaged fields. However this is not straightforward because \(u', v',\) and \(T'\), in a linear wave, are sinusoidal in longitude and time. These fields therefore have zero mean over any latitude circle or any time which is a multiple of the wave’s period. There are some fields, however, which are always positive at all phases of a wave - such as \(|u|, u^2, T^2\). If these fields were averaged around a latitude circle and over a season, the result would be positive. The latitudinal structure of the average of these fields in the atmosphere could be compared with the latitudinal structure of these fields in a linear wave.

---

\(^1\)The size of the domain depends on the resolution of the data and the waves requiring detection
In addition to checking that the waves have a consistent latitudinal structure it is useful if the wave detection method is able to check that there are reasonable interactions between fields. For example, in a Kelvin wave with $|m| \gg \frac{1}{2H}$, temperature perturbations will lead zonal wind perturbations by 1/4 of a cycle, which means that $u$ and $\frac{\partial T}{\partial x}$ are always in phase. Indeed it can be shown that for any given type of wave with $|m| \gg \frac{1}{2H}$ multiplying $u$ by $\frac{\partial T}{\partial x}$ will give the same sign regardless of the phase of the wave. In a Kelvin wave $u\frac{\partial T}{\partial x}$ will always be positive and in a westward propagating Rossby-gravity wave $u\frac{\partial T}{\partial x}$ will always be negative. Similarly $v\frac{\partial T}{\partial y}$ can be shown to be of constant sign in any wave. Because of these properties of $u\frac{\partial T}{\partial x}$ and $v\frac{\partial T}{\partial y}$ it is these fields that will be compared with linear wave theory to detect wave activity.

### 2.5.2 Estimates of waves in a wavenumber-frequency bin

The method for detecting wave activity in a seasonally averaged field for the ‘odd’ waves of table 2.1 is now explained. The method for detecting ‘even’ waves is similar. Firstly, for each latitude, the longitude time data is reconstructed from power in a wavenumber-frequency bin corresponding to the symmetric or antisymmetric part of a field as appropriate. (Recall that a wavenumber-frequency bin contains the spectral power attributable to a small group of wavenumbers and frequencies.) The zonally and seasonally averaged values of $u\frac{\partial T}{\partial x}$ and $v\frac{\partial T}{\partial y}$ are found from the reconstructed data and are written as $u\frac{\partial T}{\partial x}_{ERA}(\text{lat})$ and $v\frac{\partial T}{\partial y}_{ERA}(\text{lat})$. Next, information about idealised waves is required. So for each possible idealised wave, that could occur in the wavenumber-frequency bin, the values of $u\frac{\partial T}{\partial x}$ and $v\frac{\partial T}{\partial y}$ at each latitude are calculated using (2.4), (2.6) and (2.9-2.11). These values are normalised so that for each wave type, $u\frac{\partial T}{\partial x}$ at 5° N is equal to 1. These idealised values are written as $u\frac{\partial T}{\partial x}_\text{wave}(\text{lat})$ and $v\frac{\partial T}{\partial y}_\text{wave}(\text{lat})$, where wave denotes the wave type (i.e. Kelvin, Rossby-gravity etc).

The values of $u\frac{\partial T}{\partial x}$ and $v\frac{\partial T}{\partial y}$ in ERA15 are now assumed to be composed of idealised waves and a residual, $\varepsilon$. This leads to the following equations
where \( A_{\text{wave}} \) represents the value of \( u \frac{\partial T}{\partial x} \) for wave type ‘wave’ at 5° N. Here the terms on the LHS represent those fields, observed in the ERA15 data, in which \( n = \text{odd} \) waves can be found, and so, from table 2.1, symmetric components of \( u \) and \( T \), and antisymmetric component of \( v \), are used. The summation terms on the RHS of (2.15) and (2.16) represent the sum of all the \( n = \text{odd} \) idealised waves which contribute to the observed field. The terms \( \varepsilon_u(lat) \) and \( \varepsilon_v(lat) \) represent the residual in the \( n = \text{odd} \) fields, which cannot be attributed to \( n = \text{odd} \) waves.

Similar equations would arise if we were considering the \( n = \text{even} \) components of each field.

For this study waves will only be looked for with \( n \leq 2 \), so for \( n = \text{odd} \) the only eastward propagating waves that can be found are the Kelvin wave and the EIG1 wave. The only westward propagating waves that can be found are the WIG1 and ER1 waves. Equations (2.15) and (2.16) are derived for \( \text{lat} = 0^\circ \text{N, } 2.5^\circ \text{N, } 5^\circ \text{N, } 7.5^\circ \text{N.} \) This gives 8 equations with 8 unknown residuals and (since only two waves are looked for in \( n = \text{odd} \) fields) two unknown wave amplitudes. The unknown wave amplitudes are estimated by minimising the sums of the squares of the residuals:

\[
\text{MIN} \left( \varepsilon_{u,\text{odd}}(0^\circ \text{N})^2 + \varepsilon_{v,\text{odd}}(0^\circ \text{N})^2 + \varepsilon_{u,\text{odd}}(2.5^\circ \text{N})^2 + \varepsilon_{v,\text{odd}}(2.5^\circ \text{N})^2 + \varepsilon_{u,\text{odd}}(5^\circ \text{N})^2 + \varepsilon_{v,\text{odd}}(5^\circ \text{N})^2 + \varepsilon_{u,\text{odd}}(7.5^\circ \text{N})^2 + \varepsilon_{v,\text{odd}}(7.5^\circ \text{N})^2 \right). \quad (2.17)
\]

Equations (2.15) and (2.16) are used to substitute for the residuals at each latitude in (2.17), and the value of \( A_{\text{wave}} \) for each possible wave type is found to minimise (2.17). Once the value of \( A_{\text{wave}} \) is found it is relatively easy to find the values of \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial x} \) at all latitudes due to that wave type.
2.5.3 Constraints on method

The only constraint that was initially imposed on the wave detection method (in addition to those caused by the limitations of the wave theory in section 2.3.1) is that minimising least squares is not allowed to produce any waves of negative amplitude. If a wave with negative amplitude was found in a wavenumber-frequency bin for any season, then the amplitude of that wave type was set to zero and the amplitudes of other waves for that wavenumber-frequency bin and season recalculated.

Imposing only this constraint generally found waves at those wavenumbers and frequencies which lie inside the dispersion relations of figure 2.3. There were two exceptions to this, these are that Kelvin waves and $n = 2$ inertio-gravity waves were often found at frequencies far below the $h_e = 12m$ dispersion relation.

The low phase speed Kelvin waves could be at least partially due to the Madden-Julian oscillation (see Madden and Julian, 1994, and references therein), hence it is reasonable to find signals of Kelvin waves at these wavenumbers and frequencies. The low frequency inertio-gravity $n = 2$ waves, however, appear less reasonable. The low frequency inertio-gravity $n = 2$ waves were found at frequencies as low as 0.1 cycle per day, which corresponds to an equivalent depth of $< 0.05m$ and vertical wavelength of less than 400m. Gill (1982) p438 shows that the wave solution has exponential decay away from the equator with decay of $e^{-1}$ at $y_e = (c/2\beta)^{1/2}$. For the very small equivalent depth at which inertio-gravity $n = 2$ waves were found, this gives $y_e \sim 1^\circ$. The problems with these waves are twofold. Firstly these waves are not going to be vertically resolved in the ERA15 data, which has vertical resolution near the tropopause of around 2km and therefore can only fully resolve waves with vertical wavelength $> 4000m$. Secondly the latitudinal structure of these waves has such strong exponential decay, that if any of the low equivalent depth waves were to exist, the only latitude in ERA15 where they would easily be detectable is $0^\circ$. This means that the method could attribute any anomalous variance that existed
on the equator to these waves of very low equivalent depth. To prevent the second problem all waves, except for the Kelvin wave, are only looked for at those wavenumbers and frequencies which would give equivalent depth \((h_e)\) greater than 1m. This constraint means that any waves found will have exponential decay from the equator with \(e^{-1}\) at \(y_e \sim 2.5^\circ\), and will be somewhat detectable up to \(7.5^\circ\). The constraint on equivalent depth means that waves can now only be found with vertical wavelength greater than 880m, which is still less than the 4000m required for the waves to be vertically resolved in ERA15. Unfortunately, constraining the method further, so that only those waves that could be vertically resolved were detected would require extending the constraint on equivalent depth to only look for those waves for which \(h_e \geq 20\)m, and would therefore leave a large proportion of the expected waves undetected (see Wheeler and Kiladis, 1999). The constraint on equivalent depth is therefore held at \(h_e > 1\). However any waves found with \(1m \leq h_e \leq 20m\) are likely to be fully dependent on observations assimilated into ERA15.

### 2.6 Discussion of the accuracy of the method

In this section the accuracy of the method described in section 2.5 is discussed, and the extent to which it is able to detect wave motions in the atmosphere is considered.

The first stage of the method, which splits the data into its symmetric and antisymmetric components, means that only disturbances which have the same symmetry as a linear wave can be attributed to that wave type. Secondly, the symmetric or antisymmetric data is Fourier transformed, split into wavenumber-frequency bins and the data reconstructed from a small group of wavenumbers and frequencies only. This means that a disturbance of a particular wavenumber and frequency will be verified against an idealised wave of the same wavenumber and frequency. Next, considering the fields \(u \frac{\partial T}{\partial x}\) and \(v \frac{\partial T}{\partial y}\) ensures that the relationship between \(u\) and \(T\), and \(v\) and \(T\) is at least partially consistent with linear wave theory; for example signals with \(u\) and \(T\) almost in phase would make very little contribution to detected wave activity.

The final part of the method, which compares the latitudinal structure of the observed data with that of idealised linear waves, is perhaps the part of the method most subject to errors. This is because minimising least squares would attribute *some* of each observed field to equatorial waves.
whatever its latitudinal structure. However this final part of the method is vital because it enables the power in a wavenumber-frequency bin to be optimally assigned between different types of linear wave. This means that no prior assumptions are required concerning which waves the power in a wavenumber-frequency bin is likely due to.

Considering all the above points implies that the method will detect signals which fulfil many of the requirements of linear wave theory, without making prior assumptions about which wave type is related to a particular spectral peak. There are, however, likely to be some issues which could affect the accuracy of the method. Some of these issues have already been discussed. For example section 2.2.1 investigated whether waves in ERA15 were likely to represent waves in the real atmosphere. The validity of ERA15 was tested in several ways by comparison with radiosondes and, though there was no evidence to suggest that waves in ERA15 were unrepresentative of waves in the atmosphere, caution must be applied when using results of wave climatologies from ERA15. Another issue, concerning the use of the ERA15 dataset, was discussed in section 2.5.3, where it was found that any waves with equivalent depth $<$ 20m could not be vertically resolved on the ERA15 grid; hence it is likely that constant forcing by observations is required to maintain these waves in the ERA15 dataset. Further issues arise from the linear wave theory of section 2.3.1, on which the method strongly relies. The linear wave theory was derived using a number of assumptions, which are not necessarily valid. For example a zero background flow ($\overline{u} = 0$) was assumed, which is not the case in the atmosphere. When the wave detection method is applied, the effects of Doppler shifting by the mean zonal wind will be accounted for (using section 2 of Hoskins and Yang, 2000), however this will only extend the validity of linear wave theory to the case where the background zonal wind is constant. In reality the background zonal wind is highly variable and also has vertical shear due to the QBO. This vertical shear can have a significant impact on the wave equations (Dunkerton, 1995) but it has not been accounted for because wave theory in a vertical shear zone has not yet been fully developed. In addition, linear wave theory includes further approximations which may affect the accuracy of the method. Firstly the equatorial beta plane approximation was used; however the errors that could arise from using this approximation within $7.5^\circ$ of the equator are negligible. Perhaps more important are the assumptions that the atmosphere is isothermal and all non-conservative processes such as friction and diabatic heating have been ignored; also only upward propagating waves have been
considered. These are issues which will almost certainly affect the validity of linear wave theory and hence the wave detection method, although it is difficult to quantify any errors that will arise from these assumptions.

All of the issues that have been discussed above can be traced to two sources. These are the validity of the ERA15 data and the accuracy of linear wave theory. In addition to these there is a mathematical problem with the method which will arise if more than one type of wave occurs in a wavenumber-frequency bin. We illustrate this problem with an example.

### 2.6.1 A mathematical issue and its effect on the wave detection method

Suppose a wavenumber-frequency bin contains two different types of wave. For this example we consider the case of a Kelvin wave and an eastwards inertia-gravity wave of \( n = 1 \) (EIG1), although any two wave types would give the same problem. Though each wavenumber-frequency bin does overlap a small group of wavenumbers and frequencies, for this example it is assumed that the two different types of waves have identical wavenumber and frequency. Of course no such assumption can be made about the phase and so it is supposed that \( \phi \) represents the phase displacement of the EIG1 wave relative to the Kelvin wave. For simplicity a fixed time and height is considered so the waves are functions of \( x \) only. The zonal wind and temperature of the Kelvin wave are represented by \( u_k \) and \( T_k \) respectively. For the Kelvin wave:

\[
\begin{align*}
    u_k &= f(y) \cos(kx) \\
    \frac{\partial T_k}{\partial x} &= Af(y) \cos(kx)
\end{align*}
\]

The zonal wind and temperature of the EIG1 wave are represented by \( u_{ig1} \) and \( T_{ig1} \) respectively. For the EIG1 wave:
Chapter 2  Detecting equatorial wave activity in re-analysis datasets

\[ u_{ig1} = g(y)\cos(kx + \phi) \]
\[ \frac{\partial T_{ig1}}{\partial x} = Bg(y)\cos(kx + \phi) \]

\( A \) and \( B \) are constants which relate the amplitude of \( u \) to the amplitude of \( \frac{\partial T}{\partial x} \). \( f(y) \) and \( g(y) \) represent the amplitude at each latitude of the Kelvin wave and inertia-gravity wave respectively.

One field that the wave detection method uses is \( u \frac{\partial T}{\partial x} \) averaged over a latitude circle which is:

\[
\left[ u \frac{\partial T}{\partial x} \right] = \frac{Af^2(y)}{2} + \frac{Bg^2(y)}{2} + \frac{(A + B)f(y)g(y)\sin\phi}{2} \quad (2.18)
\]

If the method is detecting waves accurately it will attribute \( \frac{Af^2(y)}{2} \) as the proportion of \( u \frac{\partial T}{\partial x} \) due to Kelvin waves and \( \frac{Bg^2(y)}{2} \) as the proportion of \( u \frac{\partial T}{\partial x} \) due to EIG1 waves. Mathematically \( \frac{(A + B)f(y)g(y)\sin\phi}{2} \) is \( u_k \frac{\partial T_{ig1}}{\partial x} + u_{ig1} \frac{\partial T_k}{\partial x} \) and represents the interaction between the different types of wave. The method cannot detect the interaction term and, depending on its value, this interaction could be attributed to Kelvin waves, EIG1 waves, a residual term or a combination of all three. If the interaction term were large it could lead to significant inaccuracies in the wave climatology. The size of this term in comparison to other terms in (2.18), therefore needs to be assessed.

From figure 2.3 it can be seen that the dispersion curves between waves of the same symmetry are often quite separated, hence often there will be no interaction between waves at all. For the few wavenumbers and frequencies of which more than 1 wave type exists, it will be seen that it is usually the case that one wave type dominates over the other (chapter 3). If this is the case the interaction term will be much smaller than the dominant wave type (even though it could be much larger than the weaker wave type), meaning that the interaction term only forms a small proportion of \( u \frac{\partial T}{\partial x} \). Finally, and perhaps most importantly, it must be noted that the interaction term is dependent on \( \sin\phi \), where \( \phi \) is the phase difference between the different waves. The
average value of $\sin \phi$ will be zero, so on the long term we expect the interaction term to average to zero. For individual seasons, wavenumbers, frequencies, and wave types, however, we must be aware that the accuracy may be compromised by this term.

From figure 2.3 it can be seen that more interaction is expected between waves of different symmetries than between waves of the same symmetry. This is because the dispersion curves between waves of different symmetries are more likely to occur at the same wavenumber and frequency. Fortunately before any analysis was performed the raw data was split into symmetric and antisymmetric components, which means that the interaction between waves of different symmetries can be quantified (see section 2.6.2). It must be noted that mathematical errors will only arise as the interaction between waves of the same symmetry and it is only the interaction between waves of different symmetries which can be quantified. However if the interaction between waves of different symmetries is small it is reasonable to expect that the error arising from interaction between waves of the same symmetry is also small.

### 2.6.2 Interaction between symmetric and antisymmetric waves

Table 2.1 shows that odd modes are found in the fields $u_s \frac{\partial (T_s)}{\partial x}$ and $v_s \frac{\partial (T_s)}{\partial y}$ and even modes are found in $u_a \frac{\partial (T_a)}{\partial x}$ and $v_s \frac{\partial (T_a)}{\partial y}$, where subscript $s$ denotes the symmetric field and subscript $a$ denotes the antisymmetric field. The terms $u_s \frac{\partial (T_a)}{\partial x} + u_a \frac{\partial (T_s)}{\partial x}$ and $v_s \frac{\partial (T_a)}{\partial y} + u_a \frac{\partial (T_s)}{\partial y}$ will contain the interaction between symmetric and antisymmetric wave types. However this may not be the only process appearing in these terms.

To find the best estimate of the symmetric-antisymmetric interaction term we consider those wavenumbers and frequencies to which the method will be applied (namely wavenumber from -20 to 20 and frequencies from 0 to 1 cpd). The value of the interaction term is considered at 50hPa only, as waves explain a larger proportion of $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ at 50hPa than they do at higher pressures (see chapter 3). Figure 2.5 compares the size of the fields in which waves will be looked for (top figures) with the symmetric-antisymmetric interaction terms (bottom figures).
While for individual wavenumbers and frequencies the magnitude of the interaction terms $u_s \frac{\partial(T_a)}{\partial x} + u_a \frac{\partial(T_s)}{\partial x}$ and $v_s \frac{\partial(T_s)}{\partial y} + v_a \frac{\partial(T_a)}{\partial y}$ can be quite large (sometimes as large as $u_s \frac{\partial(T_s)}{\partial x}$ or $v_a \frac{\partial(T_a)}{\partial y}$), the interaction term is sometimes positive and sometimes negative, such that, considering the whole spectrum of wavenumbers and frequencies, the interaction term is comparatively small (see figure 2.5). Averaged over all the years of the dataset, the interaction field, $u_s \frac{\partial(T_a)}{\partial x} + u_a \frac{\partial(T_s)}{\partial x}$ (figure 2.5(c)), is non negative at all latitudes between $0^\circ$ and $7.5^\circ N$, and its average value over these latitudes is over 30 times smaller than either of $u_s \frac{\partial(T_s)}{\partial x}$ and $u_a \frac{\partial(T_a)}{\partial x}$.
The interaction between waves of the same symmetry should be even smaller than this and so is unlikely to be the cause of large errors. The term $v_s \frac{\partial(T_s)}{\partial y} + v_a \frac{\partial(T_a)}{\partial y}$ (figure 2.5(d)) is not positive at all latitudes but its magnitude is small compared to $v_a \frac{\partial(T_s)}{\partial y}$ and $v_s \frac{\partial(T_a)}{\partial y}$. Depending on latitude this term is between 10 and 300 times smaller than both $v_a \frac{\partial(T_s)}{\partial y}$ and $v_s \frac{\partial(T_a)}{\partial y}$. Interaction between waves of the same symmetry will be smaller than this and hence are not expected to cause large errors when the method is used.

While it is noteworthy that the interaction terms in $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ do not appear random but exhibit latitudinal and temporal structure, it is beyond the scope of this thesis to study the structure in more detail.

### 2.7 Summary

This chapter began by considering the likely accuracy of equatorial waves in the ECMWF re-analysis dataset, ERA15. The variance and spectral power of ERA15 dynamical fields showed that equatorial waves are neither more nor less likely in regions close to radiosonde observations than they are in regions spatially distant from a radiosonde station. In section 2.4, wavenumber-frequency spectral analysis was used to show that there are spectral peaks in ERA15 corresponding to dispersion relations for equatorial modes. Hence there is a strong suggestion that the ERA15 data contains equatorial waves. It is therefore reasonable to derive a method which can produce a climatology of equatorial waves in the ERA15 dataset. The climatology produced could be cautiously applied to the real atmosphere.

The method used to produce a climatology of waves in ERA15 is detailed in section 2.5. This method is different to most previous wave detection methods in two ways. Firstly it is able to find wave activity averaged over the tropics and a season, and secondly wave events have not been assumed, a priori, to exist at a limited number of wavenumbers and frequencies. The method will look for waves at all those wavenumbers and frequencies where the existence of a wave does not directly contradict linear wave theory. However, it was shown in section 2.5.3 that it was desirable to impose a constraint on equivalent depth so that the latitudinal structure of the wave
Chapter 2  Detecting equatorial wave activity in re-analysis datasets

is resolved by ERA15; this meant that waves were only looked for at those wavenumbers and frequencies corresponding to an equivalent depth of greater than 1m. For most wave types this is not a very strong constraint, and the method is able to detect linear waves over a large portion of the wavenumber-frequency domain.

The method involves comparing the zonally and seasonally averaged latitudinal structure of \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial y} \) with the expected latitudinal structure of \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial y} \) predicted from linear wave theory. It will therefore detect linear waves as those wave-like disturbances which are consistent with the latitudinal structure, and interactions between dynamical fields that linear wave theory predicts. If the detected linear waves are found at the expected dispersion relations then confidence in the wave climatology increases. Despite this, the method is not without limitations. These are mainly attributable to the limitations of linear wave theory and the ERA15 dataset. However section 2.6.1 shows that a mathematical problem arises, caused by wave-wave interactions, because the method searches for waves in a quadratic quantity. It has been shown that although these wave-wave interactions are not negligible, they are expected to be small compared to the fields in which waves can be detected.

In the following chapter a climatology of equatorial waves in ERA15 data, which has been produced from this method, will be presented.
CHAPTER 3

A climatology of equatorial waves in the lower stratosphere

3.1 Introduction

Chapter 2 described a new method of quantifying linear equatorial waves. This method is now used to detect linear equatorial wave activity in the ERA15 dataset. Climatologies are produced at 100hPa, 70hPa and 50hPa for Kelvin waves, Rossby-gravity waves, eastward propagating inertio-gravity (EIG) waves (with n=0,1,2), westward propagating inertio-gravity (WIG) waves (with n=1,2) and equatorial Rossby (ER) waves (with n=1,2).

Initially the wave climatology is presented for the fields $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$, as these fields arise naturally from the wave detection method. The climatology for each wave type is shown in section 3.2, firstly as a time series showing how wave activity changes with season (1981-1993) and secondly by showing those wavenumbers and frequencies (averaged over 1981-1993) at which each wave is detected.

The fields $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ are useful for studying equatorial waves (chapter 2) and are important when using the thermodynamic equation to model the annual temperature cycle near the tropical tropopause (chapter 5). Section 3.3 considers the total value of the $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ fields averaged over a five year period. It will quantitatively attribute these fields to a linear trend in $u$, $v$ and $T$, to disturbances which are of too high a frequency to be detected as equatorial waves in this study, to the equatorial waves found, and to a residual. Results from this section will be used in chapter 5.

While the wave climatology in the fields $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ is useful for modelling the temperature cycle near the tropical tropopause (chapter 5), it is desirable to produce climatologies of equatorial
waves in more widely used fields so that the climatologies can be compared with previous wave observations and applied to future studies. Section 3.4 presents equations for converting the wave climatologies from $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ into variance fields of temperature, zonal wind and meridional wind. These equations are then used to estimate the contribution of linear waves to the variance of $T$, $u$ and $v$ in the ERA15 data. The linear wave variance is then compared with total variance in ERA15 data to investigate the extent to which disturbances in dynamical fields can be attributed to linear waves.

A summary and discussion of the results of this chapter is presented in section 3.5

### 3.2 Wave climatologies

Figures 3.1 to 3.9 show the climatologies produced for wave types of $n \leq 2$, at 100hPa, 70hPa and 50hPa. Each figure represents the activity of one particular wave type (e.g. Kelvin, Rossby-gravity) on the equator. If a wave of $n = -1$ or $n = 1$ is being presented the climatologies are shown for the field $u \frac{\partial T}{\partial x}$. If a wave of $n = 0$ or $n = 2$ is being presented the climatologies are shown for the field $v \frac{\partial T}{\partial y}$. These fields are being presented, initially, as these fields arise naturally from the wave detection method of chapter 2. Different fields are presented for ‘odd’ and ‘even’ waves because of their different symmetry properties (see chapter 2). The non-zero field is presented for each wave type. Wave activity in $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ has not been documented elsewhere and so there are no studies with which to compare the values of these fields. Therefore the wave climatologies in $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ will be used only to compare seasonal variation in wave activity and also to see which wavenumbers and frequencies contribute to each type of wave.

The wave climatologies were first produced using data directly from ERA15. However the observed frequencies of disturbances in ERA15 are likely to be different from the intrinsic frequencies of the wave motions they represent because of Doppler shifting by the mean zonal wind. To overcome this, the wave climatologies were also produced by replacing the observed frequency, $\nu$, of the disturbances with the intrinsic frequency, $\hat{\nu}$, where $\hat{\nu} = \nu - k \Pi(z)$. However, the intrinsic frequency of each wave was derived using the mean zonal wind averaged over a season, $\Pi(z)$, which will clearly not be as accurate as finding the intrinsic frequency using.
observed zonal winds at smaller spatial and temporal scales.

In each of the wave climatologies the subfigures on the left show seasonal and interannual variation of wave activity between 1981 and 1993. For clarity vertical lines are marked at the start of each year. Here the black line shows the seasonal variation by applying the method of chapter 2 directly to the ERA15 data. The red line shows the seasonal variation when considering the intrinsic frequency ($\hat{\nu} = \nu - k\bar{u}(z)$) of the data. For most of the wave types there is very little difference between total wave activity calculated using the intrinsic frequency or the observed frequency. This is because, while the Doppler shifting of a wave will alter its frequency, it is often likely to be Doppler shifted to a new frequency where it will be partially consistent with a wave of the same type but at a different frequency.

The subfigures on the right of figures 3.1 to 3.9 show those wavenumbers and intrinsic frequencies at which each wave is detected, averaged over 1981-1993. Here the white lines show dispersion relationships for the appropriate wave type. In each case the upper dispersion relationship was calculated using equivalent depth of 100m and the lower dispersion relationship was calculated using an equivalent depth of 12m.

### 3.2.1 Kelvin waves

Figures 3.1(a), 3.1(c) and 3.1(e) present time series of the seasonal variation of Kelvin wave activity between 1981 and 1993 at 100hPa, 70hPa and 50hPa respectively. Kelvin wave activity has roughly equal magnitude at 100hPa and 70hPa, although by 50hPa its magnitude has been reduced by approximately one half. In the absence of non conservative processes, wave activity would be expected to increase with height due to the decrease in density. The fact that wave activity decreases with height shows that there is some damping mechanism operating which has greater effect between 70hPa and 50hPa than it does between 100hPa and 70hPa. This will be investigated in more detail in chapter 4 which considers how waves influence and are influenced by the QBO.
The seasonal variability of Kelvin waves is seen to be very different at the three levels. At 100hPa most of the record shows a biannual peak in Kelvin wave activity with maximums in DJF and JJA (although in 1981 1982 and 1983 the DJF peak is absent and in 1984 the JJA peak is absent). By 70hPa Kelvin wave activity follows an annual cycle with peak in DJF and minimum in JJA. The green line on figure 3.1(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore. It is clear that the 50hPa Kelvin wave activity is correlated with the QBO and there is maximum Kelvin wave activity as the QBO is changing from its easterly phase to its westerly phase.

Despite showing very different seasonal variability at the three levels, these results are consistent with expectations. The biannual peak at 100hPa is consistent with forcing by tropical convection (Holton 1972, Bergman and Salby 1994), since tropical cloud amount exhibits a maximum in JJA in the northern hemisphere, a maximum in DJF in the southern hemisphere (Zhang, 1993), and forcing from both hemispheres will lead to the generation of Kelvin waves. It is unclear why Kelvin wave activity in the early part of the record does not show this biannual peak.

Kelvin wave activity at 70hPa (figure 3.1(c)), appears to vary in phase with the annual cycle of the tropical tropopause. There is maximum activity in DJF when the tropical tropopause is at its coldest and highest (Reed and Vleck 1969, Reid and Gage 1981) and minimum activity in JJA/SON when the tropopause is at its warmest and lowest. This annual cycle in the tropical tropopause means that the buoyancy frequency, $N$, in the 100hPa-70hPa layer will be smaller in DJF than in JJA. Since wave dissipation is directly proportional to the buoyancy frequency (Andrews et al., 1987), Kelvin waves are likely to experience more damping in JJA and this would give the annual cycle that is observed.

Kelvin wave activity at 50hPa reaches a maximum as the zonal flow is changing from easterly to westerly because slow Kelvin waves are damped in westerly winds (see Shiotani and Horinouchi (1993) and references therein). This is consistent with the fact that Salby et al. (1984) and
Figure 3.1  Figures 3.1(a), 3.1(c) and 3.1(e) are time series showing the strength of $u \frac{\partial T}{\partial x}$ due to Kelvin waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.1(b), 3.1(d) and 3.1(f) show those wavenumbers and intrinsic frequencies (averaged over 1981-1993) at which the method detects Kelvin waves. Levels are: 3.1(a) and 3.1(b) 100hPa, 3.1(c) and 3.1(d) 70hPa, 3.1(e) and 3.1(f) 50hPa. The green line in figure 3.1(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore.
Hitchman and Leovy (1988) could only find fast Kelvin wave signals in NIMBUS7-LIMS data for a period when the QBO was westerly (Oct 1978-May 1979) but slow Kelvin wave signals were found in CLAES temperature data on the UARS by Canziani et al. (1995) and Shiotani et al. (1997) when the QBO was easterly. The relationship between Kelvin waves and the QBO is explained more fully in chapter 4.

Figures 3.1(b), 3.1(d) and 3.1(f) show those wavenumbers and frequencies at which Kelvin waves are detected, averaged over all seasons from 1981-1993. At all levels, peak Kelvin wave activity is found in the wavenumber-frequency bin which contains wavenumbers 1-3 and at a frequency of approximately 0.07cpd corresponding to a period of around 15 days. This period and zonal wavenumber is consistent with the first Kelvin waves found by Wallace and Kausky (1968a).

Many Kelvin wave signals that our study finds are within the dispersion relations calculated using equivalent depths 12m and 100m. However Kelvin waves are also found at much lower equivalent depths (particularly at 100hPa) and because of this it is suggested that some of the signal that is consistent with dry, free, linear Kelvin waves is actually caused by the Madden-Julian oscillation (Madden and Julian, 1994). Throughout this thesis we make no distinction between Kelvin waves found at Madden-Julian oscillation frequencies and Kelvin waves found between Kelvin waves dispersion curves of equivalent depth 12m and 100m.

Many previous studies (e.g. Mote et al. 2002, Canziani and Holton 1998) have only looked for Kelvin waves of wavenumber 1 or 2, but figures 3.1(b), 3.1(d) and 3.1(f) suggests that structures similar to Kelvin waves are found at least up to wavenumber 18. In addition Kelvin waves at higher wavenumbers are found over a greater range of frequencies. It has already been mentioned that the strongest Kelvin wave signal exists at wavenumbers 1-3, and this is perhaps why Kelvin waves at these wavenumbers are easier to detect. However integrating over all the frequencies at which Kelvin waves are found shows that total Kelvin wave activity in the field $u \frac{\partial T}{\partial x}$ has only 20% of its contribution from wavenumbers 1-3. It is shown in section 3.4 that wave activity in one dynamical field will not necessarily be composed of the same wavenumbers and frequencies as wave activity in another dynamical field. However, the fact that Kelvin wave activity is clearly
not limited to wavenumbers 1 and 2, shows that considering only these wavenumbers could underestimate the role of this wave in the dynamics of the tropical lower stratosphere.

Finally the wave activity detected using the intrinsic frequency of the data (red line on 3.1(a), 3.1(c) and 3.1(e)) is compared with the wave activity detected using the observed frequency of the data (black line on 3.1(a), 3.1(c) and 3.1(e)). It is seen that Doppler shifting has relatively little effect on the amount of waves detected. However it appears to have greatest effect at 100hPa. This result is counter-intuitive since Doppler shifting would be expected to have the greatest effect at low pressures where the mean zonal wind is stronger. Here Doppler shifting has the greatest effect at higher pressures due to the strong low-frequency, low-wavenumber signal at 100hPa (seen in figure 3.1(b)), which appears as a Kelvin wave but is probably attributable to the Madden-Julian oscillation (MJO). Because this ‘MJO’ signal appears at such low frequencies, only a weak easterly wind would be required to Doppler shift it to westward propagating frequencies, where none of its power could be attributed to Kelvin waves under this method. The 100hPa winds are often easterly, which means that less Kelvin wave activity would be found using the observed frequencies of the data than using the intrinsic frequencies of the data. There is no similar effect of Doppler shifting at 70hPa or 50hPa because at these levels the Kelvin wave signal is dominated by waves of a slightly higher frequency (see figures 3.1(d) and 3.1(f)) and Doppler shifting is not sufficient to change their observed direction of propagation.

3.2.2 Rossby-gravity waves

Apart from the Kelvin wave, the most commonly observed wave in the equatorial upper troposphere and lower stratosphere is the westward propagating Rossby-gravity (RG) wave (e.g. Randel 1992, Stanford and Ziemke 1993, Magaña and Yanai 1995, Maruyama 1991, Dunkerton 1993, Dunkerton and Baldwin 1995). Figures 3.2(a), 3.2(c) and 3.2(e) show time series of Rossby-gravity wave activity in $v \frac{\partial T}{\partial y}$. The red line shows Rossby-gravity wave activity detected using the intrinsic frequency of the data and the black line shows Rossby-gravity wave activity detected using the observed frequency of the data. Figures 3.2(b), 3.2(d) and 3.2(f) show those wavenumbers and frequencies for which Rossby-gravity waves in $v \frac{\partial T}{\partial y}$ are found (averaged
over 1981-1993), and these wavenumbers and frequencies are generally between the dispersion relations of linear wave theory with equivalent depths 12m and 100m. Maximum Rossby-gravity wave activity is found in the wavenumber-frequency bin containing wavenumbers 3-5 and at frequency 0.15cpd which corresponds to a period of around 6 days. These wavenumbers and frequencies compare well with those of the first Rossby-gravity wave observations (Yanai and Marayanna, 1966), which had zonal wavenumber 4, and a period of 4-5 days.

Overall the magnitude of Rossby-gravity wave activity decreases with height (figures 3.2(b), 3.2(d), 3.2(f)) which suggests that there is some mechanism which is causing these waves to dissipate. However the amount of this wave dissipation varies throughout the dataset (compare figures 3.2(a), 3.2(c) and 3.2(e)). The wave dissipation between 70hPa and 50hPa is shown to be consistent with QBO activity in chapter 4.

The time series of Rossby-gravity wave activity as shown in figures 3.2(a), 3.2(c) and 3.2(e) is now considered. At 100hPa Rossby-gravity wave activity follows an annual cycle with peak in JJA or SON. This annual cycle is also apparent at 70hPa for many of the years. However by 70hPa there is also a suggestion of QBO related variation. At 50hPa, variations in Rossby-gravity wave activity are dominated by the QBO (chapter 4). This can be seen by comparing the Rossby-gravity wave activity with the QBO index as shown by the green line in figure 3.2(e). Peak Rossby-gravity wave activity at 50hPa occurs as the flow is changing from westerly to easterly.

Previous results of seasonal variation in Rossby-gravity wave activity suggest an annual cycle at 100hPa and 70hPa, although there is strong disagreement as to its exact phase. This phase disagreement is probably due to the fact that the seasonal cycle in Rossby-gravity waves is dependent on where the Rossby-gravity wave is being observed (Dunkerton 1991a, Dunkerton and Baldwin 1995). Dunkerton (1991a) used data from 32 equatorial rawinsonde stations at 70hPa and found the largest amplitude of Rossby-gravity waves to be in the NH late winter/spring. This contrasts with the NH summer/autumn peak in Rossby-gravity wave activity that we find. Wikle et al. (1997) also considered the seasonal cycle in Rossby-gravity waves. They used 6 rawinsonde
Figure 3.2  Figures 3.2(a), 3.2(c) and 3.2(e) are time series showing the strength of $v^\partial T/\partial y$ due to westward propagating Rossby-gravity waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.2(b), 3.2(d) and 3.2(f) show those wavenumbers and intrinsic frequencies at which Rossby-gravity waves are detected. Levels are 3.2(a) and 3.2(b) 100hPa, 3.2(c) and 3.2(d) 70hPa, 3.2(e) and 3.2(f) 50hPa. The green line on 3.2(e) shows the QBO index as represented by the zonal wind at 50hPa at Singapore.
stations, in the region 90°E-180°E, and found a biannual peak in Rossby-gravity waves with peaks occurring in NH late winter/spring and NH late summer/autumn. They found that while the NH late summer/autumn peak was fairly robust across all the rawinsonde stations considered, the NH late winter/spring peak was not and varied between January and April depending on the location. This meant that averaging Rossby-gravity wave activity over all their stations gave only one peak in Rossby-gravity wave activity in the NH late summer/autumn. Therefore the tropical average in Rossby-gravity wave activity that our study finds has a seasonal cycle consistent with the average seasonal cycle found by Wikle et al. (1997). However the results found here will probably not have the same Rossby-gravity wave seasonal cycle as would be observed at a single location.

The seasonal cycle of Rossby-gravity waves in figure 3.2 also agrees with the seasonal cycle of Rossby-gravity waves in OLR (e.g. Hendon and Liebmann 1991, Wheeler et al. 2000). This indicates that some of the Rossby-gravity waves found here are likely to be convectively forced. If Rossby-gravity waves were due to asymmetric lateral forcing from midlatitudes (Mak, 1969) then maximum activity should occur during solstice seasons.

The seasonal cycle in Rossby-gravity wave activity found here appears reasonable when compared with previous results. Because our study covers the tropical band rather than isolated radiosonde stations these results will be particularly useful when trying to determine tropics wide variations in temperature and momentum flux due to the Rossby-gravity wave.

### 3.2.3 Inertio-gravity waves

Climatologies are presented for EIG0 waves (figure 3.3), EIG1 waves (figure 3.4), WIG1 waves (figure 3.5), EIG2 waves (figure 3.6) and WIG2 waves (figure 3.7). The subfigures on the right of figures 3.3 to 3.7 show the wavenumbers and frequencies for which waves in $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ are detected, and the white lines show the dispersion relations for each wave type with equivalent depths of 12m and 100m. For waves with $n \leq 1$ the wavenumbers and frequencies for which
each wave occurs are consistent with the dispersion relations of equatorial wave theory. The dominant wavenumber and frequencies for \( n \leq 1 \) waves is consistent between pressure levels suggesting that the wave signal covers a relatively deep layer of the atmosphere.

Like the Kelvin wave, all inertia-gravity waves of \( n \leq 1 \) show around a 50% decrease in wave activity between 70hPa and 50hPa. This is despite the fact that wave activity is expected to increase as density decreases and shows there is some damping operating between these levels. This damping will be discussed, along with the QBO in chapter 4. Between 100hPa and 70hPa there is no corresponding decrease in wave activity, (indeed an increase in wave activity is seen for the eastward propagating waves) suggesting that there is much less wave damping between these levels.

The EIG0 wave (figure 3.3) is expected to be the strongest eastward propagating signal after the Kelvin wave. Wheeler and Kiladis (1999) found this wave to explain about the same amount of variance in OLR as the westward propagating Rossby-gravity wave. Here we find \( v \frac{\partial T}{\partial y} \) on the equator attributable to EIG0 waves to be around two-thirds of the \( v \frac{\partial T}{\partial y} \) on the equator that is attributable to Rossby-gravity waves. Maximum EIG0 wave activity is found in the wavenumber-frequency bin with zonal wavenumbers 1-3 and at a frequency of around 0.3cpd. However the value of \( v \frac{\partial T}{\partial y} \) found here is only about one-third of the \( v \frac{\partial T}{\partial y} \) that is due to the strongest wavenumber-frequency bin for Rossby-gravity waves. This, combined with the dominance of Kelvin waves in eastward propagating fields, means that the EIG0 wave would be more difficult to detect than the Rossby-gravity wave. Despite prior lack of evidence for the existence of the EIG0 wave, here it appears extremely robust and shows similar characteristics at 100hPa, 70hPa and 50hPa. Figures 3.3(a) 3.3(c) and 3.3(e) show the seasonal variation in EIG0 wave activity at 100hPa, 70hPa and 50hPa respectively, it is seen that this wave varies annually with peak activity in JJA at all levels. The reason for the seasonal cycle is unclear from this study, as we would need to look at the spatial distribution of the wave activity before any theories about the forcing of this wave or its seasonal cycle could be suggested. Wheeler et al. (2000) did not find the same seasonal cycle for this wave as we find here, instead they found a double peak with maxima in May-July and October-December. However Wheeler et al. (2000) considered the
seasonal cycle in EIG0 wave activity at a single location, in a different field (OLR), at different heights and using a different methodology to what is considered in our study. It is reasonable to suggest that the seasonal cycle of this wave, like that of the Rossby-gravity wave, may not be the same for all variables and pressure levels.

Unlike the Rossby-gravity wave, the EIG0 wave at 50hPa is not strongly correlated with the QBO (discussed in chapter 4). This can be seen by comparison with the green line in figure 3.3(e) which shows the QBO index, as defined by the 50hPa zonal wind at Singapore.

The strongest signal found which is consistent with EIG1 waves occurs at an observed frequency of around 1 cycle per day at 100hPa, 70hPa and 50hPa. This signal is slightly distorted in figures 3.4(b), 3.4(d) and 3.4(f) as these figures show the intrinsic frequency of the disturbances. Eastward propagating disturbances of around 1 cycle per day are most likely due to nonmigrating diurnal tides (e.g. Lieberman, 1991). Nonmigrating diurnal tides, like migrating diurnal tides, are driven by diurnal variations in heating. However while migrating tides are sun synchronised (i.e. westward propagating and zonal wavenumber 1), nonmigrating tides arise from zonal inhomogeneities in the diurnally varying heating constituents and can propagate eastwards, westwards or be stationary (Tsuda et al., 1997). Under our method, some signals that could be nonmigrating diurnal tides appear in the EIG1 wave climatology. Perhaps this is a consequence of the attempt to maximise the amount of $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$ that can be attributed to waves of $n \leq 2$. Although signals of EIG1 waves at frequency 1cpd may not be due to EIG1 waves, they will be included in the EIG1 climatology due to the difficulty in quantifying the amount of this signal which should be removed.

The time series of EIG1 wave activity is shown in figures 3.4(a), 3.4(c) and 3.4(e). Initially the most noticeable feature is the suggestion of a small trend in wave activity between 1985 and 1994, associated with this wave at higher pressures. This trend also occurs in detected WIG1 wave activity and will be discussed later. Apart from this the EIG1 wave has a similar seasonal cycle to that of the EIG0 wave. It also has other features consistent with the EIG0 wave such as the minimums in DJF 1985 and DJF 1993. This raises questions as to whether these waves are
Figure 3.3  Figures 3.3(a), 3.3(c) and 3.3(e) are time series showing the strength of $v \partial T / \partial y$ due to eastwards inertio gravity $n = 0$ waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.3(b), 3.3(d) and 3.3(f) show those wavenumbers and intrinsic frequencies (averaged over 1981-1993) at which EIG0 waves are detected. Levels are 3.3(a) and 3.3(b) 100hPa, 3.3(c) and 3.3(d) 70hPa, 3.3(e) and 3.3(f) 50hPa. The green line on figure 3.3(e) shows the QBO index as represented by the 50hPa zonal wind at Singapore.
forced independently, especially when the only study to have found disturbances consistent with either of these waves types has found disturbances consistent with them both (Wada et al., 1999). However much more work needs to be done on EIG waves before questions about the forcing of these waves can be answered. The green line on figure 3.4(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore. Though the EIG1 wave at 50hPa often varies biennially, the maximum and minimum wave activity do not always occur at the same phase of the QBO. The QBO probably does have some influence over the amount of this wave that is able to reach 50hPa (and will be discussed in chapter 4) but it appears that the interannual variability of this wave at 50hPa is also related to other factors.

Figure 3.5 shows the climatology at 100hPa, 70hPa and 50hPa, for the westward propagating inertio-gravity wave of $n = 1$ (WIG1). Although there are EIG1 wave signals which could be diurnal tides, there is no strong signal which could be attributed to diurnal tides in the WIG1 climatology. At 100hPa, 70hPa and 50hPa the dominant signal has wavenumbers between 12 and 18 although it occurs at a wider range of frequencies at 50hPa and 70hPa then at 100hPa. Like the EIG1 wave, the activity of the WIG1 wave appears to be increasing with time between 1985 and 1993, particularly at 100hPa. Here the trend appears more pronounced in MAM when there is peak wave activity, which has the effect of increasing the amplitude of the seasonal variation towards the end of the dataset. Since wave activity was derived for each season independently, it is unlikely that this trend is an artifact of the wave detection method. It is also unlikely that the linear trend is due to standing waves as the westward and eastward inertio-gravity $n = 1$ waves do not peak at the same magnitudes of wavenumber or frequency. The trend in IG1 wave activity at 100hPa is absent before 1985, indeed EIG1 wave activity shows a decrease between 1983 and 1985. From the data available it is not clear whether the trend between 1985 and 1993 represents a robust increase in wave activity or whether it is part of some long term cyclic variability. Furthermore, it must be remembered that here waves are found in ERA15 data and so changes in the data assimilated into ERA15 between 1985 and 1993 could lead to a trend in ERA15 data that is not representative of the real atmosphere. It would be interesting to produce a longer term climatology of IG1 waves, perhaps using ERA40, to see if any long term cyclic variability in IG1 wave activity is apparent or whether the trend continues after 1993. Until this is done, comments on the cause, or even the robustness, of the trend would be pure speculation.
Figure 3.4  Figures 3.4(a), 3.4(c) and 3.4(e) are time series showing the strength of $u \frac{\partial T}{\partial x}$ due to eastwards inertio-gravity $n = 1$ waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.4(b), 3.4(d) and 3.4(f) are those wavenumbers and intrinsic frequency (averaged over 1981-1993) at which EIG1 waves are detected. Levels are 3.4(a) and 3.4(b) 100hPa, 3.4(c) and 3.4(d) 70hPa, 3.4(e) and 3.4(f) 50hPa. The green line on 3.4(e) shows the QBO index as represented by the 50hPa zonal wind at Singapore.
Figure 3.5  Figures 3.5(a), 3.5(c) and 3.5(e) are time series showing the strength of $\frac{\partial u}{\partial x}$ due to westward inertio-gravity $n = 1$ (WIG1) waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.5(b), 3.5(d) and 3.5(f) show those wavenumbers and intrinsic frequencies at which WIG1 waves are detected. Levels are 3.5(a) and 3.5(b) 100hPa, 3.5(c) and 3.5(d) 70hPa, 3.5(e) and 3.5(f) 50hPa. The green line on figure 3.5(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore.
Ignoring the linear trend for WIG1 waves leaves a clear annual cycle with maximum in MAM and minimum in DJF at 100hPa and 70hPa. As with many other wave types variability at the 50hPa level is dominated by the QBO (discussed in chapter 4). Here, maximum wave activity is found at, or just before, the start of the westerly QBO phase.

Figure 3.7(b) shows that at 100hPa the WIG2 wave climatology is dominated by a signal at wavenumber 1 and period 1 day. This is probably due to migrating diurnal tides. Figures 3.6(b) and 3.6(d) show that at 100hPa and 70hPa there is a strong EIG2 wave signal between wavenumbers 1 and 3 and with period 1 day. This is probably due to nonmigrating diurnal tides. At lower pressures, where the diurnal variability is less dominant, both WIG2 and EIG2 waves generally occur at wavenumbers and frequencies between the \( n = 2 \) IG waves dispersion relations for equivalent depths 12m and 100m. The exception to this is the very large contribution from wavenumber \( \approx 7 \), frequency \( \approx 0.3 \text{cpd} \) to the WIG2 wave at 50hPa, which is seen in figure 3.7(f). If this signal were a WIG2 wave then it would have an equivalent depth of approximately 1.7m, a vertical wavelength of approximately 1200m, and could not be vertically resolved by ERA15. The meridional structure of this wave would have exponential decay away from the equator with decay of \( e^{-1} \) at around 2.5°, and so the values of \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial y} \) at higher latitudes would not affect the detection of this wave. It is noteworthy that before the constraint on equivalent depth \( \geq 1 \text{m} \) was introduced a large number of WIG2 and EIG2 waves were found at very small (\(< 0.05 \text{m}\)) equivalent depths for all pressures. This means the wave detection method has been ‘tuned’ to ensure that inertio-gravity \( n = 2 \) waves were found at appropriate wavenumbers and frequencies. Since large wave signals are still found at unexpected wavenumbers and frequencies after this tuning, doubt is cast upon the accuracy of the method for producing a climatology of inertio-gravity \( n = 2 \) waves.

### 3.2.4 Equatorial Rossby waves

As has already been mentioned, the category of westward propagating waves also includes a low frequency wave named the equatorial Rossby (ER) wave. The climatology found for this wave
Figure 3.6 Figures 3.6(a), 3.6(c) and 3.6(e) are time series showing the strength of $v \frac{\partial T}{\partial y}$ due to eastwards inertio-gravity $n = 2$ waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.6(b), 3.6(d) and 3.6(f) show those wavenumbers and intrinsic frequencies (averaged over 1981-1993) at which the method detects EIG2 waves. Levels are 3.6(a) and 3.6(b) 100hPa, 3.6(c) and 3.6(d) 70hPa, 3.6(e) and 3.6(f) 50hPa. The green line on figure 3.6(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore.
Figure 3.7  Figures 3.7(a), 3.7(c) and 3.7(e) are time series showing the strength of $\nu_{\partial T/\partial y}$ due to westward inertio-gravity $n = 2$ (WIG2) waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.7(b), 3.7(d) and 3.7(f) show those wavenumbers and intrinsic frequencies at which the method detects WIG2 waves. Levels are 3.7(a) and 3.7(b) 100hPa, 3.7(c) and 3.7(d) 70hPa, 3.7(e) and 3.7(f) 50hPa. The green line on figure 3.7(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore.
Chapter 3  
A climatology of equatorial waves in the lower stratosphere

is shown in figure 3.8 for the $n = 1$ wave and figure 3.9 for the $n = 2$ wave. At all levels peak ER1 wave activity is found in the wavenumber-frequency bin containing zonal wavenumbers 3, 4 and 5, and at a frequency of around 0.04 cycles per day, which corresponds to a period of around 25 days. The time series of ER1 wave activity at 100hPa is shown in figure 3.8(a). As usual, the red line shows wave activity calculated using the intrinsic frequency of the data and the black line shows wave activity calculated using the observed frequency of the data. Of all the waves detected, the ER1 wave at 100hPa shows the largest percentage difference between wave activity calculated using the intrinsic and observed frequencies. Larger wave activity is detected using the observed frequency of the data, which is initially surprising given that linear wave theory represents waves which have not been Doppler shifted by the mean zonal wind. The extra wave activity detected using the observed frequency of the data is likely due to some very low frequency Kelvin waves which have been Doppler shifted to westward propagating frequencies and have latitudinal structure partially consistent with ER1 waves.

There have been very few previous studies of equatorial Rossby waves in the lower stratosphere, perhaps because the amplitude of this wave is very small in the stratosphere compared to the troposphere (Wheeler et al., 2000). However, convectively coupled ER1 wave activity is thought to peak in NH winter (Kiladis and Wheeler, 1995) the same as is shown in figure 3.8 at 100hPa and 70hPa. At 70hPa where there is minimum effect of the MJO and the QBO, the seasonal cycle in ER1 wave activity is similar to the seasonal cycle in Kelvin wave activity. This is consistent with the results of Highwood and Hoskins (1998) and Jin and Hoskins (1995) who found that idealised steady heating on the equator would trigger a Kelvin wave to the east of the heating and an equatorial Rossby wave to the west. If these waves are forced simultaneously, for example by idealised heating, then the similarity in the seasonal cycle of these waves is explained.

As with many other waves the temporal variation of ER1 waves at 50hPa is dominated by the QBO (see chapter 4). It is interesting to note that this wave is statistically more correlated with the QBO than either the Kelvin wave or the Rossby-gravity wave.

The nature of the method used here means that even in the absence of wave activity we will
Figure 3.8  Figures 3.8(a), 3.8(c) and 3.8(e) are time series showing the strength of $u \frac{\partial T}{\partial x}$ due to equatorial Rossby $n = 1$ (ER1) waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.8(b), 3.8(d) and 3.8(f) show those wavenumbers and intrinsic frequencies at which the method detects ER1 waves. Levels are 3.8(a) and 3.8(b) 100hPa, 3.8(c) and 3.8(d) 70hPa, 3.8(e) and 3.8(f) 50hPa. The green line on figure 3.8(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore.
Chapter 3 \hspace{1cm} A climatology of equatorial waves in the lower stratosphere

Figure 3.9  Figures 3.9(a), 3.9(c) and 3.9(e) are time series showing the strength of $v_{\partial T/\partial y}$ due to equatorial Rossby $n = 2$ (ER2) waves on the equator. The red line is when Doppler shifting is taken into account, the black line is for when the mean zonal flow is not considered. Figures 3.9(b), 3.9(d) and 3.9(f) show those wavenumbers and intrinsic frequencies (averaged over 1981-1993) at which the method detects ER2 waves. Levels are 3.9(a) and 3.9(b) 100hPa, 3.9(c) and 3.9(d) 70hPa, 3.9(e) and 3.9(f) 50hPa. The green line on figure 3.9(e) shows the QBO index as defined by the 50hPa zonal wind at Singapore.

80
still find traces of any wave looked for provided there is some power at those wavenumbers and frequencies at which the wave is able to exist. Here, we have found traces of ER2 waves, but the values of ER2 waves found are small compared to the other waves found. This is despite the fact that the vertical structure is expected to be resolved in ERA15 for this wave type. The ER2 waves do not seem to have the characteristics of many of the other waves: the temporal variation in ER2 wave activity greatly differs between the levels and there is no annual or QBO related variability. Also, unlike other waves, including Doppler shifting for ER2 waves drastically alters the value of ER2 wave activity that is found. All these factors leads us to suggest that there is no convincing evidence for real ER2 wave activity in the ERA15 dataset and what has been found for this wave type will be discarded.

3.3 Explaining $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$

Figures 3.10, 3.11 and 3.12 show how much of the $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$ fields attributable to wavenumbers -20 to 20 and frequencies 0 to 1 cpd can be attributed to linear waves at 100hPa, 70hPa and 50hPa respectively. The top row of each of these figures shows the field $\frac{\partial T}{\partial x}$ and the bottom row of each figure shows the field $\frac{\partial T}{\partial y}$. The figures on the left show the total value of the field caused by signals with wavenumber from -20 to 20 and frequency from 0 to 1 cycle per day. The middle figures show the amount of each field that is attributed to the linear waves found in this study. The figures on the right show the residual field. In each case the zero contour is shown by the white line. The residual could be due to a variety of factors including interaction between different wave types, non linear waves, forced waves and factors not related to waves. It is possible that some variance attributed to waves is in fact a manifestation of the method of minimising the unexplained variance. This effect remains to be quantified but is expected to be small because any disturbances that are attributed to waves must be at least partially consistent with linear wave theory.

For each level the majority of the $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$ fields with the appropriate wavenumbers and frequencies can be attributed to linear waves. It must be noted that the residuals in both $\frac{\partial T}{\partial x}$ and
Figure 3.10  For wavenumber -20 to 20, frequency 0 to 1cpd and 100hPa. Shows total $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ (left), $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ attributable to wave climatologies (middle) and unexplained $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ (right). The zero contour is shown by the white line.

$v \frac{\partial T}{\partial y}$ are not random but rather have a definite pattern. The pattern may give a clue as to what other processes are operating. Residuals, much like the wave activity, change from following an annual variation near 100hPa to a QBO variation near 50hPa. At all levels, but particularly 100hPa and 70hPa, the residuals are such that the waves seem less equatorially trapped than linear wave theory would suggest.

Total $v \frac{\partial T}{\partial y}$ at 50hPa is negative in DJF 1985 for all latitudes south of the equator, total $v \frac{\partial T}{\partial y}$ at 70hPa for DJF 1985 is smaller than for other seasons of the dataset. The small $v \frac{\partial T}{\partial y}$ in DJF 1985 perhaps explains why EIG wave activity is found to have a minimum for this season. There is no corresponding minimum in Kelvin wave activity because Kelvin wave activity is calculated from the $u \frac{\partial T}{\partial x}$ field only.

Table 3.1 shows the contribution to the $[u \frac{\partial T}{\partial x}]$ and $[v \frac{\partial T}{\partial y}]$ fields by various factors so that the relative importance of the waves can be seen. Data is averaged over the years 1985-1989.
Chapter 3  

A climatology of equatorial waves in the lower stratosphere

Figure 3.11  For wavenumber -20 to 20, frequency 0 to 1 cpd and 70hPa. Shows total $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ (left), $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ attributable to wave climatologies (middle) and unexplained $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ (right). The zero contour is shown by the white line.

inclusive and over the latitudes 7.5° N – 7.5° S. The first row of the table shows the total $|u \frac{\partial T}{\partial x}|$ and $|v \frac{\partial T}{\partial y}|$ at each pressure level. The second row is the amount of $|u \frac{\partial T}{\partial x}|$ and $|v \frac{\partial T}{\partial y}|$ due to the time mean and linear trend of $u$, $v$ and $T$, and subtracting this from the first row gives the third row of the table which shows $|u' \frac{\partial T}{\partial x}|$ or $|v' \frac{\partial T}{\partial y}|$. Finally when waves were looked for in $|u' \frac{\partial T}{\partial x}|$ and $|v' \frac{\partial T}{\partial y}|$ only wavenumbers between -20 and 20 and frequencies in the range 0 - 1 cycle per day were considered. The resolution of the data was such that wavenumbers and frequencies were obtained for wavenumbers up to ±72 and frequencies in the range 0-2 cycles per day. The field denoted ‘outside range’ is the value of the field that is accounted for by wavenumbers of magnitude greater than 20 or frequencies greater than 1 cycle per day. The table shows that in both $|u \frac{\partial T}{\partial x}|$ and $|v \frac{\partial T}{\partial y}|$, waves are relatively more dominant at 50hPa than at 70hPa or 100hPa. At 50hPa linear waves as found in this study explain 63% of $|u \frac{\partial T}{\partial x}|$ and 54% of $|v \frac{\partial T}{\partial y}|$, at 100hPa waves explain 32% of $|u \frac{\partial T}{\partial x}|$ and 33% of $|v \frac{\partial T}{\partial y}|$. In contrast to this, the mean of $u$, $v$ and $T$ are relatively more important at 100hPa than at 50hPa. At 50hPa the magnitude of the residual is smaller than at other levels. For example at 50hPa the residual has magnitude equal to 2.8% of $|u \frac{\partial T}{\partial x}|$ and 16% of $|v \frac{\partial T}{\partial y}|$, at 100hPa the residual has magnitude equal to 7.5% of $|u \frac{\partial T}{\partial x}|$ and
Chapter 3 A climatology of equatorial waves in the lower stratosphere

Figure 3.12 For wavenumber -20 to 20, frequency 0 to 1 cpd and 50 hPa. Shows total $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ (left), $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ attributable to wave climatologies (middle) and unexplained $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ (right). The zero contour is shown by the white line.

34% of $v \frac{\partial T}{\partial y}$. This implies that idealised linear waves are able to represent those disturbances at 50 hPa better than they can represent those disturbances at higher pressures.

3.4 Variance in Dynamical Fields

Though section 3.2 has considered the seasonal variation of wave activity in some detail, the magnitude of wave activity has not yet been considered. This is because, so far, the wave climatologies have been presented in the fields $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ and (with the exception of chapter 5 of this thesis) wave activity in these fields has little practical use.

This section will use the linear wave climatologies in $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ to calculate the variance in
Table 3.1 Decomposition of $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial y}$ at 50hPa, 70hPa and 100hPa, averaged over 1985-1989 and 7.5°N - 7.5°S. Units are Kelvin per day. (*) range is defined as frequency of less than 1 cycle per day and $|\text{zonal wavenumber}| < 20$.

3.4.1 Converting between $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial y}$ and $T^2$, $u'^2$, $v'^2$ 

It can be shown using the results in chapter 2 section 2.3.1.1 that for a Kelvin wave:
\[
\overline{T'^2} = \frac{1}{2} e^{z/H} \frac{H^2 \nu^2}{R^2 k^2} \hat{U}_m^2 \exp(-\beta k y^2/\nu) \left[(1/2H)^2 + m^2\right] \tag{3.1}
\]

\[
\overline{u'^2} = \frac{1}{2} e^{z/H} \hat{U}_m^2 \exp(-\beta k y^2/\nu) \tag{3.2}
\]

\[
\overline{v'^2} = 0 \tag{3.3}
\]

\[
\overline{u' \partial T'/\partial x} = -\frac{1}{2} e^{z/H} \frac{H \nu}{R} \hat{U}_m^2 \exp(-\beta k y^2/\nu) m \tag{3.4}
\]

where all notation is as defined in chapter 2. Using (3.1) to (3.4) the variance in zonal wind and temperature for a Kelvin wave can be calculated from \(\overline{u' \partial T'/\partial x}\) as follows:

\[
\overline{T'^2} = -\left[ \overline{u' \partial T'/\partial x} \right] H \nu (m^2 + (1/2H)^2) R k m \tag{3.5}
\]

\[
\overline{u'^2} = -\left[ \overline{u' \partial T'/\partial x} \right] \frac{R}{m H \nu} \tag{3.6}
\]

Using the results of chapter 2 section 2.3.1.2 it can be shown that, for waves with non zero meridional velocity \(\overline{u' \partial T'/\partial x}\) and \(\overline{v' \partial T'/\partial y}\) are:

\[
\overline{u' \partial T'/\partial x} = -\frac{H m}{2R} v_0^2 \beta N^2 k e^{z/H} e^{-\eta^2} \times \left[ \left( \frac{A}{2} H_{n+1}(\eta) - n B H_{n-1}(\eta) \right) \left( \frac{A}{2} H_{n+1}(\eta) + n B H_{n-1}(\eta) \right) \right] \tag{3.7}
\]

\[
\overline{v' \partial T'/\partial y} = -\frac{H}{2R} m v_0^2 \beta N e^{z/H} e^{-\eta^2} H_n \eta \times \left[ (n+1) A H_n(\eta) - 2n(n-1) B H_{n-2}(\eta) - \frac{\eta}{2} A H_{n+1}(\eta) + n B \eta H_{n-1}(\eta) \right] \tag{3.8}
\]
where \( A = 1/(|m|\nu - Nk) \) and \( B = 1/(|m|\nu + Nk) \).

Using the results of chapter 2 section 2.3.1.2, variance in temperature due to idealised linear waves with non zero meridional velocity is:

\[
\left[ T'^2 \right] = -\hat{v}_0^2 \frac{H^2}{2R^2} \left( \frac{1}{2H} \right)^2 + m^2 \frac{\beta N^3}{|m|} \left[ \frac{A}{2} H_{n+1}(\eta) - n B H_{n-1}(\eta) \right]^2 e^{-\eta^2} e^{z/H} \tag{3.9}
\]

which can be derived from \( \left[ u' \frac{\partial T'}{\partial x} \right] \) using (3.7) as:

\[
\left[ T'^2 \right] = -\left[ u' \frac{\partial T'}{\partial x} \right] H(m^2 + 1/2H^2) N \frac{Rk|m|mH_n(\eta)}{[AH_{n+1}(\eta) - n B H_{n-1}(\eta)]} \left[ \frac{AH_{n+1}(\eta) - 2 n B H_{n-1}(\eta)}{AH_{n+1}(\eta) + 2 n B H_{n-1}(\eta)} \right] \tag{3.10}
\]

or from \( \left[ v' \frac{\partial T'}{\partial y} \right] \) using (3.8) as:

\[
\left[ T'^2 \right] = \left[ v' \frac{\partial T'}{\partial y} \right] H \left( \frac{1}{2H} \right)^2 + m^2 \frac{N^2}{R|m|mH_n(\eta)} \times \frac{[4AH_{n+1}(\eta) - n B H_{n-1}(\eta)]^2}{[(n+1)AH_n(\eta) - 2n(n-1) BH_n-2(\eta) - 4AH_{n+1}(\eta) + n B \eta H_{n-1}(\eta)]} \tag{3.11}
\]

Variance in zonal wind due to idealised linear waves with nonzero meridional velocity is:

\[
\left[ u'^2 \right] = \frac{\hat{v}_0^2}{2} \beta m |N| \left[ \frac{A}{2} H_{n+1}(\eta) + n B H_{n-1}(\eta) \right]^2 e^{-\eta^2} e^{z/H} \tag{3.12}
\]

which can be derived from \( \left[ u' \frac{\partial T'}{\partial x} \right] \) using (3.7) as:
\[
[u'^2] = -\frac{\left| u' \frac{\partial T'}{\partial x} \right|}{H m N k} R |m| \frac{[A H_{n+1}(\eta) + 2 n B H_{n-1}(\eta)]}{[A H_{n+1}(\eta) - 2 n B H_{n-1}(\eta)]} \tag{3.13}
\]

or from \[\frac{u'^2}{\partial y} \] using (3.8) as:

\[
[u'^2] = -\frac{\left| u' \frac{\partial T'}{\partial y} \right|}{H m} |m| R \frac{[A H_{n+1}(\eta) + n B H_{n-1}(\eta)]^2}{H_n(\eta)(n + 1) A H_n(\eta) - 2 n(n - 1) B H_{n-2}(\eta) - \frac{n}{2} A H_{n+1}(\eta) + n B \eta H_{n-1}(\eta)} \tag{3.14}
\]

Variance in meridional wind due to idealised linear waves with non zero meridional velocity is:

\[
[v'^2] = \frac{1}{2} v_o^2 e^{-\eta^2} H_n(\eta) e^{\pm i H} \tag{3.15}
\]

which can be derived from \[\frac{u'^2}{\partial x} \] using (3.7) as:

\[
[v'^2] = -\frac{\left| v' \frac{\partial T'}{\partial x} \right|}{H m B^2 N^2 k} \frac{1}{H m} \frac{R(H_n(\eta))^2}{[\left( \frac{A}{2} H_{n+1}(\eta) - n B H_{n-1}(\eta) \right) \left( \frac{A}{2} H_{n+1}(\eta) + n B H_{n-1}(\eta) \right)]} \tag{3.16}
\]

or from \[\frac{v'^2}{\partial y} \] using (3.8) as:

\[
[v'^2] = -\frac{\left| v' \frac{\partial T'}{\partial y} \right|}{H m B N} \frac{1}{H m} \frac{R H_n(\eta)}{[(n + 1) A H_n(\eta) - 2 n(n - 1) B H_{n-2}(\eta) - \frac{n}{2} A H_{n+1}(\eta) + n B \eta H_{n-1}(\eta)]} \tag{3.17}
\]
These equations can now be used to find the variance in dynamical fields that can be attributed to equatorial waves.

### 3.4.2 Results

Figure 3.13 shows the contribution of equatorial waves to the variance in dynamical fields of the ERA15 data at 50hPa. On figures 3.13(a), 3.13(c), and 3.13(e) the black line is the total variance, the green line is the variance obtained after removing a time mean and linear trend from each latitude-longitude gridbox, and the red line is the variance due to the idealised linear waves found in this study. Figure 3.13 considers the variance on the equator and so only waves with \( n=\text{odd} \) will contribute to the variance in temperature and zonal wind and only waves with \( n=\text{even} \) will contribute to the variance in meridional wind. Figure 3.13 shows that at 50hPa the idealised linear waves found do not explain the majority of the variance in any of the fields, and the same is true of 70hPa and 100hPa (not shown). This is despite the fact that section 3.3 showed that idealised linear waves were able to explain 63% of \( u \frac{\partial T}{\partial x} \) and 54% of \( v \frac{\partial T}{\partial y} \) at 50hPa (see table 3.1). This suggests that \( u, T, \) and \( v \) will not have the same relationship in the data as they would if the data were fully composed of idealised linear waves. This point will be investigated further in section 3.4.3.

Figures 3.13(b), 3.13(d) and 3.13(f) show the contribution of each type of idealised wave to the variance in temperature, zonal wind, and meridional wind respectively. These figures show variance due to Kelvin waves (black), Rossby-gravity waves (turquoise), EIG0 waves (blue), WIG1 waves (light green), EIG1 waves (red), and ER1 waves (dark green). It can be clearly seen by comparing figures 3.13(b) and 3.13(d) that a particular type of wave will have a different relative contribution to the variance of different fields. For example the ER1 wave has a relatively large contribution to the zonal wind variance but a relatively small contribution to the temperature variance. A similar point will arise in chapter 4 where it will be seen that the WIG1 wave has a relatively more important contribution to the momentum flux than to either the temperature or zonal wind variance. The fact that a wave will contribute different amounts to different fields is demonstrated in figure 3.14. This shows those wavenumbers and frequencies that account for Kelvin wave activity at 50hPa (1981-1993) in 3.14(a) temperature, 3.14(b) \( u \frac{\partial T}{\partial x} \), and 3.14(c)
Figure 3.13 3.13(a), 3.13(c) and 3.13(e) show variance in $T$, $u$ and $v$ respectively. Colours are: total variance - black, variance after removing a time linear trend - green, variance due to waves detected - red. 3.13(b), 3.13(d) and 3.13(f) show the contribution from different wave types to the variance in $T$, $u$ and $v$ respectively. Colours are: Kelvin waves - black, WIG1 waves - yellow, EIG1 waves - red, ER1 waves - green, Rossby-gravity waves - turquoise, and EIG0 waves - blue. All figures are on the equator and so only even waves will contribute to $v$ variance while only odd waves will contribute to $T$ and $u$ variance.
momentum flux. It can be seen that lower wavenumbers and frequencies of Kelvin waves are relatively more important to the temperature variance than to the momentum flux or $u \partial T / \partial x$.

It will be shown in chapter 4 that the momentum flux, $u'w'$, due to equatorial waves, can be calculated from $u \partial T / \partial x$, due to equatorial waves, as follows:

$$u'w' = u \frac{\partial T}{\partial x} \frac{R}{HN^2} \frac{\nu}{k}$$  \hspace{1cm} (3.18)

This quantitatively represents the difference between figures 3.14(b) and 3.14(c). It shows that waves of high frequency and low wavenumber are relatively more important to $u'w'$ than they are to $u \partial T / \partial x$. The quantitative difference between the temperature variance and $u \partial T / \partial x$ due to Kelvin waves can also be estimated. This is done by substituting for the Kelvin wave dispersion relation, (2.7), into (3.5) and assuming that $m \gg 1/2H$ to give:

$$\frac{T'}{T^2} = \frac{u \partial T}{\partial x} \frac{HN}{R} \frac{1}{k}$$  \hspace{1cm} (3.19)

This quantitatively represents the difference between figures 3.14(a) and 3.14(b). It shows that small wavenumbers are relatively more important to the temperature variance than they are to $u \partial T / \partial x$.

Kelvin waves have most frequently been looked for in fields of temperature and zonal wind, which could explain why many studies of Kelvin waves have focused on wavenumber 1 and 2. However higher wavenumbers and frequencies are relatively more important to the momentum flux (figure 3.14(c)) which is important for driving the QBO (chapter 4). This does imply that there is no ideal field in which to look for equatorial waves and it will depend on the study in question as to which field should be chosen.
Figure 3.14  Shows those wavenumbers and frequencies that contribute to 3.14(a) temperature variance for a Kelvin wave, 3.14(b) $u \frac{\partial T}{\partial x}$ for a Kelvin wave, and 3.14(c) momentum flux ($u'w'$) for a Kelvin wave, at 50hPa averaged over 1981-1993. The white lines show the dispersion relations for 12m and 100m.

3.4.3 Additional variance

It was shown in the previous subsection that the equatorial waves found in this study are only able to explain a small proportion of the variance in $u$, $v$ and $T$, despite the fact that they are able to explain most of $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$. This will now be investigated.

Figure 3.15 presents a wavenumber-frequency spectrum of 50hPa temperature variance for DJF 1987 on the equator. This will be used to consider why the wave climatology contributes so
little to the variance in dynamical fields. In addition the dispersion relations for equatorial waves with \( H_e = 12 \)m and \( H_e = 100 \)m will be used to define a region in the wavenumber-frequency space for which each equatorial wave type is likely to exist. The wavenumber-frequency region associated with each odd wave type is shown in figure 3.16. Hereafter the blue, black, green, and red areas on figure 3.16 will be referred to as the Kelvin wave region, the ER1 wave region, the WIG1 wave region, and the EIG1 wave region respectively.

**Figure 3.15** Temperature variance for 1987 DJF on the equator due to each wavenumber and frequency. The black lines are the dispersion relations for Kelvin waves and EIG1 waves for equivalent depths 12m and 100m

After removing a linear trend 50hPa temperature variance for DJF 1987 on the equator is \( 3.51K^2 \) (figure 3.13(a)). Of this \( 3.18K^2 \) is due to those wavenumbers and frequencies in figure 3.15. The variance in figure 3.15 is distributed among the wave regions in figure 3.16 as follows: Kelvin=\( 0.89K^2 \), ER1=\( 0.33K^2 \), EIG1=\( 0.33K^2 \), WIG=\( 0.30K^2 \), unmarked=\( 1.33K^2 \). This shows that about half of the temperature variance in figure 3.15 is within wave dispersion relations between \( H_e = 12 \)m and \( H_e = 100 \)m. The variance in each wave region compares with the contribution of our wave climatology to the temperature variance (figure 3.13(b)) of: Kelvin=\( 0.68K^2 \), ER1=\( 0.07K^2 \), EIG1=\( 0.05K^2 \), WIG=\( 0.01K^2 \). Hence linear Kelvin waves explain a significant proportion of the variance in the Kelvin wave region, but waves of \( n = 1 \)

<table>
<thead>
<tr>
<th>Wavenumber(s)</th>
<th>Frequency (cycles per day)</th>
<th>Kelvin wave temperature variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>-20</td>
<td>0</td>
<td>1E-1</td>
</tr>
<tr>
<td>-10</td>
<td>0.2</td>
<td>1E-2</td>
</tr>
<tr>
<td>0</td>
<td>0.4</td>
<td>1E-3</td>
</tr>
<tr>
<td>10</td>
<td>0.6</td>
<td>1E-4</td>
</tr>
<tr>
<td>20</td>
<td>0.8</td>
<td>1E-5</td>
</tr>
<tr>
<td>-10</td>
<td>0.2</td>
<td>1E-6</td>
</tr>
<tr>
<td>0</td>
<td>0.4</td>
<td>1E-7</td>
</tr>
</tbody>
</table>
account for very little of the temperature variance in their wave region. This shows that much of the unexplained temperature variance in figure 3.13(a) lies in the regions appropriate to $n = 1$ waves.

**Figure 3.16** Wavenumbers and frequencies for which each wave is likely to exist, as defined by the area between the dispersion curves for $H_e = 12$ and $H_e = 100$. Blue, black, red and green areas are Kelvin, ER1, EIG1 and WIG1 regions respectively.

It has just been shown that the temperature variance in regions of $n = 1$ waves is not explained by the idealised wave climatologies. We will now investigate why this is the case. For this study we are detecting linear waves by verifying three of their characteristics, namely: 1) realistic wavenumber and frequency, 2) $u$, $v$, and $T$ are related in accordance with linear wave theory, and 3) the latitudinal structure of the wave is as predicted by wave theory. It is clear that the wavenumber and frequency of disturbances in the $n = 1$ wave regions are appropriate to a $n = 1$ wave type, hence it is now considered whether two fields $u$ and $T$ are related in a way consistent with theory.

For each of the wave regions in figure 3.16, the relationship between $u$ and $T$ is investigated as is now described. Firstly, it is assumed that all of the temperature variance in the wave region is due to the suggested type of idealised linear wave. For example, the temperature variance in
the red region of 3.16 is assumed to be due to idealised linear EIG1 waves. Secondly (3.5) or (3.10) is used to calculate the value of $u \frac{\partial T}{\partial x}$ from the idealised linear wave induced temperature variance. This is denoted $u \frac{\partial T}{\partial x}$ (ideal). Next the observed value of $u \frac{\partial T}{\partial x}$ is calculated for the wave region and is denoted $u \frac{\partial T}{\partial x}$ (obs). Finally $u \frac{\partial T}{\partial x}$ (obs) is found as a percentage of $u \frac{\partial T}{\partial x}$ (ideal). This is shown in figure 3.17 for DJF of 1987, at 50hPa, averaged over 7.5°N-0°.

Figure 3.17  Shows for each wave type $u \frac{\partial T}{\partial x}$ in the wave regions of figure 3.16 as a percentage of what it would be if all the temperature variance in the wave region was due to linear idealised waves. These are averaged over 7.5°N-0°

Figure 3.17 shows that, in the Kelvin, Rossby-gravity, and EIG0 wave regions, $u \frac{\partial T}{\partial x}$ (obs) represents a reasonable percentage of $u \frac{\partial T}{\partial x}$ (ideal). Hence a significant proportion of temperature disturbances in these regions vary with zonal wind as predicted by linear wave theory. In the EIG1, WIG1, and particularly the ER1, wave regions, the value of $u \frac{\partial T}{\partial x}$ (obs) is small compared to $u \frac{\partial T}{\partial x}$ (ideal). Hence the temperature disturbances in these regions are not related to the zonal wind disturbances in a way consistent with linear wave theory. It is noteworthy that a greater proportion of the EIG1 wave region is consistent with linear waves than the WIG1 wave region; this is likely due to the fact that there are some fast Kelvin waves occurring in the EIG1 region (see figure 3.1(f)), and that the Kelvin waves agree more with linear wave theory than inertio-gravity waves.
From figure 3.17 it appears that the disturbances in the regions corresponding to Kelvin, Rossby-Gravity, and EIG0 waves have $u$ and $T$ in better agreement with linear wave theory than disturbances at those wavenumbers and frequencies which correspond to $n = 1$ wave regions. The results of figure 3.17 could occur because disturbances in $n = 1$ wave regions have $u$ and $T$ more in phase than linear wave theory suggests. Alternatively, the results of figure 3.17 could occur if $u$ and $T$ have the same phase as linear wave theory predicts, but the amplitude of disturbances in $u$ are smaller than is consistent with linear wave theory, given the amplitude of disturbances in $T$. To investigate the relationship between $u$ and $T$ further, the analysis used to produce figure 3.17 is repeated using the zonal wind variance instead of the temperature variance. In other words, the zonal wind variance in each of the wave regions of figure 3.16 was used with (3.6) or (3.13) to find the value of $u \frac{\partial T}{\partial x}$ that would result if all of the zonal wind variance in each wave region was due to idealised linear waves. The observed $u \frac{\partial T}{\partial x}$ in each wave region was then found as a percentage of the $u \frac{\partial T}{\partial x}$ calculated from the zonal wind variance. This is shown in figure 3.18.

**Figure 3.18** Shows for each wave type $u \frac{\partial T}{\partial x}$ in the wave regions of figure 3.16 as a percentage of what it would be if all the zonal wind variance in the wave region was due to linear idealised waves. These are averaged over 7.5N-0N

From figures 3.17 and 3.18 it is seen that the proportion of observed $u \frac{\partial T}{\partial x}$ in Kelvin, Rossby-
gravity and EIG0 wave bands that is consistent with zonal wind variance is smaller than is consistent with temperature variance. This means that the amplitude of zonal wind disturbances in these wave regions are larger than would be consistent with the amplitude of temperature disturbances under linear wave theory. This explains why idealised linear waves account for relatively more of the temperature variance at 50hPa (figure 3.13(a)) than the zonal wind variance at 50hPa (figure 3.13(c)).

To make the relative amplitudes of temperature variance and zonal wind variance in the Kelvin, Rossby-gravity, and EIG0 wave regions fully consistent with linear wave theory, the temperature variance would need to be approximately doubled or the zonal wind variance would need to be approximately halved. Figure 2.2 of chapter 2 shows that if the ERA15 data were altered by this amount than it would no longer be in agreement with radiosonde observations, and would be unlikely to be representative of the real atmosphere. The fact that the amplitude of the zonal wind disturbances are larger than is consistent with the amplitude of the temperature disturbances under linear wave theory, is therefore likely to be the case for the real atmosphere and not just an artifact of the ERA15 data. This could mean that there are physical processes in the atmosphere that cause much stronger wavelike disturbances in zonal wind than in temperature. Alternatively, if linear wave theory were modified to account for vertical wind shear, non-isothermal background state and non conservative processes, then it might be found that the observed temperature variance and zonal wind variance were more consistent with equatorial waves.

The proportion of \( \frac{u}{\partial T/\partial x} \) that is in agreement with temperature variance in the WIG1 wave region (figure 3.17) is similar to the proportion of \( \frac{u}{\partial T/\partial x} \) that is in agreement with zonal wind variance (figure 3.18). This means that in the WIG1 wave region the amplitude of disturbances in temperature and zonal wind are consistent with linear wave theory, although the phase displacement between temperature and zonal wind disturbances is not. (The same is probably true of EIG1 waves, but the EIG1 wave region is contaminated by Kelvin waves). The low values in figures 3.17 and 3.18 for WIG1 and EIG1 waves is therefore due to the fact that the inertio-gravity wave regions have temperature and zonal wind much more in phase than linear wave theory suggests. The ER1 wave region also has very small values in figures 3.17 and 3.18.
These low values are partly due to the fact that disturbances in this region is more in phase than linear wave theory suggests and partly due to the fact that the observed $u \frac{\partial T}{\partial x}$ has different signs at different latitudes and hence observed $u \frac{\partial T}{\partial x}$ is very small when averaged over $7.5^\circ 0 - 0^\circ N$.

So far this section has shown that linear waves can only explain a very small proportion of the variance in dynamical fields. The temperature variance is better explained than zonal wind variance and reasons for this have already been discussed. We now consider why equatorial waves can only account for a very small proportion of the temperature variance. Figure 3.17 shows that the value of $u \frac{\partial T}{\partial x}$ in the wavenumber-frequency regions corresponding to Kelvin, Rossby-gravity, and EIG0 waves could explain a significant proportion of the temperature variance. However for wavenumber-frequency regions appropriate to $n = 1$ waves, zonal wind and temperature are more in phase than is predicted by linear wave theory. This means that the wave detection method filters out much of the variance in dynamical fields before waves are assigned. Possible reasons why the phase relationship between $u$ and $T$ do not agree with linear wave theory for $n = 1$ waves are considered in the next subsection.

### 3.4.4 Considering $n = 1$ waves

It was shown in the previous subsection that disturbances in regions appropriate to waves of $n = 1$ do not have the required phase relationships between $u$ and $T$ for them to be explained by linear wave theory. There are several possible reasons for this, these are:

1. The disturbances in the $n = 1$ wave regions may be due to equatorial waves, but the approximations made in linear wave theory are such that idealised linear waves are not consistent with waves in the real atmosphere. For example the linear wave theory of chapter 2 assumes a zero background flow and an isothermal atmosphere and neglects non conservative processes. These assumptions could lead to the idealised linear waves that the wave detection method uses being very different to waves in the real atmosphere, and could cause many atmospheric waves in the ERA15 dataset to remain undetected.
2. Disturbance in \( n = 1 \) wave regions may not be due to linear waves at all, but due to some, as yet, unknown disturbance in the atmosphere.

3. Dynamical fields (particularly \( u \) and \( v \)) at those wavenumbers and frequencies appropriate to \( n = 1 \) waves may not be well represented by the ERA15 data. For example ERA15 may be accurately representing the amplitude of disturbances in wind fields, but may not be accurately representing their phase. In the atmosphere \( u \) and \( T \) may have a phase relationship more consistent with linear wave theory. However this is difficult to verify due to a lack of observations.

Before considering the validity of the ERA15 data, we first discuss whether disturbances in the \( n = 1 \) wave regions are likely to be due to equatorial waves which are poorly represented by linear wave theory.

This study and many others have found disturbances associated with idealised Kelvin and Rossby-gravity waves. However very few previous studies have found disturbances that are directly attributable to \( n = 1 \) idealised linear waves. It is therefore possible that the assumptions which are made in the idealised linear wave equations lead to a reasonable representation of waves with low meridional structure, but are not able to represent more complex waves with higher meridional structure. However, much more work is needed to investigate the effects of the assumptions used in linear wave theory before this possibility can be investigated.

For idealised linear waves, \( u \) and \( T \) are approximately in quadrature such that \( u \frac{\partial T}{\partial x} \) is the same sign at all phases of a wave (see chapter 2). However the last subsection has shown that this phase relationship does not hold exactly for disturbances in ERA15 data. At wavenumbers and frequencies corresponding to \( n = 1 \) wave regions this relationship barely holds at all. Sato et al. (1994) found similar results in radiosonde data at Singapore. They considered the in-phase variation of \( T \) and \( u \), represented by the cospectra (hereafter known as \( C_{TU} \)), and the out of phase variation of \( T \) and \( u \), represented by the quadrature spectra (hereafter known as \( Q_{TU} \)). If atmospheric disturbances were idealised linear waves \( Q_{TU} \) would be much larger than \( C_{TU} \). However their results (figure 3.19) do not show this. For waves with a period of greater
than about 5 days (which approximately corresponds to our Kelvin waves) \( Q_{TU} \) is larger than \( C_{TU} \), but \( C_{TU} \) is too large to be fully explained by linear wave theory. This is approximately the same as what is found in figures 3.17 and 3.18, where it is seen that \( u \frac{dT}{dx} \) is too small to fully predict temperature variance or zonal wind variance under the assumptions of linear wave theory. For waves with a period of less than around 2 days (which approximately corresponds to inertio-gravity waves), Sato et al. (1994) find that \( C_{TU} \gg Q_{TU} \) which directly contradicts the linear wave theory prediction that \( Q_{TU} \gg C_{TU} \). However the results of Sato et al. (1994) are in agreement with our study (figures 3.17 and 3.18), which finds that \( u \) and \( T \) are sufficiently in phase that \( u \frac{dT}{dx} \) can only explain a small proportion of temperature variance or zonal wind variance under the assumptions of linear wave theory.

Ultimately Sato et al. (1994) concluded that disturbances with periods of less than 2 days could not be explained by zonally or meridionally propagating linear wave theory as it stands. It is possible that modifying linear wave theory, to include the effects of vertical shear, a non-isothermal atmosphere and non-conservative processes may lead to the predicted cospectra between temperature and zonal wind being more in line with observations. To date vertical shear has been included in equatorial wave theory for the 2 dimensional case (Dunkerton, 1995), and it has been shown that this could increase the cospectra between temperature and zonal wind, particularly in strong shear zones. This means that the in-phase variation between temperature and zonal wind, which was filtered out of the ERA15 data before any waves were looked for, could be due to equatorial waves in a vertical shear zone. Sato et al. (1994) also noted that disturbances with periods less than 2 days may not be directly due to short period equatorial waves but rather due to convective/turbulent motions associated with breaking of larger scale Kelvin and Rossby-gravity waves. Many details about these high frequency disturbances in the stratosphere remain unknown. It appears that while idealised linear wave theory has reasonable skill at predicting the motions of lower frequency waves, the theory will need to be enhanced considerably before these high frequency wave motions can be understood.

The phase relationship that we find between \( u \) and \( T \) is similar to what Sato et al. (1994) found in radiosonde data and is probably robust. However it must be remembered that our study uses
Figure 3.19  Reproduced from Sato et al. (1994). shows a) Quadrature spectra $Q_{TU}$ and b) cospectra $C_{TU}$ averaged for a height region of 20-24.5km of $T$ and $u$ fluctuations. Contour intervals are $2.5\text{Kms}^{-1}$. Thick curves show monthly-mean zonal wind averaged for the same height region. Areas with positive value are shown in red.

Reanalysis data which may not be fully representative of the real atmosphere. Section 2.2.1 showed that the amplitudes of disturbances in $u$ and $T$ compare reasonably with radiosonde data, however we have not considered whether the phase relationship between $u$ and $T$ is representative. It must be mentioned that data input into ERA15 is weakly constrained by thermal wind balance, and that only the Kelvin wave is consistent with this constraint. To fully verify our results the effect of the thermal wind balance constraint on wave motions input into ERA15 would need to be quantified, a task which is not undertaken here. However, because the covariance of $T$ and $u$ in ERA15 is in such strong agreement with the covariance of $T$ and $u$ in the radiosonde data at Singapore (Sato et al., 1994) (a dataset which is not constrained by thermal wind balance), it is expected that the effect of the thermal wind balance constraint will be small.
3.5 Summary

One of the points that has consistently arisen throughout this chapter is that there are still many uncertainties regarding equatorial waves. To help understand some of these uncertainties we have produced a tropics-wide climatology of idealised linear equatorial waves in ERA15 data and estimated their effects on some of the dynamics in the lower stratosphere. Throughout this chapter we have compared our results with previous observational and modelling studies of equatorial waves and a coherent picture is starting to emerge. Before the results are summarised it must be noted that the climatology here refers to waves in the ERA15 data and the extent to which this data is representative of the real atmosphere is uncertain; it is because of this uncertainty that comparisons with previous studies of equatorial waves in radiosonde, rawinsonde, satellite and radar data are invaluable. This study, and studies of waves in observed data should be mutually beneficial. Studies of waves in observed data are more accurate locally, while this study applies to the whole of the deep tropics and results presented here are not biased towards any one location.

The wave climatologies are first presented in section 3.2. Evidence is found for the existence of all waves with \( n \leq 1 \), but there is no evidence for ER2 waves. Many signals are found which are consistent with IG2 waves, however these must be treated with caution because not enough latitudes are used in our study to fully resolve the meridional structure of these waves.

Kelvin wave activity and Rossby-gravity wave activity have seasonal and interannual variations consistent with previous studies. Temporal variations in Kelvin waves are consistent with: forcing by convection, less damping in the troposphere than the stratosphere, and preferential damping in the easterly phase of the QBO. Temporal variations of Rossby-gravity waves are similar to what has been found for Rossby-gravity waves in OLR (e.g. Hendon and Liebmann, 1991) suggesting that the Rossby-gravity waves found here could be forced by convection.

The fact that there have been few previous observations of EIG0, EIG1, and WIG1 waves, combined with the fact that we are using reanalysis data, implies that results for these wave types must be used with caution. However the evidence for the EIG0 wave is so strong in the dataset
that it is likely to be relevant to the real atmosphere. Future studies on the EIG0 wave may note
that in this dataset it has the strongest signal in JJA, frequency $\approx 0.3 \text{ cpd}$ and wavenumbers $1 - 3$.

Disturbances consistent with EIG1 and WIG1 waves are weak compared to the other waves
observed in this study, which is perhaps why they have not been frequently found in the
atmosphere. One benefit of this study over most previous studies is that it is able to quantitatively
attribute signals of around 2 days to an optimal combination of Kelvin waves, EIG0, EIG1 and
WIG1 waves.

Section 3.3 has shown idealised linear equatorial waves are able to explain a significant propor-
tion of the $u\frac{\partial T}{\partial x}$ field and the $v\frac{\partial T}{\partial y}$ field in the tropics. The total amount of $u\frac{\partial T}{\partial x}$ and $v\frac{\partial T}{\partial y}$ that can
be attributable to linear equatorial waves is larger at 50hPa than 70hPa, both of which are larger
than at 100hPa.

Finally this chapter has shown that the contribution of equatorial waves to the variance in $T$, $u$,
and $v$ is small. This is partly due to temperature and zonal wind at inertio-gravity wavenumbers
and frequencies being more in phase than equatorial wave theory would suggest. It is unclear
whether this is because of poor representation of these wavenumbers and frequencies in ERA15
or whether this issue is appropriate to the real atmosphere. It is likely that the large in-phase
variation between temperature and zonal wind will have at least some relevance to the real
atmosphere as Sato et al. (1994) found similar results in radiosonde data at Singapore. Including
vertical shear and non-conservative processes in linear wave theory may account for at least
some of the in-phase variation. To understand inertio-gravity waves more clearly, high frequency
disturbances should be checked against equatorial waves that are more representative of the
atmosphere than the idealised linear waves detected in this thesis.
CHAPTER 4
The interaction between equatorial waves and the QBO

4.1 Introduction

The interaction between equatorial waves and the QBO has been discussed qualitatively in chapter 1, and was observed for some wave types at 50hPa in chapter 3. Here the interaction between equatorial waves and the QBO is discussed in more detail, both theoretically and with reference to the wave climatologies of chapter 3.

As a wave propagates vertically and zonally through the stratosphere it will experience radiative damping. The longer the wave takes to reach a particular level of the atmosphere the more radiative damping it will experience, and hence damping is dependent on the vertical group velocity of the wave. Chapter 1 showed that the vertical group velocity of a wave will become very small, and hence a wave will be strongly damped, as the wave approaches its critical level (where the background zonal wind is equal to the wave’s phase speed). This effect is such that the wave will have fully dissipated before its critical level is reached. The wave dissipation just below a critical level has two main effects. Firstly, the critical levels of the mean zonal flow acts as a filter for the wave spectrum, and secondly the wave dissipation below the critical level will lead to a momentum flux convergence which will accelerate the mean zonal flow. Both of these effects will be discussed in this chapter.

In section 4.2 the effect of the background flow on the vertical group velocity of each wave type will be considered in more detail, both theoretically and in relation to the wave climatology at 50hPa that was derived in chapter 3. It will be considered whether waves in ERA15 are affected by the background flow in a way qualitatively consistent with theory.
A strong motivation for studying equatorial waves was to understand their role in driving the QBO (see chapter 1). This is the focus of section 4.3, which considers whether the momentum flux convergence, due to the equatorial waves detected in chapter 3, is able to accelerate the QBO between 70hPa and 50hPa.

The acceleration of the QBO provided by individual wave types with frequency less than 1 cycle per day and wavenumber less than 20 is also quantified. The wave types considered are: Kelvin waves, Rossby-gravity waves, eastward propagating inertia-gravity (EIG) $n = 0$ waves, inertia-gravity $n = 1$ waves, and equatorial Rossby (ER) $n = 1$ waves. Quantifying the acceleration due to each of these wave types should provide observational insights into the mechanisms involved in driving the QBO and should provide a useful reference for future modelling studies.

### 4.2 Variability of the waves in different QBO phases

Chapter 3 has shown that at 50hPa the activity of some wave types in ERA15 is strongly correlated with the phase of the QBO, while the activity of other wave types is not. It was found that ER1 wave activity was most strongly correlated with the QBO phase, but a correlation was also seen between the phase of the QBO and the activity of Kelvin waves, Rossby-gravity waves and WIG1 waves. No clear correlation between the phase of the QBO and EIG0 or EIG1 wave activity was found. In this section it is considered whether the wave climatologies at 50hPa are related to the QBO in a way qualitatively consistent with linear wave theory.

Chapter 3, figure 3.1(e) shows that Kelvin wave activity at 50hPa reaches a maximum as the zonal flow is changing from easterly to westerly. This pattern has been found in many studies (e.g. Maruyama 1991, Shiotani and Horinouchi 1993, Shiotani et al. 1997, Canziani and Holton 1998). This peak in Kelvin wave activity at the transition between easterlies and westerlies was accurately modelled by Shiotani and Horinouchi (1993). Their model was based on the fact that a wave suffers damping, per unit vertical distance propagated, which is inversely proportional to the wave’s vertical group velocity.
Equatorial waves have vertical group velocity, \( c_g^{(z)} \), at height \( z \), calculated as \( c_g^{(z)}(z) = \partial \nu / \partial m \) (Andrews et al., 1987, p205), where \( \nu \) is the frequency of the wave and \( 2\pi/m \) is the vertical wavelength. The vertical group velocity for a typical Kelvin wave (frequency of 0.05 cycles per day and zonal wavenumber 1), in various background zonal winds is shown by the black line in figure 4.1(a). The red line shows the number of days it would take for this wave to vertically propagate through 5km. It can be seen that in strong westerly winds the wave has much smaller group velocity, and hence takes longer to vertically propagate 5km, than in easterly winds. If the radiative damping timescale in the lower stratosphere is around 20 days (as was suggested by Haynes et al., 1991 and Holton et al., 1995), then this wave would be strongly damped if it propagated 5km through westerly winds of more than around 10 m/s. However the wave would experience relatively little damping if it propagated 5km through easterly winds. Peak Kelvin wave activity in ERA15 is observed when the 50hPa flow is changing from easterly to westerly. At this time the Kelvin wave would have propagated upwards through easterly winds and figure 4.1(a) shows that peak Kelvin wave activity at this time is theoretically reasonable. However it is unclear from figure 4.1(a) why there is less observed Kelvin wave activity in other seasons which have easterly flow at 100hPa, 70hPa and 50hPa.

For the Rossby-gravity wave, which peaks in ERA15 as the QBO flow changes from westerly to easterly, a similar argument is made. Figure 4.1(b) shows the vertical group velocity (black line), and time to vertically propagate through 5km (red line), for a typical Rossby-gravity wave (frequency of 0.15 cycles per day and zonal wavenumber of 4). This wave will experience extremely strong damping when propagating through 5km of easterly winds. It will therefore obtain its largest amplitude at those times when it has propagated through westerly zonal winds. Peak Rossby-gravity wave activity at 50hPa in the westerly to easterly QBO transition is consistent with this analysis.

Figure 4.1(c) shows that for a typical EIG0 wave (frequency of 0.3 cycles per day and zonal wavenumber of 2), there is much less variation in vertical group velocity (black line), or time taken to propagate vertically through 5km (red line), with differing QBO phase, than for either the Kelvin or the Rossby-gravity wave. Indeed the EIG0 wave takes less than 20 days to vertically...
Chapter 4

The interaction between equatorial waves and the QBO

propagate through 5km in all the wind regimes considered. Unlike the Kelvin wave and the Rossby-gravity wave, some of the EIG0 wave signal will, therefore, be able to propagate upwards to 50hPa in all phases of the QBO. From figure 3.3 it appears that the activity of EIG0 waves at 50hPa is more dependent on the activity of EIG0 waves at 70hPa than the phase of the QBO. This suggestion is supported by the analysis from linear wave theory that is shown in figure 4.1(c).

Figure 4.1(d) shows that an ER1 wave (frequency of 0.05 cycles per day and zonal wavenumber 4) has very small group velocity in all easterly winds. The phase speed of this wave is approximately -6m/s, and hence if the background wind was 6m/s westerly, this wave could not vertically propagate at all. This is consistent with figure 3.8(e), where it is shown that the ER1 wave peaks in the westerly phase of the QBO and almost disappears in the easterly phase of the QBO. It has already been mentioned that, of all the waves, this one is the most correlated with the QBO. This is reasonable given the strong limitations on its propagation abilities.

Figures 4.1(e) and 4.1(f) show how the group velocity, and time taken to vertically propagate 5km, varies with mean zonal wind for eastwards and westwards propagating inertio-gravity waves of \( n = 1 \). Figures 4.1(e) and 4.1(f) are calculated using those wavenumbers and frequencies for which the corresponding wave shows peak activity (frequency of 0.8 cycles per day and wavenumber 10 for EIG1, frequency of -0.8 cycles per day and wavenumber 15 for WIG1). Though figures 4.1(e) and 4.1(f) show the effect of background wind on the group velocity of an inertio-gravity wave of a particular wavenumber and frequency it must be noted that the group velocity would be different if alternative wavenumbers and frequencies had been used in its calculation. This fact applies to all waves, but it is noted here because there are strong inertio-gravity wave signals at a multitude of wavenumbers and frequencies. In general the higher the wavenumber the stronger the dependence of the group velocity on the mean zonal wind. The reason the WIG1 wave has smaller group velocity than the EIG1 wave in extreme zonal winds is because peak WIG1 wave activity is found at a higher zonal wavenumber than peak EIG1 wave activity. This shows why the EIG1 wave at 50hPa is less correlated with the QBO than the WIG1 wave at 50hPa.
Chapter 4  The interaction between equatorial waves and the QBO

(a) a Kelvin wave of 0.05 cycles per day and zonal wavenumber 1, corresponding to a phase speed of 23m/s

(b) a Rossby-gravity wave of 0.15 cycles per day and zonal wavenumber 4, corresponding to a phase speed of -17m/s

(c) an EIG0 wave of 0.3 cycles per day and zonal wavenumber 2 corresponding to a phase speed of 69m/s

(d) an ER1 wave of 0.05 cycles per day and zonal wavenumber 4 corresponding to a phase speed of -6m/s

(e) an EIG1 wave of 0.8 cycles per day and zonal wavenumber 10 corresponding to a phase speed of 37m/s

(f) a WIG1 wave of 0.8 cycles per day and zonal wavenumber 15 corresponding to a phase speed of -24m/s

Figure 4.1  Black line shows vertical group velocity of different types of wave in various background winds. Red line shows the number of days for the wave to propagate 5km in various background winds.
This section has shown that the ERA15 wave climatologies at 50hPa are related to the QBO in a way qualitatively consistent with theory. Those waves found in ERA15 which are correlated with the QBO have very small vertical group velocities in particular QBO phases, such that the time taken for them to vertically propagate through 5km is large compared to a radiative damping timescale of 20 days. These waves can therefore be subject to strong radiative damping in certain QBO phases. Waves in ERA15 which have been found to exhibit little, or no, QBO related variability have been found to propagate through 5km of the atmosphere in under 20 days for all background wind speeds of magnitude less than 20m/s. These waves are therefore not subject to strong radiative damping in any phase of the QBO.

The relationship between waves and the QBO is mutually dependent: the waves propagate preferentially in certain phases of the QBO, and when waves are damped they will transfer their momentum to, and accelerate, the mean zonal flow in the lower stratosphere. The next section will focus on the extent to which the waves found in ERA15 are able to accelerate the mean zonal flow and therefore drive the QBO.

### 4.3 Forcing of the QBO by equatorial waves

#### 4.3.1 Background

To date the only process suggested to drive the QBO that can successfully account for some of its basic characteristics is vertical transfer of momentum by equatorial waves. This was first proposed by Lindzen and Holton (1968), who showed in a 2D model how a QBO could be driven by a broad spectrum of vertically propagating gravity waves. Around this time there were the first observations of Kelvin waves (Wallace and Kausky, 1968a) and Rossby-gravity waves (Yanai and Marayanna, 1966), such that subsequent models of the QBO (e.g. Holton and Lindzen, 1972) considered forcing by the observed Kelvin and Rossby-gravity waves only.

The Holton and Lindzen (1972) model of the QBO was accepted for many years, although there were concerns that the Rossby-gravity wave was unable to account for the easterly acceleration of
Chapter 4  The interaction between equatorial waves and the QBO

the QBO. It was suggested that WIG1 and ER waves were also required (Lindzen and Tsay 1975, Andrews and McIntyre 1976). Later it was found that neither Kelvin waves nor Rossby-gravity waves were sufficient to drive either phase of the QBO, when Dunkerton (1991b) showed that including tropical upwelling in a model of the QBO would increase the momentum forcing required to drive the oscillation. The existence of tropical upwelling meant that the momentum flux needed to drive the QBO was roughly twice what had previously been assumed and could therefore not be fully accounted for by observed Kelvin and Rossby-gravity waves. For example Dunkerton (1997) estimated the vertical flux of momentum due to Kelvin waves between 20-25km in active periods to be $6 \times 10^{-3} m^2 s^{-2}$. He suggested that the vertical component of the Eliassen-Palm flux at the tropopause due to Rossby-gravity waves was in the range $-(1 - 2.5) \times 10^{-3} m^2 s^{-2}$. Models, which included tropical upwelling and reproduced the QBO from Kelvin and Rossby-gravity waves, were assuming much greater momentum fluxes due to these waves. For example Gray and Pyle (1989) were able to produce a QBO in a two dimensional model using vertical momentum flux at the tropopause of: $12.5 \times 10^{-3} m^2 s^{-2}$ and $3.5 \times 10^{-3} m^2 s^{-2}$ for two Kelvin waves, and $15 \times 10^{-3} m^2 s^{-2}$ for a Rossby-gravity wave. In an early 3 dimensional model, Takahashi and Boville (1992) noted that the waves being produced to drive a realistic QBO were of amplitudes $\approx 10$ times that observed for Rossby-gravity waves and $\approx 3$ times that observed for Kelvin waves.

The current theory of the QBO (Dunkerton, 1997) is that it arises from a broad spectrum of waves, similar to the original theory of Lindzen and Holton (1968). However the relative contribution of different wave types to the forcing of the QBO is still unknown.

Recently a number of modelling studies have produced a QBO like oscillation using various combinations of equatorial waves. These modelling studies are 3 dimensional which means that they are able to mechanically force waves (ie by imposing heating) rather than by parameterising them. Takahashi (1996, 1997) found that a model could produce a realistic QBO when forced with Kelvin waves, eastward and westward propagating inertio-gravity and gravity waves, and Rossby waves. The Rossby-gravity wave did provide some forcing but it was small compared to other waves. Horinouchi and Yoden (1998) found that in the lower stratosphere the dominant eastward propagating wave was the Kelvin wave and the dominant westward propagating wave
was the $n = 1$ WIG wave. Takahashi (1999) showed that the easterly forcing of the QBO was predominantly due to gravity waves with only a small contribution from Rossby waves. Different models are therefore able to produce a QBO-like oscillation using different combinations of equatorial waves, and so observations are required to verify which combination of waves are appropriate to the atmosphere. This provides the motivation for obtaining the acceleration to the mean zonal wind that is due to the waves of period $> 1$ day that were detected in ERA15.

Before the acceleration due to waves in ERA15 is discussed, it must be noted that recently work has focused on the relative contribution to the QBO of waves which are resolved by a model and waves which require parameterisation. Scaife et al. (2000) simulated the QBO in the Met. Office Unified Model using specifically adjusted gravity wave drag parameterisations. They found that at 10hPa resolved waves generally contributed less to the QBO than parameterised waves. However the relative roles of different wave types depended on the phase of the QBO and the gravity wave parameterisation used. Giorgetta et al. (2002) considered the QBO, using the middle atmosphere configuration of the ECHAM5 model, without adjusting the parameterisations of momentum flux due to gravity waves that the model contained. They found that waves which could be resolved with model vertical resolution of 1.5km dominated slightly over parameterised waves in the lower stratosphere. However there was still a strong contribution to the QBO from parameterised waves. Also the resolved waves that their model contained were more intermittent than the parameterised waves.

It thus appears that much of the momentum flux required to drive the QBO will be due to waves of period $< 1$ day that have not been detected in this study. Nonetheless, the waves that have been detected will contribute to the momentum flux in the middle atmosphere and will therefore have some role to play in driving the QBO. Knowledge of momentum flux due to observed waves will be useful for comparisons with model simulations of the QBO which rely on a broad spectrum of atmospheric waves, many of which have never been directly observed.
4.3.2 Calculating momentum flux and zonal acceleration

The QBO can be modelled using the Transformed Eulerian Mean form of the zonal momentum equation. This is:

$$\frac{\partial \overline{u}}{\partial t} + \overline{u}' \left( \frac{\partial \overline{u}}{\partial y} - f \right) + \overline{w}' \frac{\partial \overline{u}}{\partial z} = \rho_0^{-1} \nabla \cdot \mathbf{F} + \overline{X}$$ (4.1)

where

$$\overline{u}' = \overline{u} - \frac{1}{\rho_0} \left( \rho_0 \frac{\overline{v'} \theta'}{\theta_z} \right)_z$$

$$\overline{w}' = \overline{w} + \left( \frac{\overline{v'} \theta'}{\theta_z} \right)_y.$$

Here $\overline{X}$ represents the zonal component of nonconservative mechanical forcing. $\nabla \cdot \mathbf{F}$ is the divergence of the Eliassen-Palm flux (see Andrews and McIntyre, 1976) and represents the effects of eddy heat flux and eddy momentum flux on the zonal flow. $\mathbf{F} \equiv (0, F^{(y)}, F^{(z)})$ where

$$F^{(y)} = \rho_0 [\overline{u_z v'} \theta'/\theta_z - \overline{u' v'}]$$

$$F^{(z)} = \rho_0 [(f - \overline{u_y}) \overline{v' \theta'}/\theta_z - \overline{w' w'}].$$

Some of the terms in (4.1) will be non zero for the waves that were detected in chapter 3. It is through these terms that waves can accelerate the mean zonal flow and hence drive the QBO. To understand, quantitatively, the role of the waves detected in chapter 3 in the QBO, some of the terms in (4.1) will be examined. The first term to be examined is the vertical momentum flux, $\overline{u' w'}$, in $F^{(z)}$, that is due to the equatorial waves detected in chapter 3. There are many previous
estimates of this term and it is expected to provide a significant contribution to the QBO. The other terms in (4.1) that are non zero for equatorial waves are those that are some multiple of $\overline{v'\theta'}$. However, since the theory used to detect equatorial waves assumed a zero background flow (chapter 2), the only term that is a multiple of $\overline{v'\theta'}$ which will be analysed is the term $f \overline{v'\theta'}/\theta z$ in $F^{(z)}$. The acceleration of the flow due to these wave induced terms, $\partial u_w/\partial t$, can then be modelled using:

$$\frac{\partial u_w}{\partial t} = \rho_0^{-1} \frac{\partial}{\partial z} [\rho_0 f \overline{v'\theta'}/\theta z - \overline{u'w'}].$$ \quad (4.2)

At present waves have been detected in the fields $\overline{w^2} \partial T'/\partial x$ and $\overline{v^2} \partial T'/\partial y$ and so it is necessary to convert to waves in the fields $\overline{v'\theta'}$ and $\overline{u'w'}$. Using the results of chapter 2 it can be shown that for an idealised linear equatorial wave with non zero meridional velocity

$$\overline{v'\theta'} = -\frac{1}{2} \left( \frac{P_s}{P} \right)^{\kappa} \frac{H}{R} \hat{\theta}_0^2 \sqrt{\frac{\beta N^2}{|m|}} H_n(\eta) \left[ \frac{A}{2} H_{n+1}(\eta) - nBH_{n-1}(\eta) \right] e^{-\eta^2} e^{z/H}. \quad (4.3)$$

where $A = \frac{1}{|m|\nu - NK}$ and $B = \frac{1}{|m|\nu + NK}$, and all other notation is defined in chapter 2. It can also be shown that for a idealised linear Kelvin wave

$$\overline{u'w'} = -\frac{\nu^2 m}{2N^2 k} \exp (z/H) \hat{U}_m \exp \left( -\frac{\beta k y^2}{2\nu} \right) \quad (4.4)$$

and for an idealised linear equatorial wave with non zero meridional velocity

$$\overline{u'w'} = -\frac{\nu \beta m}{2} \hat{\theta}_0^2 \left[ \frac{A}{2} H_{n+1}(\eta) - nBH_{n-1}(\eta) \right] \times \left[ \frac{A}{2} H_{n+1}(\eta) + nBH_{n-1}(\eta) \right] e^{z/H} e^{-\eta^2}. \quad (4.5)$$
Chapter 4 The interaction between equatorial waves and the QBO

Using (3.7), \( \overline{v'\theta'} \) for an idealised linear equatorial wave of non zero meridional velocity can be calculated from \( u'\frac{\partial T'}{\partial x} \) as

\[
\overline{v'\theta'} = \frac{u'}{\nu R} \frac{\sqrt{|m|\beta}}{kN} \left( \frac{P_s}{P_s} \right)^n H_n(\eta) \frac{\kappa}{\sqrt{|m|\beta N}} \left[ A_2 H_n + 1 \right] \eta_n \left( \eta \right) \right].
\] (4.6)

Using (3.4) and (3.7), \( u'w' \) for an equatorial wave can be calculated from \( u'\frac{\partial T'}{\partial x} \) for an equatorial wave as

\[
\overline{u'w'} = \frac{u'}{\nu R} kH N \left( \eta \right). \] (4.7)

However \( \overline{v'\theta'} \) and \( u'w' \) for an idealised linear equatorial wave with non zero meridional velocity can also be calculated from \( v'\frac{\partial T'}{\partial y} \) using (3.8) as follows:

\[
\overline{v'\theta'} = v' \frac{\partial T'}{\partial y} \left( \frac{P_s}{P_s} \right)^n \sqrt{N} \frac{\sqrt{|m|\beta}}{kN} \left( \eta \right) \left[ \frac{A_2 H_n + 1}{A_2 H_n + 1} \right] \eta_n \left( \eta \right) \left[ A_2 H_n + 1 \right] \eta_n \left( \eta \right) \left[ (n + 1)AH_n - 2n(n - 1)BH_{n-2} + \frac{2}{V}AH_{n+1} + nB\eta H_{n-1} \right] \right].
\] (4.8)

\[
\overline{u'w'} = u' \frac{\partial T'}{\partial y} \frac{R\nu}{H N H_n(\eta)} \left( \frac{P_s}{P_s} \right)^n \sqrt{N} \frac{\sqrt{|m|\beta}}{kN} \left( \eta \right) \left[ \frac{A_2 H_n + 1}{A_2 H_n + 1} \right] \eta_n \left( \eta \right) \left[ A_2 H_n + 1 \right] \eta_n \left( \eta \right) \left[ (n + 1)AH_n - 2n(n - 1)BH_{n-2} + \frac{2}{V}AH_{n+1} + nB\eta H_{n-1} \right] \right].
\] (4.9)
4.3.3 Momentum Flux Results

Equations (4.6) to (4.9) have been used to calculate $u'w'$ and $f'\theta'/\theta_z$ at 70hPa and 50hPa, for various waves averaged over 7.5°N - 0N. Considering $u'w'$ and $f'\theta'/\theta_z$ separately is somewhat artificial since a linear wave will affect both these terms, and the terms will act together to accelerate the mean zonal winds of the QBO. However there are advantages to considering these terms separately. Many previous studies have only considered the values of $u'w'$ due to equatorial waves. Since we consider the values of $u'w'$ and $f'\theta'/\theta_z$, we are able to compare our results for $u'w'$ with previous studies and also comment on the relative size of $f'\theta'/\theta_z$.

The field $u'w'$ is shown in figures 4.2 and 4.3, while the field $f'\theta'/\theta_z$ is shown in figure 4.4. In these figures the black line shows the field calculated using the equation which is a function of $u'\frac{\partial T'}{\partial x}$ (4.6) or (4.7)) and the red line shows the field calculated using the equation which is a function of $v'\frac{\partial T'}{\partial y}$ (4.8) or (4.9)). In nearly all cases, the two equations give the same result, and so the red and black lines occur in the same place and are difficult to visually distinguish. Calculating $u'w'$ and $f'\theta'/\theta_z$ in two different ways is a useful check against errors in the algebra.

In a very small number of cases the value of $u'w'$ or $f'\theta'/\theta_z$ calculated from the equation which is a function of $u'\frac{\partial T'}{\partial x}$ is not the same as the value calculated from the equation which is a function of $v'\frac{\partial T'}{\partial y}$. An example is seen in the field $f'\theta'/\theta_z$ due to the ER1 wave on figure 4.4. Here the two equations give different results in some of the seasons. However it is unusual to obtain different results when $u'w'$ and $v'\theta'$ are calculated in the two different ways. Any differences can always be traced to a particular wavenumber and frequency, and are due to rounding errors caused by a very small denominator in at least one of the equations, (4.6) to (4.9). These differences are too small to affect the results and can therefore be ignored.

The estimates of $u'w'$ and $f'\theta'/\theta_z$ due to the waves found in this study will first be used to consider the direction that each wave will accelerate the mean zonal flow between 70hPa and 50hPa. As each wave propagates between 70hPa and 50hPa it will be subject to some radiative damping. The effects of this damping can be partially seen in figures 4.2 to 4.4, where $|u'w'|$ and $|f'\theta'/\theta_z|$ at 70hPa are greater than they are at 50hPa. However the damping would really act
Chapter 4 The interaction between equatorial waves and the QBO

on $\rho_0 |u'w'|$ and $\rho_0 |f v' \theta' / \theta_z|$ (see (4.2)) and hence would have larger effects than what is seen. Because all waves are damped between 70hPa and 50hPa (4.2) shows that a positive value of $u'w'$ would lead to westerly acceleration while a negative value of $u'w'$ would lead to easterly acceleration. Similarly a positive value of $f v' \theta' / \theta_z$ would lead to easterly acceleration while a negative value of $f v' \theta' / \theta_z$ would lead to westerly acceleration. To see which direction the waves are accelerating the mean zonal flow the sign and relative sizes of $u'w'$ and $f v' \theta' / \theta_z$ will be considered.

In a Kelvin wave $v' = 0$ and hence the vertical Eliassen-Palm flux due to Kelvin waves is equal to $-\rho_0 u'w'$. Since $u'w' > 0$ for a Kelvin wave, this wave will accelerate the mean zonal flow westerly as is expected. Both Rossby-gravity and EIG0 waves have positive $u'w'$ and $f v' \theta' / \theta_z$ and so these terms considered separately would accelerate the flow in opposing directions. However these terms cannot act separately. It is found that for Rossby-gravity waves $u'w' < f v' \theta' / \theta_z$ and for EIG0 waves $u'w' > f v' \theta' / \theta_z$. This means that any dissipation of the Rossby-gravity waves or EIG0 waves found in this study will accelerate the flow in the direction of wave propagation. For inertio-gravity $n = 1$ waves it is seen that $f v' \theta' / \theta_z < 0$ for both eastward and westward propagating waves so this term, alone, would accelerate the flow eastwards were these waves to dissipate. However because inertio-gravity waves have $|f v' \theta' / \theta_z| < |u'w'|$, the forcing from $u'w'$ will dominate. This will mean that dissipation of the WIG1 wave will accelerate the flow westwards while dissipation of the EIG1 wave will accelerate the flow eastwards. The overall contribution of inertio-gravity waves to the QBO will be westward because WIG1 waves are stronger and more susceptible to dissipation than EIG1 waves. Both $f v' \theta' / \theta_z$ and $u'w'$ due to ER1 waves are small. However in contrast to some other waves (e.g. Rossby-gravity) both the terms will accelerate the flow in the same direction (namely westerly). This allows the ER1 wave to have a relatively strong effect on the acceleration, as will be shown in section 4.3.4.

This section has shown that, under the linear wave theory described in chapter 2, dissipation of all the waves found in this thesis will accelerate the mean zonal flow in the same direction that the wave was propagating, as is expected. The partitioning between the eddy heat flux and the eddy momentum flux has been considered out of theoretical interest; however it is of little
practical importance since these terms will not occur separately. There have been many previous estimates of $u^'w^'$ (although rather less of $\int \theta^'/\theta_z$). The values of $u^'w^'$ that have been found to be due to equatorial waves here are now compared with those of previous studies.

Figure 4.2 shows that $u^'w^'$ due to Kelvin waves varies from approximately $5 \times 10^{-4} m^2 s^{-2}$ in the extreme westerly QBO phase to approximately $2 \times 10^{-3} m^2 s^{-2}$ in the easterly QBO phase. Previous estimates of $u^'w^'$ are: $4 \times 10^{-3} m^2 s^{-2}$ for a Kelvin wave with $8ms^{-1}$ amplitude in zonal wind (Wallace and Kausky, 1968b), $2 - 9 \times 10^{-3} m^2 s^{-2}$ for 5-20 day period waves in the lower stratosphere above Singapore when there was a westerly shear zone (Sato and Dunkerton, 1997), and $1 \times 10^{-4} m^2 s^{-2}$ and $2 \times 10^{-4} m^2 s^{-2}$ for two Kelvin waves observed between 20km and 25km when the QBO was in the westerly phase (Holton et al., 2001). These previous results are in reasonable agreement with $u^'w^'$ due to Kelvin waves detected in ERA15 (figure 4.2), however these other estimates are in general larger than are shown in figure 4.2. This could be partially due to the fact that previous studies often used data from a single radiosonde station near the equator while our results are averages over a tropical band. Also Sato and Dunkerton (1997) assumed that everything in a 5-20 day period was due to Kelvin waves. Here no such assumption has been made, instead only those signals which project onto idealised linear Kelvin waves are considered.

Sasi and Deepa (2001) considered $u^'w^'$ due to Rossby-gravity waves using data from the Indian MST radar at Gadanki (13.5N, 79.2E) when the QBO was westerly (Sept 1995 - Aug 1996). They find $u^'w^'$ due to Rossby-gravity waves to be $2.1 - 6.9 \times 10^{-3} m^2 s^{-2}$, which is substantially larger than what we find here. However their study also found $u^'w^'$ due to Kelvin waves in the westerly QBO phase to be $5.5 - 27 \times 10^{-3} m^2 s^{-2}$. Hence their results overestimate what we find for both Kelvin waves and Rossby-gravity waves by a similar proportion. Our estimates of Rossby-gravity waves vertical momentum fluxes are also smaller than what was estimated by Dunkerton (1997): he calculated $u^'w^' = 1.9 \times 10^{-3} m^2 s^{-2}$. However his estimate was for a typical Rossby-gravity wave and hence the average momentum flux due to Rossby-gravity waves would be smaller than this since Rossby-gravity wave activity is intermittent (Dunkerton, 1991a). Although we find $u^'w^'$ due to Rossby-gravity waves is many times smaller than $u^'w^'$ due to Kelvin waves we note
that its seasonal evolution is almost exactly opposite to the Kelvin wave. This is as expected since the Rossby-gravity wave should be strongest in the westerly phase of the QBO when the Kelvin wave is weakest, and weakest in the easterly phase of the QBO when the Kelvin wave is strongest.

Figures 4.2 and 4.3 show that $\overline{\mathbf{u}}'\mathbf{w}'$ due to inertio-gravity waves is small in comparison with $\overline{\mathbf{u}}'\mathbf{w}'$ due to Kelvin waves. This is not the same as was found by Sato and Dunkerton (1997), who found $|\overline{\mathbf{u}}'\mathbf{w}'|$ due to waves with period 1-3 days to be $20 - 60 \times 10^{-3} m^2 s^{-2}$, which was larger than the $|\overline{\mathbf{u}}'\mathbf{w}'|$ they found due to waves with period 5-20 days. Sato and Dunkerton...
Figure 4.3 Shows $\overline{u'w'}$ (m$^2$/s$^{-2}$) due to the EIG1, WIG1 and ER1 waves at 70hPa and 50hPa.

(1997) found the momentum flux for waves with period 1-3 days to be approximately 2 orders of magnitude larger than the momentum flux that we attribute to the inertio-gravity waves. Momentum flux due to inertio-gravity waves found in our study is also much smaller than that suggested by Vincent and Alexander (2000), they estimated $\overline{u'w'}$ due to inertio-gravity waves at Cocos Island ($12^\circ$S, $97^\circ$E) between September 1992 and June 1998 to be $8 - 49 \times 10^{-3}$ m$^2$/s$^{-2}$.

Previous studies (Sato and Dunkerton 1997, Maruyama 1994, Vincent and Alexander 2000) have found that waves with frequencies consistent with inertio-gravity waves lead to an overall net upward transport of westerly momentum. Here we find that the net upward transport of vertical momentum due to inertio-gravity $n = 1$ waves is easterly (ie, $|\overline{u'w'}|$ for WIG1 waves $> |\overline{u'w'}|$ for
Figure 4.4 $\frac{\partial^2 v}{\partial z^2}$ due to Rossby-gravity, EIG0, EIG1, WIG1 and ER1 waves at 70hPa and 50hPa.
EIG1 waves). However we find net upward transport of momentum associated with waves with a period of \( \approx 2 \) days to be westerly because of the inclusion of fast Kelvin waves and EIG0 waves. Our study is different from many other studies in that we are able to consider inertio-gravity waves separately from other disturbances at inertio-gravity wave frequencies. This means that additional insights into the role of inertio-gravity waves in the lower stratosphere (such as the fact that they are associated with net upward transport of westerly momentum) are possible.

The next section will quantify the effects of each wave on the acceleration of the mean zonal flow between 70hPa and 50hPa. It will do this by considering the convergence of the vertical Eliassen-Palm flux between 70hPa and 50hPa.

### 4.3.4 QBO Results

The red line on figure 4.5 shows the acceleration of the 50hPa zonal wind at Singapore. The acceleration needed to drive the QBO would be about twice as large as this because of the effects of tropical upwelling. The blue line on figure 4.5 is the acceleration of the ERA15 data in the 70hPa-50hPa layer averaged over 7.5°N-7.5°S. It can be seen that the acceleration in the ERA15 data is in good agreement with observations at Singapore. The acceleration of the mean zonal flow provided by the waves detected in chapter 3 is the black line on figure 4.5. This was calculated using (4.2) and the values of \( \frac{f \bar{v}' \bar{\theta}'}{\bar{\theta} z} \) and \( \bar{u}' \bar{w}' \) for each wave type at 50hPa and 70hPa.

Though ERA15 provides a good representation of the QBO at 50hPa, the waves detected in ERA15 do not provide enough acceleration to drive either the extreme easterly or the extreme westerly phase of the QBO at any time of the dataset. This is as expected since chapter 3 only detected waves with frequency \(< 1 \) cycle per day and much of the QBO will be driven by waves of higher frequency than this (section 4.3.1). In the early part of the dataset the acceleration due to waves of frequency \(< 1 \) cycle per day, and the acceleration of the flow do appear to vary in phase, however towards the end of the dataset there is less correlation. At no time of the dataset do the waves have sufficient magnitude to drive the QBO.
Figure 4.5 Red line shows the acceleration of the 50hPa zonal wind at Singapore. Blue line shows acceleration of the mean zonal wind in ERA15, averaged over 7.5°N and 7.5°S in the 70hPa-50hPa layer. Black line shows acceleration due to waves of chapter 3, averaged over 7.5°N - 7.5°S in the 70hPa - 50hPa layer. (Note the acceleration required to drive the QBO is approximately twice that shown by the blue and red lines because of tropical upwelling).
Chapter 4

The interaction between equatorial waves and the QBO

Figure 4.5 shows that the acceleration due to equatorial waves found in this study is always westerly, while the QBO alternates between easterly and westerly acceleration. Finding westerly acceleration throughout is consistent with other studies as low frequency waves (periods 5-10 days) drive the mean flow westerly (Sato and Dunkerton, 1997), and waves with period of around 2 days also drive the mean flow westerly (Maruyama, 1994). Waves with a period of less than 1 day are not considered in our study and are not included in figure 4.5.

Figure 4.6(a) shows the acceleration of the mean zonal flow due to eastward propagating waves. The accelerations provided by Kelvin waves, EIG0 waves and EIG1 waves are shown by the black, red and green lines respectively. Figure 4.6(b) shows the acceleration of the mean zonal flow due to westward propagating waves. The accelerations provided by Rossby-gravity waves, ER1 waves and WIG1 waves are shown by the green, red and black lines respectively. It is seen that the Kelvin wave provides the strongest acceleration at all times of the year and this is why figure 4.5 shows that the acceleration due to all equatorial waves detected is westerly at all times of the dataset. Canziani and Holton (1998) showed that the peak values of acceleration due to Kelvin waves were \(0.18 \text{ms}^{-1}\text{d}^{-1}\). However these peak values occurred at 10-20hPa and Kelvin waves in their data were intermittent. Here we consider momentum flux averaged over 90 days and 70hPa-50hPa and so what is seen in figure 4.6 is much smaller than, but not inconsistent with, Canziani and Holton’s results.

Figure 4.6(b) shows that easterly acceleration is predominantly accounted for by the WIG1 wave and to a lesser extent by the ER1 wave. The Rossby-gravity wave only explains a very small amount of the easterly acceleration. This is unsurprising in light of section 4.3.1 but very surprising in light of early QBO results (e.g. Holton and Lindzen, 1972). It is noteworthy that here we find the total acceleration due to westward and eastward propagating inertio-gravity \(n = 1\) waves to be easterly. This is because more WIG1 waves are found in the data than EIG1 waves. If westward and eastward propagating unresolved inertio-gravity waves existed in the same relative proportions as WIG1 and EIG1 waves with periods of greater than 1 day then perhaps equatorial waves could drive a quantitatively correct QBO.
Chapter 4  The interaction between equatorial waves and the QBO

4.6(a) shows acceleration due to eastward propagating waves in the 70hPa-50hPa layer; black line is acceleration due to Kelvin waves, red line is acceleration due to EIG0 waves, and green line is acceleration due to EIG1 waves. 4.6(b) shows acceleration due to westward propagating waves in the 70hPa-50hPa layer; green line is acceleration due to Rossby-gravity waves, red line is acceleration due to ER1 waves and black line is acceleration due to WIG1 waves.

Although figure 4.6 provides a reasonable estimate of QBO forcing due to each type of equatorial wave, there are several issues which must be considered. Firstly, we have assumed that acceleration due to each wave type is independent of other wave types. This will probably not be the case, as Takahashi and Boville (1992) showed that significant wave-wave interactions arose by just considering Kelvin and Rossby-gravity waves in their 3-D model. These wave-wave interactions are similar to what was discussed in section 2.6.1 and occur because the eddy momentum flux and eddy heat flux are quadratic quantities. However the model of Takahashi and Boville (1992) contained waves of unrealistically large amplitude and hence is likely to overestimate wave-wave interactions in the real atmosphere. Secondly it was shown in chapter 3 section 3.4 that the idealised linear waves could not explain much of the variance in dynamical fields, particularly at those frequencies appropriate to inertio-gravity waves. Perhaps there is also additional Eliassen-Palm flux at inertio-gravity wave frequencies that cannot be explained by waves. If this were the case it would explain why the momentum flux due to the inertio-gravity waves found in ERA15 is so much smaller than the momentum flux of disturbances with periods 1-3 days that
was calculated by Sato and Dunkerton (1997).

4.4 Summary

This chapter began by considering theoretically the extent to which the activity of each equatorial wave type was likely to be correlated with the QBO at 50hPa. This was done by calculating the vertical group velocity of each wave (which is inversely proportional to the wave’s dissipation per unit vertical distance propagated) in various zonal winds. The vertical group velocity for each wave type was calculated only for that wavenumber and frequency at which maximum wave activity was detected.

All waves had largest group velocity when propagating through background zonal winds which flowed in the direction opposite to the zonal direction of wave propagation. However it was found that small changes in the background wind had a greater effect on the vertical group velocity of some waves than of other waves. Those waves which had vertical group velocity most sensitive to changes in the background wind were the ER1 wave, the Rossby-gravity wave, the WIG1 wave and the Kelvin wave. The activity of these waves at 50hPa was correlated with the QBO. The EIG0 and EIG1 waves which had vertical group velocity less sensitive to changes in the background wind were less well correlated with the QBO at 50hPa. This shows that dissipation of waves in ERA15 is qualitatively consistent with linear wave theory.

This chapter has also considered the extent to which waves in ERA15, with frequency $< 1$ cycle per day, are able to drive the QBO. It was found that dissipation of all the wave types detected in this study would accelerate the zonal flow in the direction that the wave was propagating. However the waves detected only provided a small proportion of the acceleration needed to drive the QBO. This is consistent with previous results (e.g. Giorgetta et al., 2002) which imply that many of the waves that force the QBO would be of higher frequency than has been considered. Of all the equatorial waves considered the Kelvin wave provides the strongest contribution to forcing the QBO. The westward wave with the strongest forcing is the WIG1 wave. The contribution of the Rossby-gravity wave to the QBO appears negligible. Overall the forcing from resolved
equatorial waves is eastwards, hence if the ERA15 data provides a reasonable representation of all the resolved equatorial waves the net forcing from unresolved waves would be westwards.
CHAPTER 5

The annual temperature cycle near the tropical tropopause

5.1 Introduction

Figures 5.1(a) and 5.1(b) show the mean annual cycle in temperature at Truk (7.45N, 151.85E) and Pago Pago (14.33S, 170.72W), generated using radiosonde data averaged over 1966-1997. The solid white line on these figures is the cold point tropopause (CPT). Above 120hPa there is a pronounced annual cycle in temperature, which has the same phase at both stations despite the fact that Truk is in the northern hemisphere (NH) and Pago Pago is in the southern hemisphere (SH). This annual cycle consists of coldest temperatures in NH winter and warmest temperatures in NH summer and exists throughout the tropics.

![Figure 5.1](image.png)

(a) TRUK - annual cycle  
(b) PAGO PAGO - annual cycle

**Figure 5.1** Pressure-time sections of the mean annual cycle in temperature at 5.1(a) - Truk and 5.1(b) - Pago Pago. These are produced from radiosonde data and are averaged over 1966-1997. The solid white line shows the average cold point tropopause.

Though the temperature structure near the tropical tropopause is similar over the tropics, some differences do exist. For example the cold point tropopause is colder in the NH than the SH
and this difference is amplified in the NH winter (Seidel et al., 2001). Typical magnitudes of this difference can be seen on figure 5.1. The cold point tropopause at Truk (in the Northern Hemisphere) is colder than the CPT at Pago Pago (in the Southern Hemisphere) by approximately 2K in DJF and approximately 1K in JJA. There is also longitudinal variation in temperature near the tropopause (Highwood and Hoskins, 1998). However horizontal variations in temperature near the tropical tropopause are small and so this chapter is concerned with modelling the annual temperature cycle at set pressure levels near the tropopause averaged over $10^\circ$N- $10^\circ$S.

The annual temperature cycle near the tropical tropopause has previously been attributed to the annual cycle in the strength of the tropical upwelling branch of the wave-driven stratospheric circulation (Yulaeva et al., 1994). This chapter will use the thermodynamic equation to quantitatively assess the extent to which this is the case. A quantitative understanding of what determines the temperature near the tropical tropopause could be used in the future to determine the response of the tropopause temperature to climate change and could help predict future values of stratospheric water vapour.

This mean annual temperature cycle is modelled using the thermodynamic equation and ERA15 data. However the model itself and the data inputs to the model can be changed substantially and realistic results still obtained. This chapter will compare and contrast three thermodynamic models which can reasonably predict the annual temperature cycle. These models will be fully described in section 5.2 but as a reference for the chapter they are summarised in table 5.1.

Two of the models to be compared ([A] and [B]) will use the Eulerian-mean thermodynamic equation, while the other ([C]) will use the Transformed Eulerian-Mean (TEM) thermodynamic equation (described in section 5.5.1). Section 3.6 of Andrews et al. (1987) shows that in the TEM thermodynamic equation eddy forcing terms, due to steady linear disturbances, on a conservative flow will be zero and will not induce any mean-flow changes. However in the untransformed Eulerian-mean thermodynamic equation the same disturbances will generally lead to non-zero eddy forcing terms, which will be precisely cancelled by a disturbance induced Eulerian-mean circulation. This has implications for the treatment of equatorial waves in each of the models.
Linear waves will have no effects on the TEM thermodynamic equation and so no explicit treatment of them in model [C] is required. In models [A] and [B], however, linear waves will have affected the Eulerian-circulation and hence a full representation of them is required in each model for consistency. The models differ in that model [A] assumes that the proportion of horizontal advection that is due to linear waves is the same as shown in table 3.1, while model [B] assumes that all of the horizontal advection in ERA15 is due to linear waves. Comparing results of these models will lead to insights about the wave climatologies of chapter 3 and equatorial waves in the atmosphere.
Because different inputs are used in each model, this can lead to conflicting results as to the cause of the annual temperature cycle. For example model [A] does not include any forcing from above 10hPa and attributes over one half of the annual temperature cycle to horizontal temperature advection. Models [B] and [C] include forcing from above 10hPa and are able to explain most of the annual temperature cycle by tropical upwelling. In a similar way the radiative equilibrium temperature and radiative timescale need to be calculated differently in each model (as is shown in table 5.1) for a good prediction of temperature. A comparison of the models will suggest values for the radiative timescale and radiative equilibrium temperature that are most appropriate to the atmosphere, in addition to suggesting the cause of the annual temperature cycle.

5.2 The thermodynamic model

This chapter will use three different models to predict the temperature near the tropical tropopause. There are some features of model design, and details about model forcing data, that are common to all the models. It is these common elements that will be discussed in this section. Firstly, all the models are based on the thermodynamic equation which is:

\[
\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega \left( \frac{\partial T}{\partial p} - \frac{T \kappa}{p} \right) = Q \tag{5.1}
\]

Here \( T \) is temperature, \( t \) is time, \( \omega \) is pressure vertical velocity, \( u \) and \( v \) are the zonal and meridional components of horizontal wind, \( \kappa \approx 0.286 = R/c_p, R = 287JK^{-1}kg^{-1} \) is the gas constant for dry air, \( c_p \approx 1004JK^{-1}kg^{-1} \) is the specific heat of dry air at constant pressure, and \( Q \) is radiative heating. \( \left( \frac{\partial T}{\partial p} - \frac{T \kappa}{p} \right) \) determines how the temperature will change in response to vertical motion and is dependent on the static stability of the air, hence this term will be denoted \( S \).

Equation (5.1) shows that the temperature tendency is balanced by dynamical factors and radiative heating. The dynamical factors are on the left of the equation and show how the temperature will change due to advection of warmer or cooler air. On the right of (5.1) is the radiative heating,
\( Q \). Equation (5.1) can be simplified by assuming that the radiative heating can be adequately represented by linear relaxation to some equilibrium value, \( T_{eq} \) with timescale \( \tau \):

\[
Q = (T_{eq} - T)/\tau.
\] (5.2)

Thus if \( T > T_{eq} \) the air experiences radiative cooling, while if \( T < T_{eq} \) the air experiences radiative heating.

The Eulerian mean of (5.1)

\[
\frac{\partial \overline{T}}{\partial t} + u \frac{\partial \overline{T}}{\partial x} + v \frac{\partial \overline{T}}{\partial y} + \omega \left( \frac{\partial \overline{T}}{\partial p} - \frac{T \kappa}{p} \right) = \frac{T_{eq} - \overline{T}}{\tau}
\] (5.3)

is directly used to model the temperature in two of the three models that are compared and is used to derive the third. Throughout this chapter overbars denote an average over \( 10^\circ N - 10^\circ S \) as this is the geographical region that will be modelled. To further simplify the radiative heating, \( T_{eq} \) and \( \tau \) are assumed to be constant over the area \( 10^\circ N - 10^\circ S \) such that \( Q = (T_{eq} - \overline{T})/\tau \). Results of Wehrbein and Leovy (1982) imply that the zonal mean of \( T_{eq} \) is reasonably constant in this region. It will be shown in section 5.6 that \( \tau \) is predominantly dependent on temperature and since the temperature varies little over the area \( 10^\circ N - 10^\circ S \), it is expected that \( \tau \) can also be approximated as spatially constant.

To obtain the temperature, (5.3) is discretised in time using a trapezoidal scheme and rearranged as:

\[
\overline{T}^{(n+1)} = \left( \frac{1}{\Delta t} - \frac{1}{2\tau} \right) \overline{T}^{(n)} + \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) T_{eq} - \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) \left( u \frac{\partial \overline{T}}{\partial x} + v \frac{\partial \overline{T}}{\partial y} + \omega S \right).
\] (5.4)
Chapter 5 The annual temperature cycle near the tropical tropopause

Here superscripts denote the timestep and terms with no superscript are calculated as an average of timestep \( n + 1 \) and timestep \( n \), \( \Delta t \) is the size of the timesteps. An initial temperature \( \overline{T}^{(0)} \) is assumed to calculate successive \( \overline{T}^{(n)} \) values. As \( n \) increases the effect of \( \overline{T}^{(0)} \) on \( \overline{T}^{(n)} \) is reduced such that after approximately 50 days \( \overline{T}^{(n)} \) is dependent solely on the dynamical and radiative history of the region. This means that each model has a spin up period of approximately 50 days, and hence no results will be shown for the first year that each model is run.

The dynamical inputs to (5.4) are obtained from the ERA15 dataset which is described in chapter 2. ERA15 is available for a long enough time period that information about the ‘average’ annual cycle can be obtained along with information about individual years. The inputs to (5.4) will now be examined in more detail.

5.2.1 Inputs common to all the models

5.2.1.1 Radiative parameters, \( T_{eq} \) and \( \tau \)

The radiative relaxation timescale, \( \tau \), and the radiative equilibrium temperature, \( T_{eq} \), at pressure \( p_0 \) were calculated using the radiative-convective model described in Thuburn and Craig (2002) and based on the broadband radiation scheme of Morcrette (1990). This radiative-convective model was forced with tropical profiles of temperature, water vapour and ozone, which were obtained from a three dimensional climatology, described by Freckleton and Forster (1995), based on satellite, aircraft and ground based observations. The temperature profile in the climatology for March consisted of temperatures of 195.35K (-77.8°C) for all pressures between 100hPa and 50hPa, which is not realistic when compared with figure 5.1. Therefore the temperature data in March was interpolated from that in February and April.

Radiative parameters at \( p_0 \) for each month were calculated as follows: Firstly typical tropical profiles were input into the Morcrette scheme and the heating, \( Q_1 \), calculated; secondly a Gaussian temperature perturbation of \( \Delta T \), of depth \( d \), and centred on \( p_0 \) was applied to the temperature profile and the heating recalculated as \( Q_2 \). Finally the radiative parameters are:
\[
\tau = \frac{\Delta T}{Q_1 - Q_2}
\] (5.5)

\[
T_{eq} = T + Q_1 \left( \frac{\Delta T}{Q_1 - Q_2} \right).
\] (5.6)

Here (5.6) has been calculated from (5.5) by assuming the linear relaxation to equilibrium approximation, (5.2).

Equation (5.2) is used throughout this chapter. However it is not expected to be quantitatively accurate for large departures of \( T \) from \( T_{eq} \) (Andrews et al., 1987). It is therefore important to determine whether \( |T_{eq} - T| \) in the tropical lower stratosphere is sufficiently small for (5.2) to be valid.

Figure 5.2 shows the heating, \( Q_2 \), that was calculated, when the tropical temperature profile for January was perturbed with various perturbations, \( \Delta T \), of depth 3000m, centred on 68hPa. It is seen that when small values of \( \Delta T \) are applied the heating is proportional to the perturbation, implying that the linear relaxation to equilibrium approximation is accurate. However for large temperature perturbations the linear relaxation to equilibrium approximation is less accurate. In the tropical lower stratosphere Wehrbein and Leovy (1982) found that \( |T_{eq} - T| \approx 10 \)K. Figure 5.2 shows this is small enough that the linear relaxation to equilibrium approximation can be used.

The radiative equilibrium temperature, and the radiative timescale are assumed to have no interannual variability. They are calculated once for a ‘typical’ year where temperature, water vapour, and ozone at each level take their ‘mean’ monthly value. This ‘typical’ annual cycle in radiative parameters is used every year the thermodynamic model is run. In reality there are interannual variations in these radiative parameters because there are interannual variations in ozone (Shiotani and Hasebe, 1994) water vapour (Geller et al., 2002) and temperature (Reid, 1994) due to factors such as the QBO, ENSO, and volcanoes. However section 5.6 will show that
Chapter 5  The annual temperature cycle near the tropical tropopause

Figure 5.2  Shows the heating, $Q_2$, that is calculated by the Morcrette scheme when the temperature is perturbed by various amounts. All temperature perturbations are of depth 3000m and centred on 68hPa.

It is the variations in temperature that are likely to have most effect on the radiative parameters. Since interannual temperature variations are smaller than the annual temperature variations (Randel et al., 2000), using a constant annual cycle in radiative parameters is reasonable.

The radiative heating after the atmosphere is perturbed, $Q_2$, is strongly dependent on the depth of the perturbation, $d$. This means that different depths of perturbation will lead to different values of radiative equilibrium temperature and radiative timescale. Figure 5.3 shows the values of $T_{eq}$ and $\tau$ that were calculated for each month at 68hPa using depths of perturbation 1000m, 2000m, 3000m, 4000m and 5000m. The smaller the depth of perturbation the lower the radiative equilibrium temperature and radiative timescale. It is seen that the radiative timescale is particularly dependent on the depth of atmosphere that is perturbed, with the radiative timescale being over twice as large for $d = 5000m$ than for $d = 1000m$. It is impossible to tell, a priori, which depth of perturbation is appropriate to the real atmosphere. Therefore the radiative parameters are calculated using the depth of perturbation that gives the best results of each thermodynamic model. Though this does amount to ‘tuning’ of the models, this tuning is limited because the depth of perturbation used in each model is constrained to be constant throughout the year. The depth of perturbation that provides good model predictions is likely to be more relevant to the
Figure 5.3  a) shows the radiative equilibrium temperature and b) shows the radiative timescale, for each month. The values are calculated using the following depths of perturbation: 1000m - black crosses, 2000m - red stars, 3000m - green diamonds, 4000m - blue triangles, and 5000m purple squares. Thin solid lines are added for clarity.
The dynamical inputs to (5.4) are the horizontal advection terms \( \frac{\partial u}{\partial x} \) and \( \frac{\partial v}{\partial y} \) and the vertical upwelling term \( \overline{\omega S} \). The horizontal advection terms are treated differently in each of the three models used and hence will be discussed later. The vertical upwelling term can be decomposed as:

\[
\overline{\omega S} = \overline{\omega S} + \overline{\omega' \frac{\partial T'}{\partial p}} - \frac{\overline{\omega' T' \kappa}}{P}
\]  

(5.7)

where primed quantities denote departures from the mean. The term \( \overline{\omega' T' \frac{\kappa}{P}} \) is difficult to obtain directly from data and is equal to zero when disturbances are linear equatorial waves on a background state with no zonal wind. It is therefore assumed that this term will not significantly affect the temperature predicted and it is set to zero in each model. The value of \( \overline{\omega' \frac{\partial T'}{\partial p}} \) will depend on how equatorial waves are treated in each of the methods and will be discussed later. However \( \overline{\omega} \) and \( \overline{S} \) will be calculated the same in each of the methods and the source of these values is now considered.

The mean upwelling, \( \overline{\omega} \), across a pressure surface, \( p_0 \), has been calculated as follows. The poleward mass flux across 10°N and 10°S above \( p_0 \) is obtained. Mass continuity implies this will be equal to the mass flux crossing \( p_0 \). The mass flux across \( p_0 \) can then be used to calculate the average pressure vertical velocity, \( \overline{\omega} \), at \( p_0 \). In each model, \( \overline{\omega} \) is obtained from the meridional velocity in ERA15 data, which is available at 6 hour timesteps.

ERA15 data has insufficient vertical resolution to provide a good estimate of \( \overline{S} \), hence radiosonde data at Truk (7.45N,151.85E) is used. Though \( S \) at Truk is not identical to \( S \) everywhere in
the tropics, it should provide a good first estimate since the annual temperature cycle is similar throughout the tropics (see section 5.1). Daily values of $S$ at Truk were calculated between 1951 and 1997, and these values were used to estimate a ‘mean’ annual cycle in $S$. This ‘mean’ annual cycle in $S$ was used every year that each thermodynamic model was run. This is more reasonable than using daily values of $S$ at Truk, as there are large day to day variations in $S$ at Truk which are unlikely to be representative of everywhere in the tropics.

All remaining inputs to (5.4) will be dependent on which of the three thermodynamic models is run. Each model will be discussed in turn and the ability of each model analysed by considering the extent to which it can predict the annual temperature cycle at 70hPa, averaged over $10^\circ$N - $10^\circ$S.

### 5.3 Model [A]

#### 5.3.1 Description

Section 5.2.1.2 explained that the vertical upwelling, $\bar{\omega}$, at $p_0$ was calculated from ERA15 data using mass flux out of the $10^\circ$N - $10^\circ$S region above $p_0$. Here it is assumed that $\bar{\omega}$ is adequately represented by considering mass flux out of the $10^\circ$N - $10^\circ$S region up to 10hPa, which is the upper limit of the ERA15 dataset. In reality there is likely to be significant upwelling generated by mass flux out of the tropics in the upper stratosphere and mesosphere (Holton et al., 1995), yet it is desirable to assess the extent to which the temperature can be modelled using the ERA15 poleward mass flux.

The dynamical terms in (5.4) which have not already been considered are $\frac{\partial T}{\partial x}$, $\frac{\partial T}{\partial y}$, and $\frac{\partial T'}{\partial p}$. It was seen in chapter 3 that a significant proportion (between 30% and 70%) of $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$ can be attributed to equatorial linear waves. Model [A] assumes that these climatologies are accurate, and further assumes that the remainder of $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$ is not due to any form of equatorial wave. The term $\frac{\partial T}{\partial p}$ cannot be obtained directly from ERA15 because of poor vertical resolution. However this term will be non-zero in the linear equatorial waves detected.
in chapter 3. It is assumed that the total value of this term is due to the linear equatorial waves detected because this thesis has not considered any other physical processes which would affect it.

Though linear waves can cause large values of \( \frac{\partial T}{\partial x} \), \( \frac{\partial T}{\partial y} \) and \( \omega' \frac{\partial T'}{\partial p} \), there is significant cancellation between these terms so that \( \frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \omega' \frac{\partial T'}{\partial p} \) is small. In a Kelvin wave \( \frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \omega' \frac{\partial T'}{\partial p} = 0 \), however for other types of wave this is not necessarily the case. From (5.4) it can be seen that non zero \( u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega' \frac{\partial T'}{\partial p} \) will lead to a change in temperature. This wave induced temperature change will be precisely cancelled by a wave induced Eulerian-mean circulation (see section 5.1). Since the wave induced Eulerian-mean circulation is included in model [A], the wave induced \( u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega' \frac{\partial T'}{\partial p} \) must be included in model [A] for consistency. The wave induced \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial y} \) are obtained directly from the wave climatologies of chapter 3. The wave induced \( \omega' \frac{\partial T'}{\partial p} \) has been calculated from the wave climatologies of chapter 3 using linear wave theory.

The proportion of \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial y} \) which chapter 3 has not attributed to linear waves will hereafter be denoted the ‘no-wave horizontal advection’. The no-wave horizontal advection will be calculated and input directly into (5.4).

Inputting the details specific to model [A] into (5.4) means that the temperature can be modelled using:

\[
T^{(n+1)} = \left( \frac{1}{\Delta t} - \frac{1}{2\tau} \right) T^{(n)} + \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) T_{eq} - \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) \left( A + B + \omega |_{p_0=10hPa}^{10hPa} \right) 
\]

(5.8)

where \( A \) is the no-wave horizontal advection, \( B \) is \( u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega' \frac{\partial T'}{\partial p} \) due to the linear waves found in chapter 3, \( p_0 \) is the pressure level being considered and \( \omega |_{p_0=10hPa}^{10hPa} \) means that \( \omega \) is calculated from the mass flux, poleward of 10°, between \( p_0 \) and 10hPa only.
Equation (5.8) is now used to model the mean temperature at 70hPa.

### 5.3.2 Mean results at 70hPa from model [A]

The radiative timescale used in model [A] was calculated as described in section 5.2.1.1, using a depth of perturbation of 5000m (shown in figure 5.3(b)). This radiative timescale allowed the most accurate prediction of the annual temperature cycle. It was found, however, that using the radiative equilibrium temperature, $T_{eq}$ corresponding to a perturbation depth of 5000m did not lead to a good temperature prediction. Instead using this value of $T_{eq}$ caused the model to overestimate the ERA15 temperature by approximately 5K. For this reason it was decided that the value of $T_{eq}$ to be used in this model be calculated from the radiatively determined temperature of Shine (1987). This value of $T_{eq}$ is shown by the purple line in figure 5.4. This is warmer and has a different annual variation to the radiative equilibrium temperatures derived from the Morcrette scheme (shown in figure 5.3(a)). The parameters used in model [A] are summarised in table 5.1. These lead to a modelled temperature at 70hPa, averaged over 1981-1993 and 10°N - 10°S as shown by the red line in figure 5.4. It can be seen that the modelled temperature compares well with the ERA15 temperature, averaged over 1981-1993 and 10°N - 10°S, which is shown by the black line on figure 5.4. The radiatively determined temperature of Shine (1987), which was used in model [A] might not be quantitatively accurate near 70hPa (pers. comm. K. Shine) as it was produced from a model which was designed for lower pressures. However it can be seen from figure 5.4 that these radiatively determined temperatures allowed reasonable temperature predictions from model [A] and hence were used in preference to choosing an arbitrary value of the radiative equilibrium temperature.

It can be seen from figure 5.4 that model [A] is able to predict the 10-17K cooling below the radiative equilibrium temperature used and also the amplitude of the annual temperature cycle at 70hPa. However the phase of the modelled annual temperature cycle lags that of ERA15. The contribution of the terms in (5.8) to the average annual temperature is now considered.

Because (5.8) is linear, it can be decomposed, to find the impact of different processes on the
Chapter 5  The annual temperature cycle near the tropical tropopause

Figure 5.4  The black solid line is temperature at 70hPa averaged over 1981-1993. The red dashed line is modelled temperature at 70hPa averaged over 1981-1993. The purple dash-dotted line (top) is the radiatively equilibrium temperature used (calculated from Shine (1987)). All curves are averaged over 10°N - 10°S.

temperature as:

\[ T_{1}^{(n+1)} = C_1 T_{1}^{(n)} - C_2 \bar{\omega}^{10hPa}_{70hPa} S \]  \hspace{1cm} (5.9)

\[ T_{2}^{(n+1)} = C_1 T_{2}^{(n)} - C_2 A \]  \hspace{1cm} (5.10)

\[ T_{3}^{(n+1)} = C_1 T_{3}^{(n)} - C_2 B \]  \hspace{1cm} (5.11)

\[ T_{rad}^{(n+1)} = C_1 T_{rad}^{(n)} - C_2 T_{eq}. \]  \hspace{1cm} (5.12)

Here \( C_1 = \frac{1}{\Delta t} - \frac{1}{\tau} \) and \( C_2 = \frac{1}{(\frac{1}{\Delta t} + \frac{1}{\tau})} \), \( A \) is the no-wave horizontal advection, and \( B \) is due to the linear waves found in chapter 3. The temperature can be written as:
\[ \mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{\text{rad}} \]  

(5.13)

where \( \mathcal{T}_1, \mathcal{T}_2 \) and \( \mathcal{T}_3 \) are the temperature changes due to vertical upwelling, no-wave horizontal advection and advection terms due to linear waves respectively. \( \mathcal{T}_{\text{rad}} \) is the radiatively determined temperature. The radiative timescale, \( \tau \), is included in \( C_1 \) and \( C_2 \) and thus appears in all the equations. \( \tau \) alone, does not affect the temperature, but instead acts to amplify the effects of other terms. The effects of \( \tau \) will be examined in section 5.6, however we first consider the contribution to temperature from the dynamical terms.

### 5.3.2.1 Contribution from each dynamical term to the temperature at 70hPa

The annual cycle in tropical lower stratospheric temperatures has been related to annual variations in tropical upwelling since Reed and Vleck (1969). It is therefore logical to consider first the contribution of mean tropical upwelling to the temperature at 70hPa as predicted by model [A]. This is calculated using (5.9) and is shown in figure 5.5. It can be seen that the mean tropical upwelling cools the temperature below its radiatively determined value and also this term will impose an annual cycle on the temperature of a similar phase to what is observed (figure 5.4). However the temperature in the ERA15 data is 10-17K below its radiatively determined value (figure 5.4) of which only 2-7K can be accounted for by the mean upwelling term; also the annual temperature cycle in ERA15 is nearly twice as large and not always of the same phase (see day 60) as what would be predicted considering only the mean upwelling.

The ‘upwelling’ contribution to temperature is dependent on the mean vertical velocity \( (\overline{\omega}) \) and the static stability \( [S = (\partial T/\partial p - T\kappa/p)] \). The average annual cycle in each of these terms is shown in figure 5.6. It can be seen that while there is a significant annual cycle in \( \overline{\omega} \), \( \overline{S} \) is fairly constant and varies only by around 12% throughout the year. The temperature tendency varies in phase with \( \overline{\omega} \) and the magnitude of the \( \overline{\omega} \) contribution to \( \partial T/\partial t \) is amplified by \( \overline{S} \). It is also noteworthy that there is a very large day to day variation in \( \overline{\omega} \), that can even be seen in
Chapter 5  The annual temperature cycle near the tropical tropopause

Figure 5.5  The contribution of mean tropical upwelling to the temperature of $10^\circ N - 10^\circ S$ at 70hPa. Averaged over 1981-1993.

The unsmoothed upwelling averaged over 1981-1993 (black line in figure 5.6(a)). This suggests that the upwelling is not due to some smoothly varying force and is instead forced episodically. This is consistent with the large variability in the mid-latitude wave breaking which drives the stratospheric circulation that was discussed by Randel et al. (2002).

Figure 5.6  a) tropical upwelling at 70hPa, averaged over 1981-1993 and $10^\circ N - 10^\circ S$ forced between 70hPa and 10hPa. The red line is 10 day running mean. b) static stability ($S = \partial T/\partial p - T \kappa/p$) at 70hPa calculated from radiosonde data at Truk averaged over 1951-1997. A running mean of 3 days is presented for static stability.

The cooling below the radiatively determined temperature which cannot be attributed to the mean
upwelling is due to $\frac{\partial u}{\partial x}$, $\frac{\partial v}{\partial y}$ and $\frac{\partial \omega}{\partial p}$ which has been decomposed into no-wave horizontal advection and advective fluxes due to linear equatorial waves.

The contribution from the linear wave induced $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p}$ is included to balance the linear wave induced Eulerian-mean circulation ($\overline{v}, \overline{w}$). The linear wave induced $\frac{\partial u}{\partial x}, \frac{\partial v}{\partial y}$ and $\frac{\partial \omega}{\partial p}$ affect the temperature as shown in figures 5.7(a), 5.7(b) and 5.7(c) respectively. Though the effects of linear waves on a single advection term can cause substantial temperature deviations, the effects of linear waves on $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p}$ only affect the temperature as shown in figure 5.8(a). The small cooling shown in figure 5.8(a), will be exactly balanced by a warming due to changes in ($\overline{v}, \overline{w}$) which will have been included in figure 5.5. Hence the existence of linear waves means that figure 5.5 slightly underestimates the effect of the mean upwelling on temperature. The total amount of this underestimation is small and its effect on the amplitude of the annual temperature cycle is smaller still. Therefore if linear equatorial waves were not included in (5.8) the modelled temperature would not have been greatly different.

There is a proportion of $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial y}$ which cannot be accounted for by linear equatorial waves, termed the ‘no-wave horizontal advection’. The contribution to temperature at 70hPa from this term is shown in figure 5.8(b). It is clear that the no-wave horizontal advection explains twice as much cooling below the radiatively determined temperature as the mean upwelling. The contribution from the no-wave horizontal advection has its minimum value around day 60, which is where the contribution from the mean upwelling has its secondary maximum. It can be seen that in model [A] the mean upwelling and the horizontal advection are both necessary in order to predict a reasonable temperature cycle. However in sections 5.4.2 and 5.5.2 it will be seen that alternative models can lead to very different conclusions.

In this section the ‘mean’ temperature at 70hPa has been modelled by assuming that the vertical velocity forced between 70hPa and 10hPa in ERA15 data was reasonable and also that the wave climatology of chapter 3 was able to detect all equatorial wave disturbances in the atmosphere. The mean annual cycle is quantitatively well predicted, yet there are some issues regarding the accuracy of this method which must be discussed. The first issue concerns the very important
no-wave horizontal advection term. There is no obvious physical process other than equatorial waves to which this term can be attributed, and furthermore it follows a very similar annual cycle to the proportion of horizontal advection due to linear waves. It is therefore likely that the no-wave horizontal advection may be due to equatorial waves which could not be detected in chapter 3 such as waves in a vertical shear zone. If this were the case then the no-wave horizontal advection in figure 5.8(b) would be largely counteracted by the term $\omega' \frac{\partial T'}{\partial p}$ and would therefore not be able to provide the additional cooling around day 60 that is required. In addition chapter 3 showed that the dominant wave type, the Kelvin wave, was influenced by the annual cycle of the tropopause at 70hPa. If horizontal advection due to undetected Kelvin waves is included in the no-wave horizontal advection term of the model but vertical advection due to undetected Kelvin waves is not included, then large errors of the same phase of the annual temperature cycle will occur in the model. This would necessarily lead to a reasonably accurate temperature prediction.
produced by a model with largely inaccurate forcing. Further doubts arise as to the validity of the method when we note that the radiatively determined temperature that was required is different to that which was predicted by the Morcrette scheme (section 5.2.1.1) and may not be quantitatively accurate. Finally the mean vertical velocity used in this model was forced up to 10hPa only. This is not in agreement with the fact that much extratropical wave driving is expected to act on the upper stratosphere and mesosphere (e.g. Holton et al., 1995).

5.4 Model [B]

5.4.1 Description

Model [B] is able to reasonably predict the annual temperature cycle at 70hPa, without many of the assumptions and uncertainties that exist in model [A]. However some different assumptions and uncertainties are required. The no-wave horizontal advection of model [A] has no obvious physical cause except for equatorial waves which have not been detected in chapter 3. It is therefore assumed, in model [B], that $u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega' \frac{\partial T}{\partial z}$ is due only to equatorial waves.
Figure 5.8(a) suggests that if this were the case then the value of \( \frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \omega \frac{\partial T'}{\partial x} \) would be small. Therefore this term is omitted from model [B] even though there would be some wave induced Eulerian-mean circulation included. The errors associated with the omission of \( \frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \omega \frac{\partial T'}{\partial x} \) in model [B] are quantified in section 5.4.2.1.

Another difference between model [B] and model [A] is that in model [B] the pressure vertical velocity is due to mass flux out of the tropics between the level of interest, \( p_0 \), and the top of the atmosphere. This is expected to be more accurate than considering only to 10hPa. There are now two stages to calculating vertical velocity, the first is calculated from the ERA15 mass flux poleward of 10°N and 10°S up to 10hPa and the second is calculated from the mass flux crossing 10hPa. The vertical velocity forced from above 10hPa is estimated from the results of Seol and Yamazaki (1999) who calculated tropical mass flux across various pressure levels in NCEP data between 1985-1995. The mass flux across 10hPa between 10°N and 10°S can not be obtained directly from Seol and Yamazaki (1999) as, like many other studies of the stratospheric circulation (e.g. Rosenlof and Holton 1993, Eluszkiewicz et al. 1996), Seol and Yamazaki (1999) defined the tropics as that region of the stratosphere where there is mean upwelling. To calculate mass flux across 10hPa between 10°N and 10°S we make the assumption that the latitudinal distribution of tropical upwelling is the same at 10hPa as it is at 100hPa. Seol and Yamazaki (1999) (their figure 8) present values of mass flux out of the tropics between 100hPa and 10hPa, \( mass_{100}^{10} \), and values of tropical mass flux crossing 10hPa, \( mass_{10} \). The ratio \( \frac{mass_{10}}{mass_{100}^{10}} \) is 0.35 in DJF, 0.27 in MAM, 0.42 in JJA and 0.27 in SON. These ratios are used, along with the ERA15 mass flux crossing 100hPa to estimate the mass flux crossing 10hPa. The exact mass flux crossing 10hPa is assumed to be the same every year and is calculated from the average annual cycle in ERA15 mass flux crossing 100hPa. The mass flux crossing 10hPa will increase the vertical velocity at each lower stratospheric level of interest.

For this method both the radiative timescale, \( \tau \), and the radiative equilibrium temperature, \( T_{eq} \), were obtained from the Morcrette scheme as described in section 5.2.1.1. The perturbation depth that gave most reasonable results was 3000m.
Inputting model [B] details into (5.4) means that the temperature is modelled using:

\[ T^{(n+1)} = \left( \frac{1}{\Delta t} - \frac{1}{2\tau} \right) T^{(n)} + \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) T_{eq} - \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) (\bar{\omega} S). \] (5.14)

The ability of this model to predict the mean temperature at 70hPa is now discussed.

### 5.4.2 Mean results at 70hPa from model [B]

Figure 5.9 shows the temperature averaged over 1981-1993 at 70hPa as predicted by (5.14). The black line is the ERA15 temperature averaged over \(10^\circ N - 10^\circ S\), the red line is the temperature predicted using (5.14) and the purple line is the radiative equilibrium temperature calculated with perturbation depth of 3000m. Since \(u \frac{\partial T}{\partial x}\), \(v \frac{\partial T}{\partial y}\) and \(\omega' \frac{\partial T'}{\partial p}\) have not been included in this model the difference between the radiative equilibrium temperature and the modelled temperature is due entirely to the cooling effects of the mean upwelling. The mean upwelling can therefore quantitatively account for most of the annual temperature cycle at 70hPa, particularly in the northern hemisphere summer and autumn.

Figure 5.10(a) and 5.10(b) show the cooling below radiative equilibrium attributable to upwelling forced between 70hPa and 10hPa and upwelling forced from above 10hPa respectively. The first of these figures represents the temperature change due to the vertical velocity calculated directly from ERA15 data while the second of these figures represents the temperature change due to the parameterised upwelling. It is seen that that the forcing between 70hPa and 10hPa and the forcing above 10hPa contribute similar amounts to the temperature at 70hPa, although the forcing from above 10hPa is slightly dominant. There is a secondary maximum in temperature around day 80 in both of the forcings; however this maximum is relatively larger in the 70hPa - 10hPa forcing layer. If we note the importance of the forcing from above 10hPa and the assumptions made in calculating this term the model provides a good prediction of the temperature particularly between day 150 and day 300. The results of this section show that at most times of the year the temperature can be reasonably modelled without the ‘no-wave horizontal advection’ that
Figure 5.9  The black solid line is temperature at 70hPa averaged over 1981-1993. The red dashed line is modelled temperature at 70hPa averaged over 1981-1993. The purple dash-dotted line (top) is a typical annual cycle in the radiatively determined temperature. All curves are averaged over $10^\circ N - 10^\circ S$.

was vital to model [A], provided a measure of the upwelling forced from above 10hPa is included.

Figure 5.10  The contribution to temperature due to upwelling forced from: a) between 70hPa and 10hPa, and b) above 10hPa.
5.4.2.1 Errors due to the omission of advection terms

Model [B] implicitly includes the effects of equatorial waves on the Eulerian-mean circulation but does not include the effects of equatorial waves on $\frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \frac{\partial T'}{\partial p}$. This will lead to an error in the model and the likely size of this error is now assessed. Figure 5.8(a) in section 5.3.2.1, showed the temperature change due to the effect of linear equatorial waves on $\frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \frac{\partial T'}{\partial p}$ for model [A]. This is slightly larger than the same term in model [B], shown in figure 5.11, since model [B] uses a smaller radiative timescale. Figure 5.11 is essentially the error in model [B] by including only some of the effects of the linear waves found in chapter 3. It is seen that the error in the predicted temperature caused by this omission is small, in absolute terms, and the error in the amplitude of the predicted annual cycle caused by this omission is even smaller. In model [B] the whole of $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$ are assumed to be due to equatorial waves, but only those waves that have been detected in chapter 3 can be quantified. If the waves that have not been detected were assumed to affect the temperature in the same way as those waves which have been detected, then the change in temperature due to $\frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \frac{\partial T'}{\partial p}$ would be at most three times what is shown in figure 5.11. Model [B] will therefore overestimate the temperature by up to 1.5K and underestimate the amplitude of the annual temperature cycle by 0.45K for this region.

Figure 5.11 The contribution of equatorial waves to the change in temperature due to the terms $\frac{\partial T}{\partial x} + \frac{\partial T}{\partial y} + \frac{\partial T'}{\partial p}$ at 70hPa averaged over $10^\circ N - 10^\circ S$ and 1981-1993.
Chapter 5  The annual temperature cycle near the tropical tropopause

5.5 Model [C]

5.5.1 Description

Sections 5.3.2 and 5.4.2 have shown that the treatment of equatorial waves in the Eulerian-mean form of the thermodynamic equation affects the relative importance of different mechanisms in producing the modelled temperature. Chapter 3 was able to obtain a proportion of \( u \frac{\partial T}{\partial x} \) and \( v \frac{\partial T}{\partial y} \) that was due to linear equatorial waves. However, contrasting models [A] and [B] suggest that it is perhaps more reasonable to assume that the whole of these terms are due to equatorial waves, some of which cannot adequately be represented by the wave theory described in chapter 2.

If the Transformed Eulerian-Mean (TEM) form of the thermodynamic equation is used, the question of how equatorial waves should be treated is somewhat redundant. This is because in the TEM form of the equations of motion, steady, linear, conservative disturbances to a flow will have no effect on the mean circulation. The TEM thermodynamic equation can be obtained from the Eulerian-mean thermodynamic equation, (5.3) as will now be shown.

Splitting (5.3) into a mean and disturbance part gives:

\[
\frac{\partial T}{\partial t} + \frac{\partial T}{\partial y} + \frac{\partial T'}{\partial x} + \frac{\partial T'}{\partial y} + \frac{\partial T'}{\partial p} + \frac{\partial T'}{\partial p} + \frac{\kappa T \omega}{\rho} = T_{eq} - \frac{T}{\tau}. \tag{5.15}
\]

Mass continuity implies that \( \frac{\partial (\bar{u} + u')}{\partial x} + \frac{\partial (\bar{v} + v')}{\partial y} + \frac{\partial \bar{\omega}}{\partial p} + \frac{\partial \omega'}{\partial p} = 0 \). This is added to (5.15). The fact that \( \frac{\partial}{\partial x} = 0 \) is also used to give:

\[
\frac{\partial T}{\partial t} + \frac{\partial T}{\partial y} + \frac{\partial (\bar{u}u')}{\partial y} + \frac{\partial T'}{\partial p} + \frac{\partial (\omega')T}{\partial p} + \frac{\kappa T \omega}{\rho} = T_{eq} - \frac{T}{\tau}. \tag{5.16}
\]

Now the residual vertical velocity is defined as:
This represents that part of the mean vertical velocity whose contribution to adiabatic temperature change is not cancelled by the eddy heat flux divergence (Holton, 1992) and approximates mean Lagrangian motion under steady state conditions (Dunkerton, 1978).

Substituting from (5.17) into (5.16) gives:

\[
\frac{\partial T}{\partial t} + \nabla \cdot \mathbf{u} + \omega^\ast \mathbf{S} + \frac{\partial}{\partial p} \langle \omega^\ast T' \rangle = \frac{T_{eq} - T}{\tau}.
\] (5.18)

The term \( \frac{\partial}{\partial p} \langle \omega^\ast T' \rangle \) is omitted as this is zero under linear conservative conditions (Andrews et al., 1987). The term \( \nabla \cdot \mathbf{u} \) is also omitted as this was found to be small in the region of interest by Rosenlof (1995). Hence the Transformed Eulerian-Mean form of the thermodynamic equation, which is used to obtain the temperature for model [C], becomes:

\[
\frac{\partial T}{\partial t} + \omega^\ast \mathbf{S} = \frac{T_{eq} - T}{\tau}.
\] (5.19)

This means that in the TEM thermodynamic model, the temperature is dependent only on the residual vertical velocity, the static stability and the radiative parameters. Equation (5.19) is solved as:

\[
T^{(n+1)} = \left( \frac{1}{\Delta t} - \frac{1}{2\tau} \right) T^{(n)} + \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) T_{eq} - \left( \frac{1}{\Delta t} + \frac{1}{2\tau} \right) \langle \omega^\ast \mathbf{S} \rangle
\] (5.20)

where all terms are as previously defined. For this model the residual pressure vertical velocity,
\( \overline{\omega} \), is calculated using the mass flux out of the tropics from the pressure level of interest to the top of the atmosphere. The mass flux out of the tropics above 10hPa has been parameterised in the same way as it was in model [B]. Equation (5.20) is now used to predict the mean temperature at 70hPa.

### 5.5.2 Mean results at 70hPa from model [C]

This is not the first time the TEM thermodynamic equation has been used to provide information about the temperature at 70hPa. Rosenlof (1995) considered the relative values of all the terms in the TEM thermodynamic equation and showed that radiative heating was largely balanced by vertical temperature advection and the potential temperature tendency \( \frac{\partial \theta}{\partial t} \) was a small residual. However Rosenlof (1995) calculated the residual circulation \((\overline{\omega}, \overline{\omega})\) from heating rates computed using data from the Upper Atmosphere Research Satellite (UARS) and so they do not obtain terms in the thermodynamic equation independently.

In our study the residual circulation has been obtained from ERA15 data and the radiative heating parameters have been independently calculated from a model. The terms of the thermodynamic equation are therefore not constrained to balance. If temperature can be predicted using these independent estimates of inputs to (5.20) then we can be more confident about the accuracy of the residual circulation and the radiative parameters that are used.

On figure 5.12, the solid black line is the ERA15 temperature averaged over 1981-1993 and \(10^\circ N - 10^\circ S\), the dotted red line is the temperature obtained using (5.20) of model [C], averaged over the same spatial region and time period. The radiative parameters of figure 5.3 used here were those calculated with a perturbation depth of 1000m, and the radiative equilibrium temperature used is the purple line on figure 5.12. The radiative timescale using this perturbation depth is around 15 days, which is smaller than the 20 days suggested by Holton et al. (1995) and Haynes et al. (1991). Throughout the ‘mean’ year, except between days 30 and 130, the TEM model predicts temperatures of 1-2K below what is observed. However the annual cycle is very well predicted. Using this model about half of the annual temperature cycle can be attributed.
to annual variations in the radiative equilibrium temperature, while the remainder is due to the
effects of residual upwelling, $\omega^*$. 

![Figure 5.12](image)

**Figure 5.12** The black solid line is temperature at 70hPa averaged over 1981-1993. The red
dashed line is modelled temperature at 70hPa averaged over 1981-1993. The purple dash-dotted
line (top) is a typical annual cycle in the radiatively determined temperature calculated with depth
of perturbation 1000m. All curves are averaged over $10^\circ N - 10^\circ S$. 

Figure 5.13 is the same as figure 5.12 but with radiative parameters calculated with a depth
of perturbation of 3000m. Using a depth of perturbation of 3000m for model [C] gives a
less accurate temperature prediction than a depth of perturbation of 1000m. However using
3000m allows a clearer comparison with model [B]. Comparing model results using different
perturbation depths will help assess the sensitivity of the model to how radiation is calculated.

Figure 5.3 shows that increasing the depth of perturbation will increase the radiative equilibrium
temperature, which will increase the modelled temperature, while at the same time it will
increase the radiative timescale which will decrease the modelled temperature. Hence it is not
immediately obvious whether increasing the perturbation depth will increase or decrease the
modelled temperature. As a guide, the effect of $\omega^*S$ on the temperature is amplified by a factor
which is dependent on the radiative timescale, while the radiative equilibrium temperature does
Figure 5.13  The black solid line is temperature at 70hPa averaged over 1981-1993. The red dashed line is modelled temperature at 70hPa averaged over 1981-1993. The purple dash-dotted line (top) is a typical annual cycle in the radiatively determined temperature calculated with depth of perturbation 3000m. All curves are averaged over 10° N – 10° S.

It has been seen that using a perturbation depth of 1000m in model [C] allows an accurate representation of the temperature for a mean year at 70hPa. Using a perturbation depth of 3000m can reproduce the annual variation in temperature but induces a cold bias of up to 7K. If the perturbation depth is 1000m then half of the annual temperature cycle is due to annual variations in radiative equilibrium temperature and half of the annual temperature cycle is due to the residual upwelling. If the perturbation depth is 3000m then all of the annual temperature
cycle can be attributed to variations in residual upwelling. Therefore the depth of perturbation used will partially determine what causes the annual temperature cycle at 70hPa in the model. Unfortunately the depth of perturbation which is most appropriate to the real atmosphere is unknown, and so the effect of upwelling on temperature for both perturbation depths is considered.

Figure 5.14(a) and 5.14(b) show the contribution to temperature from the Eulerian-mean upwelling, $\bar{\omega}$, between 70hPa and 10hPa and above 10hPa respectively. In each figure the red line is the contribution to temperature when the radiative timescale was calculated with depth of perturbation 3000m and the black line is the contribution to temperature when the radiative timescale was calculated with depth of perturbation of 1000m. (Note the red line is identical to what was used in model [B]). It is clearly seen that while the shape of the annual cycle cannot be changed by varying the perturbation depth, the amplitude of the annual cycle and the absolute cooling attributable to each term is relatively tunable. For both cases the upwelling forced from above 10hPa is slightly more important than the upwelling forced between 70hPa and 10hPa.

Recall that the residual upwelling, $\omega^*$, is related to the Eulerian upwelling, $\bar{\omega}$, by:

$$\omega^* = \bar{\omega} + \frac{\partial}{\partial y} \left( \frac{v'T' S}{S} \right)$$

(5.21)

So in the TEM thermodynamic equation the temperature will depend on the Eulerian upwelling, $\bar{\omega}$, but there will be an additional contribution to temperature from the term $\frac{\partial}{\partial y} \left( \frac{v'T' S}{S} \right)$. Following Holton (1992), $\frac{\partial}{\partial y} \left( \frac{v'T' S}{S} \right)$, is hereafter referred to as the ‘eddy heat flux divergence term’. The contribution to temperature from the eddy heat flux divergence term is shown in figure 5.14(c), where again the red line is obtained using a depth of perturbation of 3000m and the black line is obtained using a depth of perturbation of 1000m. In the NH summer this term explains more of the cooling below radiative equilibrium than the Eulerian upwelling, and in the NH winter the cooling due to this term is approximately 1/4 of that due to Eulerian upwelling. The temperature change due to eddy heat flux divergence term is important. This term represents the difference
between model [B] and model [C], and it was the introduction of this term in model [C] which led to the necessary reduction of the perturbation depth used to calculate the radiative parameters.

Figure 5.14  *Black lines are temperature changes when radiative parameters are calculated using a depth of perturbation 1000m, red lines are temperature changes when radiative parameters are calculated using a depth of perturbation of 3000m.*

Figure 5.14(c) shows that the temperature change due to the heat flux divergence term is between 1 and 2K for perturbation depth 1000m, and between 2 and 3K for perturbation depth 3000m. This difference does not have a clear annual cycle.

Section 5.4.2, which modelled the temperature using the Eulerian-mean thermodynamic equation (model [B]), found that when the perturbation depth was 3000m the maximum cooling due to the proportion of $u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega \frac{\partial T}{\partial p}$ attributable to equatorial waves was 1.5K. This was thought to
be the contribution to temperature that could not be explained by the Eulerian upwelling. Figure 5.14(c) shows that this contribution to temperature is up to twice as large as what was predicted from the linear wave climatologies. This means that the linear waves found in chapter 3 are not typical of disturbances in the atmosphere. Instead the disturbances that have not been accounted for would have a relatively larger effect on the components of the thermodynamic model than the disturbances consistent with idealised linear waves in a zero background flow. If more realistic background conditions were included in the linear wave theory of chapter 2 then equatorial waves could be detected which were consistent with the observed background state. These equatorial waves may affect the temperature in a way more consistent with the eddy heat flux divergence term than when a zero background flow was assumed. These results highlight the importance of extending linear wave theory to find equatorial waves more representative of the real atmosphere. A first step would be to estimate the contribution to temperature of $u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + \omega' \frac{\partial T'}{\partial p}$ that is due to equatorial waves in a vertical shear, and compare this with the temperature contribution due to the eddy heat flux divergence term in the TEM equation.

To conclude this section we consider which depth of perturbation should be used to calculate the radiative parameters. Though the annual temperature cycle can be reasonably predicted using either a perturbation depth of 3000m or 1000m, the absolute value of the predicted temperature is much closer to what is observed when a perturbation depth of 1000m is used. For the TEM thermodynamic equation in this study the 1000m perturbation depth is assumed to be the most appropriate.

### 5.6 Considering the radiative timescale

In the three models that have been used to predict the temperature, the effects of dynamics in the models have been relatively easy to determine. The role of the radiative equilibrium temperature in the models has also been clear. Equation (5.4) shows that the temperature will also be influenced by the radiative timescale. The radiative timescale was partially considered in section 5.5 when the TEM model was forced with different perturbation depths. However the
role of the radiative timescale in all the models will now be investigated more fully.

Figure 5.3(b) showed the annual variation of the radiative timescale at 70hPa, calculated as described in section 5.2.1.1, for various depths of perturbation. In model [A] the perturbation depth used was 5000m which meant that the radiative timescale varied annually between 29 and 38 days. In model [B] the perturbation depth used was 3000m which meant that the radiative timescale varied annually between 22 and 28 days. In model [C] the perturbation depth used was 1000m which meant that the radiative timescale varied annually between 13 and 16 days. This range of radiative parameters is roughly in line with the 20 days suggested by Holton et al. (1995) and Haynes et al. (1991) for the lower stratosphere.

The annual variation in the radiative timescale, $\tau$, is predominantly due to annual variations in the temperature, with a minor contribution from annual variations in water vapour and ozone. This is seen by calculating an annual cycle in $\tau$, with various depths of perturbation, as described in section 5.2.1.1 but fixing the temperature profile as January. This leads to practically no annual variation in $\tau$, as shown in figure 5.15, and is in agreement with Shine (1987) who showed that $\tau$ is a function of temperature.

![Figure 5.15](image_url)  
*Figure 5.15*  As figure 5.3(b) but holding the temperature profile as perpetual January, and allowing the ozone and water vapour profiles annual variation.
It is seen from (5.4) that the radiative timescale, $\tau$, will not affect the temperature directly, but will rather amplify the effects of the dynamical forcing. The role of the radiative timescale is independent of what dynamical forcing is considered, so for simplicity the dynamical forcing is denoted $\Psi$. $\Psi$ could represent $A + B + \overline{\omega}^{10hP_o}S$ in (5.8), or $\overline{\omega}S$ in (5.19).

The equation used to model temperature is then:

$$\frac{\partial T}{\partial t} + \Psi = T_{eq} - \overline{T}.$$  

(5.22)

It is now assumed that $\Psi$ is sinusoidal and can be written:

$$\Psi = \Psi_0 \exp \left\{ \frac{2\pi it}{\Omega} \right\}.$$  

(5.23)

where $\Psi_0$ is the amplitude of the dynamical forcing, $\Omega$ is the period of the dynamical forcing and $t$ is time. Solving (5.22) shows that the temperature depends on the dynamical forcing, the equilibrium temperature, and the radiative timescale as follows:

$$T = T_{eq} - \frac{\Psi \tau \Omega^2}{\Omega^2 + 4\pi^2 \tau^2}.$$  

(5.24)

Though this analysis does not show the exact role of $\tau$ in the real atmosphere (because the dynamical forcing is not sinusoidal), it is able to show how the radiative timescale can interact with different dynamical forcings to affect the temperature. Equation (5.24) shows that the contribution to temperature of the dynamical forcing, $\Psi$, in the model, is amplified by $\frac{\tau \Omega^2}{\Omega^2 + 4\pi^2 \tau^2}$. This amplification is quantified for various values of $\tau$ and $\Omega$ in table 5.2. As $\Omega \to \infty$, the dynamical forcing will approach a constant value, and table 5.2 shows that the amplification of the dynamical forcing approaches $\tau$. If the dynamical forcing has sinusoidal variation with period of 1 year (365
days) then the amplification is close to $\tau$. However if the period of the dynamical forcing, $\Omega$ is similar to the radiative timescale then $\frac{\tau \Omega^2}{1 + 4 \pi^2 \tau^2}$ is small and will greatly reduce the impact of the dynamical forcing. The reason for this can be seen from looking at (5.4). Equation (5.4) shows that the radiative timescale will not just amplify the dynamical forcing at a particular timestep. It will also determine the ‘memory’ of the system. This is because the temperature also depends upon the state of the atmosphere at previous timesteps by an amount dependent on the radiative timescale. If the forcing oscillates around zero with a small period then the temperature would be expected to oscillate around $T_{eq}$ with a small period. This means that the cumulative effect of the dynamics cannot drive the temperature far from equilibrium and hence the amplification factor in table 5.2 is small.

\[
\frac{\tau \Omega^2}{1 + 4 \pi^2 \tau^2} \quad \text{(days)}
\]

<table>
<thead>
<tr>
<th>$\tau$ (days)</th>
<th>$\Omega = \infty$</th>
<th>$\Omega = 365$ (days)</th>
<th>$\Omega = \tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.0</td>
<td>13.0</td>
<td>12.4</td>
<td>0.3</td>
</tr>
<tr>
<td>16.0</td>
<td>16.0</td>
<td>14.9</td>
<td>0.4</td>
</tr>
<tr>
<td>20.0</td>
<td>20.0</td>
<td>17.9</td>
<td>0.5</td>
</tr>
<tr>
<td>22.0</td>
<td>22.0</td>
<td>19.2</td>
<td>0.5</td>
</tr>
<tr>
<td>30.0</td>
<td>30.0</td>
<td>23.7</td>
<td>0.7</td>
</tr>
<tr>
<td>38.0</td>
<td>38.0</td>
<td>26.6</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 5.2 $\frac{\tau \Omega^2}{1 + 4 \pi^2 \tau^2}$ calculated with various values of $\tau$ and $\Omega$.

It is now considered how the results of this study are affected by changing the value of the radiative timescale. This is shown in figure 5.16 for model [A]. Here the red line shows the modelled temperature assuming a constant $\tau$ of 38 days, the green line shows the modelled temperature assuming a constant $\tau$ of 29 days and the black line shows the modelled temperature where $\tau$ varies as in figure 5.3(b) with depth of perturbation 5000m. It can be seen that increasing $\tau$ will increase the impact of the dynamics as the model is able to cool the temperature further below radiative equilibrium. Increasing $\tau$ from 29 to 38 days leads to a decrease in the modelled temperature by up to 5K. It can be seen that the annual cycle in $\tau$ can explain an extra 2K of the annual temperature cycle than could be explained if $\tau$ were constant. This is because the annual cycle in $\tau$ is such that cooling below radiative equilibrium in DJF will be enhanced more than cooling below radiative equilibrium in JJA.
Chapter 5 The annual temperature cycle near the tropical tropopause

Figure 5.16 How $\tau$ affects the modelled temperature in model [A]. The green line shows the modelled temperature assuming a constant $\tau$ of 29 days, the red line shows the modelled temperature assuming a constant $\tau$ of 38 days, the black line shows the modelled temperature allowing $\tau$ to vary as in figure 5.3(b) with depth of perturbation 5000m.

Model [C], which considers the TEM form of the thermodynamic equation, is forced with a depth of perturbation 1000m, which means that the radiative timescale varies between 13 and 16 days. On figure 5.17 green and red lines show the temperature that this model would predict if constant $\tau$ values of 13 and 16 days respectively were used. The black line shows the temperature predicted when allowing $\tau$ to vary annually as calculated using depth of perturbation 1000m. In this model $\tau$ affects the temperature much less than it does in model [A]. This is mainly because $\tau$ has a much smaller annual variation with depth of perturbation 1000m than it does when calculated with depth of perturbation 5000m. The difference between using a value of $\tau = 13$ days as opposed to $\tau = 16$ days is only around 1K. The annual variation in temperature that is attributable to the annual variation in $\tau$ is also around 1K.

Figures 5.16 and 5.17 show that a larger perturbation depth leads to larger annual variations in $\tau$ which in turn lead to a larger amplitude of the modelled annual temperature cycle.

So far the models have been tuned by calculating the radiative relaxation timescale with a depth
Figure 5.17  How $\tau$ effects the modelled temperature. The red line shows the modelled temperature assuming a constant $\tau$ of 13 days, the black line shows the modelled temperature assuming a constant $\tau$ of 16 days, the black line shows the modelled temperature allowing $\tau$ to vary as in figure 5.3(b) with depth of perturbation 1000m.

of perturbation which is most able to model the temperature. This tuning is reasonable because it is unclear what the appropriate depth of perturbation is for the real atmosphere. For example Hartmann et al. (2001) calculated radiative relaxation times of 18 days and 35 days using a sinusoidal atmospheric perturbation with depths 7km and 14km respectively, Fels (1982) found the relaxation timescales in the dominant 15$\mu$m CO2 band at 20km could be between 9 and 140 days, the radiative relaxation timescale at 20km used by Dunkerton (1989) was approximately 60 days and Shine (1987) suggested that for perturbations of deep atmospheric extent the relaxation times are around 100 days in the lower stratosphere.

It is clear that none of the models considered in this chapter would provide a good prediction of temperature with the radiative timescale calculated using the deep atmospheric perturbations that many previous studies suggested. All of the perturbation depths used here are relatively shallow and in particular the very small perturbation depth of 1km, which was used in model [C], gives a radiative relaxation time much smaller than used by these previous studies. However the considerable uncertainty in previous studies regarding which radiative timescale should be used, suggests that the most appropriate relaxation times are those that are able to provide a good prediction of temperature.
Finally, it must be noted that this study has found those radiative parameters which are most consistent with the dynamical forcings at 70hPa. If similar radiative parameters were required by other studies then this would suggest a relatively shallow perturbation depth was most appropriate to calculate radiative parameters for this region of the atmosphere.

5.7 How should the temperature be modelled?

Sections 5.3, 5.4 and 5.5 presented three models which can obtain a reasonable prediction of the annual temperature cycle averaged over 10°N to 10°S, and 1981 to 1993 at 70hPa. These models differ in their treatment of equatorial waves, how their radiative parameters are calculated and in their forcing regions of vertical velocity. In order to gain insights about the radiative and dynamical properties of the atmosphere a decision must be reached as to which model best represents the physical processes in the atmosphere.

Model [A] is perhaps the least physically representative, as this assumes that only a proportion of horizontal advection terms in ERA15 is due to equatorial waves, while being unable to give any physical explanation for the remainder. This model also requires the invalid assumption that there is no tropical upwelling forced from above 10hPa, and the radiative equilibrium temperature it requires is not consistent with the radiative timescale.

Models [B] and [C] are similar in that both assume extratropical wave driving to the top of the atmosphere, and both contain a radiative equilibrium temperature consistent with the radiative timescale when using the Morcrette scheme. The radiative parameters in model [C] are calculated with a depth of perturbation of 1000m and radiative parameters in model [B] are calculated with a depth of perturbation of 3000m. The larger perturbation depth of model [B] is the most reasonable of these, but even this is small in comparison with other studies (see section 5.6).
The greatest advantage to using model [C] is that it requires no assumptions about equatorial waves in the dataset and instead uses the eddy heat flux divergence term which is obtainable directly from the ERA15 data. The increase in accuracy due to including observable wave characteristics in the model is expected to outweigh any potential error due to using the smaller (and perhaps less realistic) perturbation depth. Of the three models described earlier in this chapter it is expected that model [C] will produce the most realistic results, and this model is now used to investigate temperature at different pressure levels and for individual years.

5.8 Using the TEM thermodynamic equation (Model [C])

5.8.1 Modelling individual years at 70hPa

It was shown in section 5.5.2 that the TEM thermodynamic equation is able to reasonably predict the average temperature cycle at 70hPa, when radiative timescale and radiative equilibrium temperature are calculated from the Morcrette radiation scheme with a depth of perturbation 1000m. Results are now presented which show the ability of this model to predict the temperature for individual years at 70hPa.

Figure 5.18 shows the ability of the TEM thermodynamic equation to predict the temperature for the years 1981-1993 averaged over 10°N - 10°S at 70hPa. The black solid line is the ERA15 temperature and the red dotted line is the modelled temperature. It can be seen that the model produces a good representation of temperature every year and is able to reproduce many of the interannual variations including the very cold winter of 1984-1985 and the relatively warm summer of 1983. Throughout, day to day variations are often very well represented and many examples can be seen of temperature changes over a few days being almost perfectly modelled (see 1991 around day 150 for one such example). The ability of the model is perhaps surprising when we consider that the upwelling forced from above 10hPa and the radiative equilibrium temperature have been assumed to be the same every year and are shown by the black line in figure 5.14(b) and the purple line in figure 5.12 respectively. This means that all of the day to
day variation which has been accurately modelled is due to physical process operating between 70hPa and 10hPa.

Figure 5.19 shows the contribution to the temperature variation for each year from the Eulerian-mean upwelling forced between 70hPa and 10hPa. It can be seen that there are many years where this does not have an annual cycle of the same phase as the observed temperature (e.g. 1987, 1989). It is these years in particular where the parameterised upwelling forced from above 10hPa is a vital input to the annual cycle model. On figure 5.19 the horizontal red lines show those times when the QBO is easterly. The QBO used here is defined by the 50hPa zonal wind at Singapore, and this will lead the 70hPa zonal winds by several months. It can be seen in figure 5.19 that during the middle and end of the easterly QBO phase the mean upwelling forced between 70hPa and 10hPa leads to lower temperatures than at the corresponding time of year when the QBO is westerly. The magnitude of this effect is 1-2K. In a ‘mean’ year, it has been shown that the temperature due to upwelling forced between 70hPa and 10hPa will be cooler in the NH winter and warmer in the NH summer. For individual years, this is not necessarily the case, because the phase of the QBO can have a very strong effect on the upwelling (see 1987). Other studies have also found tropical upward mass flux to be enhanced during the easterly QBO phase and reduced during the westerly QBO phase (e.g. Seol and Yamazaki 1998, Yang and Tung 1996). This has been attributed to enhanced wave breaking in the narrower westerly waveguide squeezed by the easterly wind (Baldwin et al., 2001). In addition the same amount of wave driving would have greater effect in the easterly QBO phase as there is less angular momentum contrast between the tropics and mid-latitudes. Dunkerton (1989) was able to produce a thermally driven meridional circulation without any sources or sinks of angular momentum only if there were strong easterly winds at the equator and strong westerly winds in mid-latitudes. Though it is known that the QBO can modulate the annual temperature cycle, modelling results presented show that if all of the QBO impacts are confined to the 70hPa-10hPa layer then the annual cycle can still be extremely well predicted. The forcing above 10hPa and the radiative timescale in the model have been calculated in such a way that they can have no QBO dependence.

Much of the interannual variation that has been reasonably modelled is due directly to the mean
Chapter 5

The annual temperature cycle near the tropical tropopause

Figure 5.18 The black solid line is the ERA15 70hPa temperature averaged over 10N-10S for each day and the red dashed line is the modelled 70hPa temperature averaged over 10N-10S for each day.
Eulerian Upwelling forced between 70hPa and 10hPa. For example in the winter of 1984-1985 there are easterly winds which force additional upwelling between 70hPa and 10hPa and lead to the cold temperatures observed in this year. The years 1983 and 1986 are when the QBO is westerly and therefore there is much less upwelling forced between 70hPa and 10hPa for these years and this explains the relatively warm temperatures in these years.

Figures 5.20 show the contributions to the 70hPa temperature from the eddy heat flux divergence term \( \frac{\partial}{\partial y} \left( \frac{\partial T}{\partial y} \right) \) for 1987-1989. This term is included in the TEM residual upwelling but not included in the Eulerian-mean upwelling, and represents the dynamical difference between model [B] and model [C]. The effect of this term on the temperature for the three years shown is similar to all the years modelled. There is not a clear pattern as to when the cooling from this term is the strongest. The role of the eddy heat flux divergence term appears to be in cooling the
temperature 1-2K below its radiative equilibrium value and also in explaining some of the day to day variation. This term does not have any impact on the amplitude of the annual temperature cycle whatsoever.

![Figure 5.20](image)

**Figure 5.20** Shows the interannual temperature variation due to the heat flux divergence component of $\omega^*$, i.e. the difference in temperature caused by substituting $\omega$ with $\omega^*$ in (5.14)

When considering individual years at 70hPa, it is seen that different components of the upwelling have different roles. These are: the tropical upwelling forced from above 10hPa is dominant in producing the annual temperature cycle, the tropical upwelling forced between 70hPa and 10hPa has a small role in the annual temperature cycle but accounts for most of the interannual variability, and the eddy heat flux divergence term can affect the day to day variability only.

We now consider whether the TEM thermodynamic equation is able to explain the annual temperature cycle at other levels.

### 5.9 Modelling the temperature at other levels

#### 5.9.1 Radiative parameters

Figure 5.21 shows the radiative equilibrium temperature and radiative timescale for each month, calculated as described in section 5.2.1.1 for various pressure levels. The pressure levels shown are: 110hPa - black crosses, 93hPa - red stars, 68hPa - green diamonds, 50hPa - blue triangles and 32hPa purple squares. It is seen that as pressure decreases the radiative equilibrium temperature
increases and the radiative timescale decreases; this is in agreement with results of Fels (1982) and Shine (1987). The radiative parameters calculated here will be used to force the TEM thermodynamic model at 100hPa and 50hPa in order to see whether the TEM thermodynamic model can predict the temperature at these levels.

Figure 5.21  a) shows the radiative equilibrium temperature and b) shows the radiative timescale, for each month. The values are shown at the following pressure levels: 110hPa - black crosses, 93hPa - red stars, 68hPa - green diamonds, 50hPa - blue triangles, and 32hPa purple squares. Thin solid lines are added for clarity.
5.9.2 The ‘average’ annual cycle at 100hPa

The temperature profiles input into the Morcrette scheme to calculate the radiative parameters did not include 100hPa. The levels nearest to 100hPa were 93hPa and 110hPa and the radiative parameters at these levels are seen in figure 5.21. The radiative timescale is very similar at 93hPa and 110hPa and so either of these values of radiative timescale could be input to the model at 100hPa and would have little effect on the results. The radiative equilibrium temperature at 93hPa is around 2K warmer than the radiative equilibrium temperature at 110hPa and so the model results at 100hPa would differ by 2K depending on which radiative equilibrium temperature was used. It was decided that the colder of the two radiative equilibrium temperatures should be used since 100hPa is close to the coldest point of the tropopause region throughout the year.

The black line of figure 5.22 shows the ERA15 temperature at 100hPa averaged over 10°N - 10°S and 1981-1993. The red line is the temperature, averaged over the same spatial region and time period obtained using the TEM thermodynamic model, and the purple line is the radiative equilibrium temperature. Figure 5.22 shows that the model is able to produce a good prediction of the annual temperature cycle at 100hPa, however the predicted temperature has a small warm bias of approximately 2-3K throughout the year. The model has not been tuned to 100hPa, yet it still appears to have a great deal of skill at this level. This increases our confidence in the model’s ability.

The 100hPa temperature has been predicted in exactly the same way as the 70hPa temperature, yet there are likely to be important differences between the two levels. The main difference is that some convection is likely to reach 100hPa (Gettelman et al., 2002b). The mixing effects of this convection may reduce the relaxation to equilibrium timescale at 100hPa. For example Dunkerton (1989) figure 3 suggests a tropospheric relaxation timescale of around 7 days which will increase slowly with height from the top of convection until it reaches its stratospheric values at around 18km. Since there is a small amount of convection reaching 100hPa it may be necessary to reinterpret the radiative parameters. At 100hPa the relaxation to equilibrium timescale (τ) and the equilibrium temperature (T_{eq}) may be better represented by a radiative-convective timescale and a radiative-convective equilibrium temperature respectively. The radiative-convective
timescale should be shorter than the radiative timescale (see Dunkerton, 1989, figure 3). The radiative-convective equilibrium temperature would also be smaller than is shown in figure 5.22 because of (5.6). It is possible that the effects of convection could reduce \( \tau \) and \( T_{eq} \) sufficiently to cause a more accurate modelled prediction of temperature at 100hPa than is seen in figure 5.22.

![Figure 5.22](image)

**Figure 5.22** The black solid line is temperature at 100hPa averaged over 1981-1993, the red dotted line is modelled temperature at 100hPa averaged over 1981-1993. Both curves are also averaged over \( 10^\circ N - 10^\circ S \). Purple dash-dotted line is the radiative equilibrium temperature used which is calculated at 110hPa using perturbation depth of 1000m.

The difference between the radiative equilibrium temperature and the modelled temperature in figure 5.22 is due to the residual upwelling term, \( \vec{\omega} \vec{S} \), and most of this is due to the Eulerian upwelling term, \( \vec{\omega}_S \). Since the Eulerian upwelling, \( \vec{\omega} \) forced above 10hPa was assumed to be an annually varying proportion of upwelling forced between 100hPa and 10hPa, the contribution to temperature forced in each of these regions is not shown separately. Instead, section 5.4.1 suggests that the contribution to temperature from the Eulerian upwelling forced from above 10hPa is about one third of that forced between 100hPa and 10hPa. This contrasts with the results at 70hPa where the Eulerian upwelling forced from above 10hPa was the dominant dynamical cause of temperature change from radiative equilibrium.
For completeness the contribution of the eddy heat flux divergence term \( \frac{\partial}{\partial y} \left( \frac{v^T S^T}{S} \right) \) to temperature at 100hPa is shown in figure 5.23. It is seen that this term (which represents the difference between the TEM and Eulerian-mean forms of the thermodynamic equation) will cool the modelled temperature by between 1 and 2K and will reduce the amplitude of the modelled annual temperature cycle by around 0.5K. It is seen that this term is relatively most important in NH summer when it explains nearly all of the cooling below radiative equilibrium that the model is able to produce.

![Graph showing contribution to temperature](image)

**Figure 5.23** Shows the contribution of eddy heat flux divergence term to the temperature of 10°N – 10°S at 100hPa. Averaged over 1981-1993. This is the same as the difference between using \( \overline{\omega} \) and \( \overline{\omega}^* \) in the thermodynamic equation.

Finally the average annual temperature cycle at 100hPa is considered by assessing the relative importance of mean upwelling and static stability in producing the cooling below radiative equilibrium due to \( \overline{\omega}^* S \). Section 5.3.2 found that at 70hPa the static stability of the atmosphere was fairly constant and while this parameter could amplify the effects of the vertical velocity on the temperature, there were no annual variations in \( S \) that could lead to significant variations in the temperature. At 100hPa, however, the situation is slightly different. Figure 5.24 shows the value of \( S \) used for each day of the year and it is seen that \(|S|\) is around twice as large in the NH summer than it is in the NH winter. The larger annual variation at 100hPa is due to the fact that 100hPa is sometimes in the troposphere (where there is low static stability) and sometimes in the stratosphere (where there is high static stability) (see figure 5.1) and \(|S|\) at 100hPa will be largest when this level is in the stratosphere. This annual cycle in \(|S|\) will have the effect of damping the annual cycle in temperature at 100hPa. This is because upwelling in the NH summer will
have twice as large an effect on temperature as the same amount of upwelling in the NH winter. This partly explains why the observed annual temperature cycle at 100hPa is smaller than the observed annual temperature cycle at 70hPa, despite a greater residual pressure vertical velocity at the 100hPa level.

![Figure 5.24](image)

**Figure 5.24** Annual variation of $S$ at 100hPa.

### 5.9.3 Interannual variations in temperature at 100hPa

Figure 5.25 shows the ERA15 temperature (black line) and modelled temperature (red line), averaged over $10^\circ$N - $10^\circ$S, for the years 1987-1992. The model results for these years are similar to those of other years in the dataset. For all years the temperature is predicted to be warmer than observed by up to 4K, which could be attributed to the effect of convection on radiative parameters (discussed earlier). For most years the shape of the modelled annual temperature cycle is reasonable, although its amplitude is often too large. Perhaps this is because the radiative timescale that was used in the model is larger than the radiative-convective timescale that should have been used. The largest discrepancies between the observed and modelled shape of the annual cycle occur in the NH late winter (between day 0 and day 90). At this time of year the 100hPa level is most likely to be in the troposphere and hence subject to a greater error in the radiative parameters.

Though there are some discrepancies between the modelled temperature and the ERA15 temperature when considering individual years at 100hPa, these results from the TEM thermodynamic
Figure 5.25 The black solid line is the ERA15 100hPa temperature averaged over 10N-10S for each day and the red dashed line is the modelled 100hPa temperature averaged over 10N-10S for each day.

Equation are more accurate than if one of the Eulerian-mean models were used. For example, figure 5.26 shows the ability of model [B] to predict the temperature at 100hPa for 1987-1989. The black line is the ERA15 temperature and the red line is the modelled temperature. Again the amplitude of modelled temperature cycle is too large, suggesting that the radiative timescale used in this model is also too large. It therefore appears that model [B] could also benefit from replacing the radiative parameters with radiative-convective parameters at 100hPa. However the results of model [B] do appear less accurate than those derived using the TEM thermodynamic equation (model [C]) throughout.
Figure 5.26  The black solid line is the ERA15 100hPa temperature averaged over 10N-10S for each day and the red dashed line is the 100hPa temperature calculated using Model [B] described in section 5.4.1

5.9.4 The ‘average’ annual cycle at 50hPa

At 100hPa any mismatches between the observed and modelled temperature could be attributed to the fact that convection was not included in the calculation of the relaxation to equilibrium timescale and the equilibrium temperature. The 50hPa level is undoubtedly in the stratosphere at all locations and at all times of year and hence any mismatches between observed and modelled temperatures at this level cannot be attributed to convection. It is therefore desirable to understand whether the model is able to predict the temperature at 50hPa.

On figure 5.27, the black line is the 50hPa temperature from ERA15 (averaged over 10°N-10°S and 1981-1993), the purple line is the radiative equilibrium temperature (calculated with perturbation depth of 1000m) and the red dotted line is the modelled temperature at 50hPa. It is seen that the modelled temperatures are around 2K warmer than the ERA15 temperature at 50hPa in the NH summer; this means that the annual cycle in the modelled temperatures has an amplitude nearly twice that of the ERA15 temperatures. The green and yellow lines on figure 5.27 are the temperatures at Truk (7.45N,151.85E) and Pago Pago (14.33S,170.72W) respectively, derived from radiosonde data and averaged over 1966-1997. It can be seen that the radiosonde temperatures are in greater agreement with those temperatures produced from our model than they are with the ERA15 temperatures. It is likely that the Truk and Pago Pago temperatures are representative of the tropical average temperature at 50hPa as these stations are
not in close proximity to each other yet they have similar annual variation. In addition horizontal gradients of temperature are expected to be small at 50hPa. It therefore appears that the model is producing a better representation of the real average temperature at 50hPa than ERA15!

![Figure 5.27](image)

**Figure 5.27**  The solid black line is temperature at 50hPa averaged over 1981-1993, dotted red line is modelled temperature at 100hPa averaged over 1981-1993. Both curves are also averaged over 10° N – 10° S. The purple dash-dotted line is the radiative equilibrium temperature used which is calculated using perturbation depth of 1000m. The green and yellow lines are the temperatures at Truk (7.45N,151.85E) and Pago Pago (14.33S,170.72W) respectively, derived from radiosonde data and averaged over 1966-1997.

To understand these results at 50hPa, the temperature change due to residual upwelling is decomposed into several components. These components are: upwelling forced between 50hPa and 10hPa, upwelling forced above 10hPa, and the eddy heat flux divergence term. These affect the temperature as shown in figures 5.28(a), 5.28(b) and 5.28(c) respectively. It is seen that of these the only component that has any significant contribution to the annual temperature cycle is the Eulerian-mean upwelling forced from above 10hPa. This component has been parameterised by our model.

It is important to understand why the ERA15 temperature at 50hPa does not agree with radiosonde temperature. There are two possible reasons for this. Firstly, at 50hPa, much of the observational data used in ERA15 is provided by satellites. Unlike radiosonde data, satellite data does not have good vertical resolution and so the 50hPa satellite temperature data may be
Figure 5.28  The contribution to 50hPa temperature of various components of the residual upwelling. All these are calculated using radiative parameters obtained with a depth of perturbation 1000m

contaminated by temperature data from other levels. Secondly, the forecast model in ERA15 could be inadequately predicting the 50hPa temperature. This is because the ERA15 model uses the thermodynamic equation and hence a good representation of poleward mass flux above 10hPa is needed to predict the temperature. The vertical velocity at 10hPa in ERA15 may be insufficient since the ERA15 model top is at 10hPa. To see whether the error in the ERA15 50hPa temperature is likely to be due to poor observations or inadequate forcing of the thermodynamic equation the NCEP-reanalysis data is considered. The NCEP reanalysis data was derived in a similar way to the ERA15 reanalysis data (described in chapter 2). One important difference between the NCEP reanalysis data and the ERA15 data is that the NCEP data contains an extra model level at 3hPa. If the NCEP-reanalysis data is in good agreement
with the radiosonde data at 50hPa, then it will imply that extra model levels in the stratosphere are important. However, if the NCEP data is in good agreement with ERA15 than it could imply that the assimilated satellite observations can contaminate the lower stratospheric temperature.

The daily NCEP temperature data at 50hPa averaged over 1968-1996 and 10°N-10°S is shown by the black line in figure 5.29. The data was provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, from their Web site at http://www.cdc.noaa.gov/. Daily NCEP temperature data at 50hPa averaged over 1981-1993 could not be obtained. However monthly averaged NCEP temperature data at 50hPa was almost identical for the periods 1981-1993 and 1968-1996. This implies that the 50hPa NCEP temperature data averaged over 1968-1996 can be directly compared with 50hPa ERA15 temperature data averaged over 1981-1993. On figure 5.29 the green and yellow lines show the 50hPa temperature at the radiosonde stations of Truk and Pago Pago respectively, averaged over 1966-1997. It can be seen that the NCEP reanalysis data is in very good agreement with the radiosonde data. This highlights the requirement of additional model levels above 10hPa in a reanalysis dataset to adequately represent dynamics below 10hPa.

![Figure 5.29](image.png)

**Figure 5.29** The solid black line is temperature at 50hPa from NCEP reanalysis data averaged over 1968-1996, and over 10°N – 10°S. The green and yellow lines are the temperatures at Truk (7.45N,151.85E) and Pago Pago (14.33S,170.72W) respectively, derived from radiosonde data and averaged over 1966-1997

At 50hPa the ‘mean’ annual temperature cycle can be almost fully attributed to the Eulerian upwelling forced from above 10hPa and the radiative equilibrium temperature. Both these terms
have a constant annual cycle in this study and so any interannual variations in 50hPa temperature could not be captured by the model. For this reason the 50hPa temperature for individual years is not considered.

5.10 Summary

This chapter has shown that it is possible to reasonably model the temperature at 100hPa, 70hPa, and 50hPa, averaged over $10^\circ \text{N} - 10^\circ \text{S}$, using the thermodynamic equation. The thermodynamic equation is input with all advective and flux terms, the radiative timescale, and the radiative equilibrium temperature. However, it has been shown that there is not a unique set of radiative and dynamical inputs that can predict the temperature. The temperature can be reasonably predicted using each of at least 3 models which have slightly different inputs. Comparing and contrasting these models provides insights into the dynamical and radiative properties of the lower stratosphere.

The Eulerian-mean upwelling used in this study was initially calculated from meridional velocity in ERA15 between 70hPa and 10hPa. This was unable to fully account for the coldness of the temperature at 70hPa, or its annual cycle, regardless of what radiative parameters are used in the thermodynamic model. However adding extra dynamics into the thermodynamic model can provide sufficient extra cooling to obtain a very realistic modelled temperature.

In the first model considered (model [A]), the extra cooling required to predict the mean temperature at 70hPa, was provided by the horizontal temperature advection which could not be attributed to equatorial linear waves (chapter 3). The addition of this extra cooling led to a reasonable temperature prediction; however the physical cause of this horizontal temperature advection was unknown. If all of the horizontal temperature advection in the ERA15 data was attributed to equatorial waves (regardless of the results of chapter 3) then the mean 70hPa temperature could still be modelled. The extra cooling required was then provided by amplifying the vertical velocity at 70hPa by adding in a parameterisation of poleward mass flux above 10hPa.
This formed the basis of model [B] and it was found that the vertical velocity forced between 70hPa and 10hPa and that forced from above 10hPa had roughly equal contributions to the 70hPa temperature difference from radiative equilibrium.

Comparing model [A] (which assumed that the wave climatology of chapter 3 was complete) and model [B] (which attributed the remaining horizontal temperature advection to undetected equatorial waves) provided insights into the relationship between horizontal temperature advection and waves in the atmosphere. Model [B] produced a more accurate representation of temperature at 70hPa than model [A] and used more physically realistic inputs. This implies that the horizontal temperature advection, which was not accounted for in chapter 3, is due to some form of wave which is not represented by the idealised linear wave theory in a zero background flow that was presented in chapter 2. Choosing model [B] over model [A] shows that the vertical velocity forced from above 10hPa is necessary to determine the temperature at 70hPa. It is therefore important to include a good representation of this forcing in models of the tropical lower stratosphere. The forcing from above 10hPa will also have implications for other levels of the atmosphere.

The temperature was also modelled by replacing the Eulerian-mean upwelling of model [B] with the residual upwelling, $\omega^*$, to give the TEM thermodynamic equation model (model [C]). This model also required a parameterisation of upwelling forced from above 10hPa. The results obtained using the TEM form of the thermodynamic equation were more accurate than those obtained using the Eulerian-mean form of the thermodynamic equation.

The difference in the modelled temperature between the TEM and Eulerian-mean form of the thermodynamic equation was due to the eddy heat flux divergence term ($\frac{\partial}{\partial y}(\frac{\zeta}{\gamma} \bar{T}^*)$). This difference was greater than what could be attributed to equatorial waves, if equatorial waves were assumed to be consistent with those found in chapter 3. This meant that those disturbances which have not been accounted for in chapter 3 have a relatively greater effect on the Eulerian-mean circulation than those disturbances that have been attributed to linear waves. This highlights the need for expanding the wave detection study of chapter 2 to detect waves more appropriate to
the real atmosphere. The first step in this process would be to devise a method to detect dry, free, linear, equatorial waves in a vertical shear zone. However it would also be useful to devise a method to detect forced waves and non-linear waves, in order to understand their impact on $\frac{\partial}{\partial y} \left( \frac{\omega' \theta'}{S} \right)$.

The lack of assumptions about the existence or linearity of equatorial waves in the TEM thermodynamic equation (model [C]) meant that, of the three models, this was the most likely to be accurate and was therefore considered further.

The radiative timescale and radiative equilibrium temperature used in the TEM thermodynamic model were calculated as in section 5.2.1.1 using a perturbation depth of 1000m. This perturbation depth and the corresponding radiative parameters were smaller than many other studies have suggested (section 5.6) however these radiative parameters led to an extremely good prediction of temperature, not only in the mean annual temperature cycle but also when individual years were considered. It is therefore likely that, if there is not some persistent bias in the ERA15 data, the radiative parameters used in this study are appropriate.

The TEM form of the thermodynamic equation has also been used to model the temperature for individual years at 70hPa, and a very good prediction of temperature obtained. It was found that the annual temperature cycle was mainly due to annual variations in upwelling forced from above 10hPa and the radiative equilibrium temperature. Interannual variability appeared to be mainly related to the QBO and this was accounted for in the model by variations in upwelling forced between 70hPa and 10hPa. The difference between the residual upwelling, $\overline{\omega'}$, and the Eulerian-mean upwelling, $\overline{\omega}$, did not appear to affect the annual temperature cycle or interannual variability but accounted for 1-2K of cooling below radiative equilibrium.

The TEM thermodynamic equation was used to model the temperature at 100hPa and 50hPa. At these levels all of the inputs to the model, including the radiative parameters, were found in exactly the same way as they were at 70hPa. Considering these levels was a valuable test of
the model’s ability, since the radiative parameters had been tuned to provide good results at 70hPa.

At 100hPa, the mean annual temperature cycle was reasonably predicted. However the absolute value of temperature was usually predicted to be around 2K warmer than observed. This discrepancy was attributed to tropospheric processes, which would be likely to reduce the radiative timescale and radiative equilibrium temperature below what the radiation scheme calculated. It was shown that the largest errors in the shape of the annual cycle often occurred at those times of year when 100hPa was in the troposphere, although the temperature predicted was also too warm in the NH summer when 100hPa was in the stratosphere.

At 50hPa, discrepancies between the observed and modelled temperature could not be attributed to tropospheric processes, but here, for most of the year, the model was able to predict the average annual cycle extremely well. Indeed the modelled temperature was in greater agreement with radiosonde data at two stations than the ERA15 temperature was. This perhaps suggests that only having 1 model level to represent 10hPa-0hPa in ERA15 is detrimental to the ERA15 temperature below 10hPa. The TEM thermodynamic model uses a parameterisation of dynamics above 10hPa and is therefore able to reasonably predict the 50hPa temperature.
CHAPTER 6

Conclusions

This thesis has considered the dynamical variability in the tropical tropopause and tropical lower stratosphere region. Understanding the variability of this region on a variety of timescales is necessary before questions such as “what processes control the amount of water vapour entering the stratosphere?” or “which waves are responsible for driving the QBO?” can be answered. A physical understanding of what causes variability in the tropical tropopause and tropical lower stratosphere region will lead to more accurate predictions of how this region will respond to any future changes in forcing that may occur.

This thesis set out to increase understanding of variability in the tropical tropopause and tropical lower stratosphere region by answering three main questions. These are:

- What is the climatology of idealised linear equatorial waves in this region?
- How do these waves interact with the QBO?
- Can the annual cycle in lower stratospheric temperatures be quantitatively modelled?

These questions have been considered using the ERA15 dataset. This dataset was chosen because it contains tropical data on an evenly distributed horizontal grid and hence any results found will be more ‘typical’ of the deep tropics than if observations at a single location were used.

Throughout the thesis, the extent to which ERA15 data is representative of the real atmosphere has been considered. It is found that ERA15 generally compares well with radiosonde data. The results obtained from ERA15 compare well with previous studies. It is therefore expected that the results of this thesis are directly applicable to the real atmosphere. As an added bonus, the
information about variability in ERA15 will be useful. This is because ERA15 has been used in many previous studies of the tropical tropopause and lower stratosphere (e.g. Jackson et al. 2001, Norton, 2001).

The conclusions of this thesis will now be discussed in relation to the research questions proposed.

6.1 The equatorial wave climatology

The method used to produce a climatology of idealised, linear, equatorial waves in ERA15 is derived in chapter 2. The main advantage of the method over previous methods is that it can assign disturbances averaged over a season to an optimal combination of idealised linear waves. Linear waves are detected in the fields $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$, as these fields are the same sign at nearly all phases of a linear wave. The method involves a least squares projection of the latitudinal structure of $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ in the ERA15 data onto the latitudinal structure of $u \frac{\partial T}{\partial x}$ and $v \frac{\partial T}{\partial y}$ in idealised linear equatorial waves. The waves detected must satisfy a dispersion relation from linear wave theory; however no assumptions are made concerning the equivalent depths at which waves should be found. The climatology produced contains estimates of wave activity in each season, and also shows the wavenumbers and frequencies at which each wave type occurs. Individual wave events are not detected using this method and therefore no information is available about typical wave amplitudes or the geographical locations of waves.

Chapter 3 presents the climatology of idealised linear equatorial waves in ERA15 at 100hPa, 70hPa, and 50hPa. Because ERA15 uses diabatic normal mode initialisation on the 5 deepest vertical modes, the only waves considered have zonal wavenumber less than 20 and frequency less than 1 cycle per day. Each wave type has been looked for at all wavenumbers and frequencies where its presence is theoretically possible, however it is found that waves predominantly occur with dispersion relations of equivalent depth between 12m and 100m. This is in agreement with previous studies and so confidence in the wave detection method and the representation of waves by ERA15 is increased.
In the ERA15 dataset the dominant linear equatorial mode is the Kelvin wave. This mode can explain $0.5K^2 - 1K^2$ of temperature variance on the equator at 50hPa. Kelvin wave activity appears correlated with the annual cycle in tropical convection at 100hPa, the annual cycle in the tropical tropopause at 70hPa, and the QBO at 50hPa. High frequency Kelvin waves are relatively more important at lower pressures which is consistent with fast Kelvin waves suffering less damping than slow Kelvin waves as they propagate upwards.

The westward propagating Rossby-gravity waves were strongest at wavenumber 4 and at a period of approximately 5 days at all levels (in agreement with Yanai and Marayanna, 1966). At 100hPa peak Rossby-gravity wave activity occurred during JJA or SON, while at 50hPa peak Rossby-gravity wave activity was strongest as the QBO was changing from westerly to easterly. At 70hPa the temporal variation of Rossby-gravity waves appeared to be influenced by both the QBO and the annual cycle.

The characteristics of Kelvin waves and Rossby-gravity waves found are in reasonable agreement with what has been found in previous studies (e.g. Wheeler et al., 2000), hence it is likely that these waves in ERA15 data are similar to those in the real atmosphere. The other waves found in this study (EIG0, EIG1, WIG1, and ER1) have not been frequently observed in the tropical lower stratosphere, despite the fact that they are likely to be important for driving the QBO (Baldwin et al., 2001). The lack of previous observations for comparisons makes it impossible to be certain that EIG0, EIG1, WIG1, and ER1 waves in ERA15 are true representations of the atmosphere. However this thesis provides the first observational evidence for some of the characteristics of these waves, and hence can be used as a guide for future wave detection studies. For example the ERA15 dataset contains very strong evidence for the existence of the EIG0 wave; it finds peak EIG0 wave activity at zonal wavenumber 1-3, at a frequency of around 0.3 cycles per day and in the season JJA. The characteristics of EIG0 waves found here could provide a valuable reference point for future studies which aim to detect EIG0 waves.

Signals of EIG1 and WIG1 waves in ERA15 are robust but were found over a greater range of wavenumbers and frequencies than other wave types. This means they have smaller power at a
Chapter 6

Conclusions

particular wavenumber and frequency and are more difficult to detect. Strong signals, which have a frequency of around 1 cycle per day, have been attributed to the EIG1 wave but are probably due to nonmigrating diurnal tides. Signals consistent with IG2 waves were detected in ERA15 and, for the most part, appeared consistent with theory. However IG2 waves were not considered in any detail because the resolution of the data was not sufficient to fully represent the latitudinal structure of some of these waves.

At 100hPa and 70hPa, ER1 wave activity was found to follow a seasonal cycle with peak activity in DJF. At 50hPa ER1 wave activity was more strongly correlated with the QBO than any other wave type and maximum activity occurred when the QBO was westerly. There was no clear evidence of ER2 wave activity in the ERA15 dataset at 100hPa, 70hPa, or 50hPa.

Linear waves contribute to the variance in temperature, zonal wind, and meridional wind of the tropical lower stratosphere. However only a small proportion of the variance of these fields in ERA15 is explained by the idealised linear equatorial waves found. This is because idealised linear waves were detected in the quadrature spectra between $u$ and $T$, while ERA15 contains a large cospectra between $u$ and $T$. The cospectra between $u$ and $T$ would not contribute to the wave climatologies but would contribute substantially to the variance of $u$ and $T$. Large cospectra between $u$ and $T$ were also found in radiosonde data at Singapore by Sato et al. (1994) and hence is characteristic of the real atmosphere as well as ERA15 data. The cospectra between $u$ and $T$ could be due to any disturbances that are not fully represented by the wave theory of chapter 2. These could include: waves in vertical shear, forced waves, non-linear waves, or disturbances not directly related to waves.

6.2 Equatorial waves and the QBO

Table 6.1 shows the phase of the QBO when there was peak activity of each wave type (in ERA15) at 50hPa. These results are qualitatively consistent with linear wave theory and add confidence to the wave climatologies and the representation of waves in ERA15. The EIG0 wave and the EIG1 wave are not correlated with the QBO at 50hPa because these waves have a larger
group velocity than other waves in the extreme phases of the zonal wind. This means that of all the waves, the EIG0 and EIG1 waves are the least susceptible to damping and will therefore be the least influenced by the QBO.

<table>
<thead>
<tr>
<th>wave type</th>
<th>Phase of QBO when there is peak wave activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kelvin</td>
<td>Easterly to westerly transition</td>
</tr>
<tr>
<td>EIG0</td>
<td>Not correlated</td>
</tr>
<tr>
<td>EIG1</td>
<td>Not correlated</td>
</tr>
<tr>
<td>Rossby-gravity</td>
<td>westerly to easterly transition</td>
</tr>
<tr>
<td>ER1</td>
<td>westerly</td>
</tr>
<tr>
<td>WIG1</td>
<td>at, or just before, start of westerly</td>
</tr>
</tbody>
</table>

Table 6.1  Shows the phase of the QBO where each wave type shows maximum activity

When waves are damped the momentum associated with the wave is transferred to the mean zonal flow (see chapter 1), providing the driving force for the QBO. The transfer of momentum associated with waves in ERA15 was therefore calculated. It was found that momentum forcing by waves of frequency $< 1$ cycle per day (in ERA15) was insufficient to drive either phase of the QBO, however the acceleration provided by the waves was mainly westerly due to the dominance of Kelvin waves. The strongest easterly acceleration was provided by WIG1 and ER1 waves, while acceleration from the Rossby-gravity wave was found to be negligible.

Pawson and Fiorino (1998b) showed that the QBO in ERA15 does not have as large an amplitude as the QBO in the real atmosphere. They also showed that the forecast model used in ERA15 cannot produce a QBO at all unless it is assimilated with observations. This is hardly surprising in light of the small accelerations produced by the waves with frequency $< 1$ cycle per day that this thesis has found. However if waves with frequency $< 1$ cycle per day were perfectly represented by ERA15 and perfectly detected in this thesis, a quantitatively accurate QBO could still not be produced because extra momentum flux would be required from higher frequency waves (e.g. Giorgetta et al., 2002). In addition it is likely that this thesis does not detect all disturbances that are due to equatorial waves at low frequencies, because the method used assumed an isothermal
atmosphere and a zero background wind, neither of which are the case in the real atmosphere.

6.3 Modelling the annual temperature cycle in the tropical lower stratosphere

One of the largest variabilities in the tropical lower stratosphere is the annual temperature cycle. This has amplitude of approximately 4.5K at 80hPa and leads to an annual variation in water vapour entering the stratosphere (Mote et al., 1996). The annual temperature cycle has been attributed to the wave-driven stratospheric circulation (Yulaeva et al., 1994).

In chapter 5 the thermodynamic equation was used to quantitatively investigate the relative roles of dynamics and radiation in determining the lower stratospheric temperature, and its annual variation. A quantitative understanding of what determines the lower stratospheric temperature would provide insights into how the temperature would change in response to a given forcing. However it was found that a reasonable prediction of temperature at 70hPa could be obtained using each of at least 3 models which contained slightly different radiative and dynamical inputs.

Firstly the 70hPa temperature was modelled using the Eulerian-mean form of the thermodynamic equation with dynamical forcing consisting of: 1) upwelling calculated from ERA15 poleward mass flux between 70hPa and 10hPa, and 2) horizontal temperature advection which could not be attributed to the wave climatologies of chapter 3. Though a good prediction of temperature was obtained, this model was unrealistic because upwelling forced from above 10hPa was not included. Also there was no clear physical cause of the horizontal temperature advection that was used to force this model.

The second model differed in that upwelling was calculated using the sum of poleward mass flux between 70hPa and 10hPa in ERA15, and a parameterisation of poleward mass flux above 10hPa. Horizontal temperature advection was ignored since, if it was due to equatorial waves, it would be almost balanced by the term $\frac{\partial \omega}{\partial p}$ in the thermodynamic equation. This term could not be
obtained from the data available. The second model, which assumed that all of the horizontal advection fields were due to waves, was more accurate than the first model, which assumed that the linear wave climatologies of chapter 3 fully represented all wave disturbances. This suggests that the wave climatologies of chapter 3 may not be detecting all disturbances due to equatorial waves, and perhaps partly explains why the wave climatologies cannot explain the variance in dynamical fields (chapter 3) or drive the QBO (chapter 4).

The final model used to predict the 70hPa temperature had similar dynamical inputs to the second model, however the Transformed Eulerian-Mean (TEM) form of the thermodynamic equation was used. It was found that the difference in the predicted temperature that occurred between using the TEM and Eulerian-mean forms of the thermodynamic equation was much greater than could possibly be explained by equatorial waves like those found in chapter 3. This shows that disturbances in ERA15 that are unaccounted for have a larger affect on the Eulerian-mean circulation than disturbances attributed to linear waves. This highlights the need to improve wave theory and the wave detection method, a matter which is discussed in more detail in section 6.4.

One of the aims of chapter 5 was to understand the relative roles of radiation and dynamics in the annual temperature cycle. At 70hPa, if radiative parameters were held constant then dynamics could account for half of the annual temperature cycle. The other half of the annual temperature variations were explained by annual variations in the radiative timescale and the radiative equilibrium temperature. However the radiative timescale could only affect the annual temperature cycle by amplifying the effects of dynamics at certain times of the year.

The TEM form of the thermodynamic equation was also used to model the temperature of individual years at 70hPa. The annual cycle in temperature was mainly attributed to upwelling forced from above 10hPa, while interannual variations in temperature were mainly attributed to upwelling forced between 70hPa and 10hPa. At 100hPa the TEM thermodynamic equation was less accurate in modelling the temperature than at 70hPa, perhaps because of the effects of tropospheric processes on the radiative parameters at 100hPa which were not included in the model. At 50hPa the TEM thermodynamic equation predicted temperature in better agreement with two
radiosonde stations than the ERA15 temperature. This showed the importance of the parameterised upwelling forced from above 10hPa, and led to the suggestion that a dataset extending above 10hPa is required for an accurate representation of data below 10hPa.

## 6.4 Future work

### 6.4.1 Improvements in linear wave theory

This thesis has provided insights into dynamical processes operating in the tropical lower stratosphere. It has produced the first long-term, tropics-wide, climatology of linear equatorial waves in a dataset like ERA15. It has found the proportion of the QBO acceleration that can be explained by these waves and it has realistically modelled the annual temperature cycle in the tropical lower stratosphere. Despite the insights gained from this large-scale study of the tropical lower stratosphere, at each stage the investigation has been limited by the assumption of linear waves in a zero background flow. The linear waves found have been consistent with previous studies, yet they are unable to explain variance in dynamical fields, they account for only a small proportion of the acceleration needed to drive the QBO, and cannot account for the difference between the Eulerian-mean and the TEM circulation.

The next logical step is therefore to ascertain whether the role of equatorial waves (with period greater than 1 day) is really as small as this study suggests, by considering waves more representative of the atmosphere. Wave theory, as used in this thesis, considered dry, free, linear, waves in a zero background flow and an isothermal atmosphere. It is important to modify wave theory to include waves which are not subject to these constraints. For example, it would be useful to understand how the non-linearities of waves, which become important in certain conditions, would affect the waves’ structures. It would also be useful to consider how forced waves differ from free waves in order to ascertain whether some of the disturbances in the tropical lower stratosphere are forced. However, perhaps the first modification to the wave theory of chapter 2 should be to include the effects of vertical shear, associated with the QBO, in the wave equations. This first modification has already been done for two dimensional waves (Dunkerton, 1995) and
it appears that equatorial waves in a vertical shear zone could explain more variance in dynamical fields than equatorial waves in a zero background flow. However three dimensional linear wave theory in a vertical shear, remains to be developed.

Once linear wave theory in a vertical shear zone is sufficiently developed, linear waves in various vertical shears could be modelled. These waves could then be compared with linear waves in a zero background flow to give an indication of how to improve the wave detection method. This comparison would also suggest whether waves in a vertical shear have a greater effect on the QBO and the Eulerian-mean circulation than waves in a zero background flow. In a similar way, if wave theory could be enhanced to include the effects of non-linear and forced waves on the tropical lower stratosphere, a greater understanding of this region should arise.

6.4.2 Detecting individual equatorial waves

If wave theory could be enhanced, as described in the previous subsection, the next step would be to look for signals in the atmosphere that were consistent with the enhanced wave theory. Initially this should be done at individual timesteps and locations, to see the extent to which the enhanced wave theory could explain atmospheric disturbances. However insights could still be obtained by detecting waves at individual locations and timesteps using the wave theory described in chapter 2. Detecting individual wave events would provide information about wave amplitudes, where the waves occur, and how intermittent they are. This information would enhance our understanding of the role of equatorial waves in lower stratospheric dynamics. Individual wave events in 8 days of ERA15 have already been found by Yang et al. (2003). Perhaps the results of this thesis could reduce the computational time required to detect individual wave events in future studies by providing guidance as to those wavenumbers and frequencies, and to those seasons, where waves are likely to occur.
6.4.3 Enhancements to the wave detection method

This thesis detected equatorial waves using data between 7.5°N and 7.5°S. However including data from latitudes polewards of 7.5°, where many of the waves are still detectable, could improve the accuracy of the climatology. If, in addition, data were available with higher latitudinal resolution, the latitudinal structure of the waves could be better represented. It may then be possible to detect the waves with \( n \geq 2 \) which could not be accurately detected in this thesis.

The accuracy of the wave detection method could also be increased by decreasing the size of the wavenumber-frequency bins in which each type of wave was looked for. To provide the maximum increase in accuracy, at the minimum increase in computational cost, it is suggested that the size of the wavenumber-frequency bins should only be reduced at low wavenumbers-frequencies. Low wavenumber-frequency bins contain the largest power (because of the red nature of the spectrum) and also would contain the largest errors in detected waves (because the differences between waves of adjacent wavenumbers (frequencies) are largest at low wavenumbers (frequencies)).

After increasing the number of latitudes used to detect waves, and decreasing the size of the wavenumber-frequency bins, the climatology of waves in ERA15 could probably not be further improved unless amendments to linear wave theory could be included. However to understand waves in the real atmosphere it is desirable to combine results from this thesis with similar results derived from alternative data sources. A new reanalysis project, ERA40, has recently been completed at ECMWF which contains data from 1957-2001. ERA40 has higher horizontal and vertical resolution than ERA15, and therefore is likely to represent waves more accurately. A comparison of the wave climatology derived from ERA40 with the wave climatology derived from ERA15 would be useful for studies which aim to compare other processes in ERA40 with those in ERA15. A wave climatology from ERA40 would most likely provide information about factors such as the robustness of the trend in WIG1 and EIG1 waves that was found in chapter 3.
6.4.4 Improvements to observations

The Kelvin and Rossby-gravity waves detected in this study have been verified against previous observations from a variety of datasets. Ideally the characteristics of inertio-gravity and equatorial Rossby waves detected in this study should also be verified against independent observations; however, previous observations of these waves are limited. An atmospheric dataset in which these waves can be detected is therefore required.

A dataset with high vertical, horizontal, and temporal resolution is required to detect inertio-gravity waves. Radiosonde data would provide sufficient vertical resolution, and so an intensive observational campaign from a network of radiosonde stations could be devised to produce a dataset in which inertio-gravity waves could be detected. The network of radiosonde stations should be uniformly distributed in longitude and latitude, with the distance between two adjacent stations small enough to resolve the waves’ structures. It is important to include data from several longitudes to provide information about the direction of propagation of any wave found. Several latitudes should be considered such that the latitudinal structure of any detected waves could be verified against linear wave theory. In previous studies of inertio-gravity waves in the lower stratosphere the QBO has been easterly, and hence only EIG waves have been detected. It would therefore be most useful if the intensive observational campaign aimed to detect WIG waves by initially concentrating on the westerly phase of the QBO. Ideally, the observational campaign should cover a full cycle of the QBO, as this would allow clear comparison between eastward and westward propagating inertio-gravity waves.

This suggested campaign to detect inertio-gravity waves would allow verification of many of the findings of this thesis. For example:

- Do the same wavenumbers and frequencies of inertio-gravity waves dominate in the atmosphere as dominate in ERA15?
- Is the QBO related variation of inertio-gravity waves at 50hPa the same as was found here? (e.g. Is there really no QBO related variation at 50hPa associated with EIG waves?)
• Is the momentum flux due to inertio-gravity waves as small as was found in chapter 4, or is it more similar to the momentum flux for waves with periods 1-3 days found by Sato and Dunkerton (1997)?

• Are $u$, and $T$ largely in-phase at inertio-gravity wave frequencies, throughout the tropics, as we find? Or is this large tropical in-phase variation an artifact of the ERA15 data?

In a similar way the characteristics this thesis has found for the equatorial Rossby waves could be verified against the radiosonde observations.

There is no reason to suppose that inertio-gravity or equatorial Rossby waves will have the same characteristics throughout the tropics. Therefore it would be useful to have several campaigns to detect these waves at various locations in order that a more robust idea about these waves will emerge. Obtaining new data in which inertio-gravity and equatorial Rossby waves could be detected would also be useful for exploring those atmospheric disturbances that this thesis has been unable to attribute to any idealised linear equatorial waves.

6.4.5 Enhancements to the prediction of the annual temperature cycle

Though chapter 5 was able to produce a realistic model of the annual temperature cycle in the lower stratosphere, some uncertainties regarding the role of different processes in producing the annual temperature cycle remain. This is because an annual temperature cycle could be produced in three different models by varying the depth of perturbation used to calculate the radiative timescale and radiative equilibrium temperature. This study suggests values of the radiative timescale and the radiative temperature that could be applicable to the stratosphere, but if their true values were known the relative accuracy of each of the models used could be determined. Knowledge of the radiative state of the stratosphere may be improved by modelling the temperature using alternative datasets which have better resolution above 10hPa, such as ERA40 and NCEP reanalysis.

The annual temperature cycle at 100hPa was less accurately modelled than the annual temperature cycle at 10hPa.
cycle at 70hPa and 50hPa. This was attributed to the effects of convection on the relaxation to
equilibrium timescale and the equilibrium temperature, but was not investigated in any detail. To
improve the accuracy of the model at 100hPa, some representation of the effects of convection
should be included. This would give a physically more realistic model and would lead to a greater
understanding of the effects of different processes on various levels of the UTLS region.
Bibliography


BADC. URL http://badc.nerc.ac.uk/data/ecmwf-era/.


References


References


References


