The Influence of the Quasi-Biennial Oscillation on the Stratospheric Polar Vortices

Peter Watson
Wadham College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Michaelmas Term 2013

Abstract

The mean strengths of the wintertime stratospheric polar vortices are known to be related to the phase of the quasi-biennial oscillation (QBO) in the tropical stratosphere from circulation statistics – the "Holton-Tan relationship". The principal topic of this thesis is improving understanding of the mechanism behind the QBO's influence.

Following the example of previous studies, the QBO influence on the Northern Hemisphere (NH) extratropics on monthly time scales in an observational reanalysis is examined, and is shown to closely resemble the stratospheric Northern annular mode (NAM). It is argued that this may not be informative about the mechanism, as the response could be NAM-like for many different mechanisms. It is suggested that examining the transient response of the NH extratropics to forcing by the QBO would be much more informative, particularly on time scales of a few days.

In a primitive equation model of the middle atmosphere, the long-term stratospheric NH response to imposed zonal torques is often found to be NAM-like under perpetual January conditions, with wave feedbacks making a very important contribution. However, the response in runs with a seasonal cycle is not NAM-like. Investigation of the transient responses indicates the wave feedbacks are qualitatively similar in each case but only strong enough under perpetual January conditions to make the long-term response NAM-like. This supports the hypothesis that feedbacks from large-scale dynamics tend to make the stratospheric response to arbitrary forcings NAM-like, and therefore indicates that the long-term response is not generally useful for understanding forcing mechanisms. Examining the short-term transient response to known torques is found to be more successful at inferring information about the torques than several other previously proposed methods.

Finally, the short-term transient response of the NH extratropics to forcing by the easterly QBO phase in a general circulation model is found to be consistent with the proposed mechanism of Holton and Tan [1980], indicating that this mechanism plays a role in the Holton-Tan relationship.

Acknowledgements

I must firstly thank my supervisor Lesley Gray for all her help and encouragement, without which this work could not have been done. I also owe a debt of gratitude to David Andrews, who acted as my second supervisor up to his retirement at the end of my second year of study, and who provided many helpful theoretical insights. I am very grateful to Scott Osprey for providing instruction on how to run the models of the atmosphere I have used. I have also benefited greatly from many helpful discussions with all the members of the Oxford Stratosphere and Climate group, many of whom proofread parts of this thesis.

I am also very thankful to all those who have provided much-needed distractions from work during my time at Oxford, both here and further afield. There are too many important people to name them all, but I think special commendations must go to Sam and Ben for showing an especially high tolerance of me.

And last but certainly not least, I must thank my parents for all their help and support, and for impressively proofreading this entire thesis.

Contents

Chapter 1

Introduction

Understanding stratospheric dynamics and correctly representing it in numerical models is important for addressing several research questions which have impacts on general society. These include being able to forecast variability and trends in the stratospheric vortex that forms around each pole in wintertime and the impacts of this at the surface [e.g. Thompson and Solomon, 2002; Baldwin *et al.*, 2003b] and forecasting changes in stratospheric ozone [e.g. Newman *et al.*, 2001; Shepherd, 2008; Newman, 2010]. The principal topic of this thesis is an investigation of the variability in the wintertime polar vortices arising due to the quasi-biennial oscillation (QBO; section 2.5) in the tropical stratosphere. This is known from the statistics of vortex behaviour to be one of the most important phenomena that modulate the strength of the vortices, and is a source of predictability of vortex behaviour on annual time scales. The observed statistical relationship between the QBO and the vortices is known as the "Holton-Tan (HT) relationship" after Holton and Tan [1980] [Baldwin *et al.*, 2001; Gray, 2010; Anstey and Shepherd, 2013]. The impact of the QBO on the Northern Hemisphere (NH) stratosphere is discussed in the most detail as this has been the focus of most previous work, though the conclusions are also likely to be relevant to understanding its impact in the Southern Hemisphere (SH).

The main outstanding issues relating to the HT relationship are:

1. Understanding how the relationship arises through the atmospheric equations of motion and what physical processes are important.

2. Understanding which features of the QBO are necessary and which are sufficient to produce a relationship like that observed.

3. Improving numerical atmospheric models so that they produce this relationship with realistic magnitude and seasonal timing.

These issues are explained in more detail in section 2.7. The first item is the main focus of this thesis, though this is also important for making progress on the other two issues.

Whilst following a path to try to get a better understanding of this problem, I have found it necessary to consider the issue of how observational data and numerical models are best used to infer information about interactions between different parts of the climate system more generally. There are a number of problems in atmospheric dynamics where behaviour in one part of the atmosphere is known to be related to an external factor based on observational statistics, but it is not understood what processes give rise to these relationships and the problems have not yet been solved analytically. The mechanism by which the QBO influences the vortex is one member of this class of problems, with other examples being the mechanism by which the extratropical stratosphere affects the troposphere (discussed in section 2.8) and how the eleven-year solar cycle affects the stratospheric polar vortices [e.g. Labitzke, 2005]. Material presented here may be useful for designing studies to understand those problems and others in this class as well.

1.1 Outline of thesis

An overview of the theory of stratospheric dynamics that is relevant to the content of this thesis and of past studies that have examined the QBO, the polar vortices and their interaction is given in chapter 2. In chapter 3 the sources of data and models used in this thesis are described and the general method used to assess statistical significance is explained.

New results are discussed beginning in chapter 4, where it is shown that the 3D structure of the extratropical response on monthly time scales to the changing QBO phase in the ECMWF Re-Analysis (section 3.1) closely resembles a pattern called the Northern annular mode (NAM; section 2.3), the "leading mode of variability" in the extratropical stratosphere. It is suggested that this behaviour may be a manifestation of a phenomenon seen in other dynamical systems that show the responses to arbitrary forcings are often very similar to the leading mode of variability. Consequently it is argued that the relationship between the QBO phase and the vortices on monthly time scales, which has been the focus of most previous studies, is not useful for understanding the mechanism behind the relationship. Instead, it is suggested that examining the transient vortex response to forcing by the QBO on short time scales would be much more useful for understanding the mechanism, as this will be less affected by feedbacks which may be causing the response of the extratropical stratosphere to closely resemble the NAM.

In chapters 5 and 6, the long-term mean and transient responses respectively of the UK Met Office (UKMO) Stratosphere-Mesophere Model (SMM; section 3.3) to a set of imposed torques is examined with two objectives: firstly, to test whether feedbacks in the stratosphere do cause the response to an arbitrary forcing to closely resemble the annular modes, and therefore whether this can explain why the observed response to QBO forcing is NAM-like; and secondly to understand how feedbacks from the wave part of the circulation affect the response to an applied torque. The latter will help understanding of how the QBO affects the vortex as well as understanding of wave-mean flow interaction more generally. In chapter 7, these experiments with the SMM in which a known simple torque is imposed are used to test several methods that have been used in previous studies to understand forcing mechanisms using observational data or model output. It is found that the only method to robustly allow the known imposed forcing to be determined, using the model output, is to examine the short-term transient response.

In chapter 8, this method is used with the UKMO HadGEM2-CCS general circulation model (GCM) to examine the short-term transient response of the NH polar

vortex to QBO forcing. This response is found to correspond to the signature of the proposed mechanism of Holton and Tan [1980] and indicates that this mechanism plays a role in the way the QBO influences the vortex. This is in contrast to several studies that have argued this mechanism is not important based on analysis of the vortex response on longer time scales of two weeks or more.

The conclusions of this thesis and the progress that has been made on the issues discussed above are summarised in chapter 9, and suggestions for extensions to this work are proposed.

The material in chapters 4 and 8 includes and expands upon the analysis of Watson and Gray [2014].

Chapter 2

Background

The stratosphere is the layer of the atmosphere between the tropopause and the stratopause. The height of the tropopause varies between about 15 km at the Equator and about 7 km at high latitudes, and that of the stratopause is about 50 km. In terms of pressure levels the tropopause is at about 100 hPa at the Equator and 400 hPa at high latitudes and the stratopause is at about 1 hPa [Holton *et al.*, 1995; Shepherd, 2002].

The traditional defining difference between the stratosphere and the troposphere is that the temperature increases with height in the stratosphere, due to the presence of ozone which absorbs solar ultraviolet (UV) radiation (though other definitions are sometimes used [Shepherd, 2002]). This makes the air stable against vertical convection and accounts in part for the differences between tropospheric and stratospheric phenomena. The main features of the stratosphere and the modern theoretical understanding that are relevant to the topic of this thesis are described below.

2.1 The NH winter stratospheric polar vortex

Each NH winter the Arctic is plunged into the long polar night and the stratosphere cools by infrared radiation to space. This increases the Equator–Pole temperature gradient which increases the vertical zonal wind shear in accordance with thermal wind balance (a relationship that is approximately satisfied on seasonal time scales, which states that the vertical zonal wind shear increases in the NH as the meridional

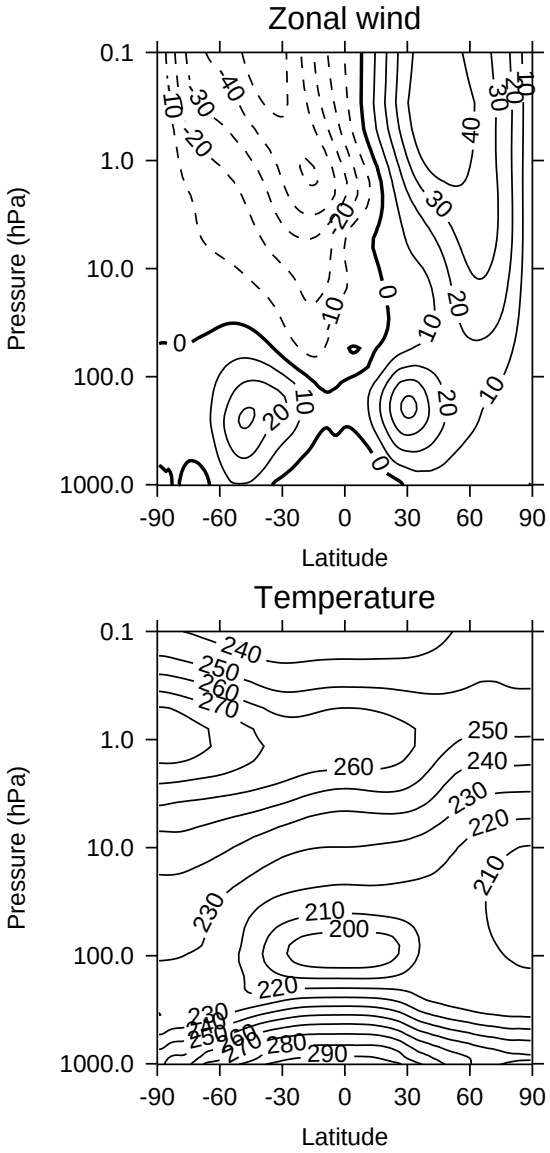

Figure 2.1: The DJF zonal mean zonal wind and temperature in ERA-40 (September 1957–August 2002) and ECMWF Operational re-analysis (September 2002–August 2008).

temperature gradient becomes more negative) [Andrews *et al.*, 1987]. This leads to a vortex of westerly winds forming in the stratosphere around the Arctic. The radiative-equilibrium state of the vortex therefore has a cold core of air above the Arctic circumscribed by a zonally-symmetric westerly jet [Shine, 1987]. Figure 2.1 shows the zonal mean zonal wind (ZMZW) and zonal mean temperature (ZMT) averaged over December-January-February (DJF) in the European Centre for Medium Range Weather Forecasting (ECMWF) Re-Analysis (ERA-40; section 3.1) and ECMWF Operational re-analysis. The westerly stratospheric vortex is visible in the NH, with the DJF ZMZW maximising between 0.1–1 hPa in mid-latitudes. The ZMZW at 10 hPa maximises near 60N at $\sim$30 ms^{-1}. There is a corresponding ZMT minimum of $\sim$210 K at the North Pole (NP) in the lower stratosphere. Whilst this is qualitatively similar to the radiative equilibrium state, the westerly vortex is less strong and the NP lower stratosphere is warmer than would be the case in equilibrium, due to the action of planetary (Rossby) waves [Shine, 1987]. These are described in more detail in section 2.2. The ZMZW and ZMT are strongly coupled by the thermal wind equation, so a "strong" vortex has high ZMZW and low high-latitude ZMT. Easterly winds are present in the SH stratosphere and the troposphere exhibits mid-latitude westerly jets.

A recent analysis of the vortex climatology using the ERA-40 re-analysis is provided by Mitchell *et al.* [2011a]. They found that the ZMZW averaged over all winters peaks in late December at (60N, 10 hPa) at $\sim$40 ms^{-1} and the mean 60–90N cap ZMT at 10 hPa reaches a minimum of $\sim$200 K. The centroid of the vortex (the centre of the best-fit ellipse to the vortex edge) is normally found near (50E, 80N) and the vortex is elongated with an aspect ratio of $\sim$1.5. The vortex variability tends to be greater at higher altitudes due to planetary waves having a greater impact, although the vortex also becomes more circular as radiative timescales decrease and there is less distortion due to the Aleutian high (a region of climatologically high geopotential height in the stratosphere, centred near (175W, 55N) at 10 hPa and tilting westward with height [Harvey and Hitchman, 1996]). Westerly NH winds are present from about August until April.

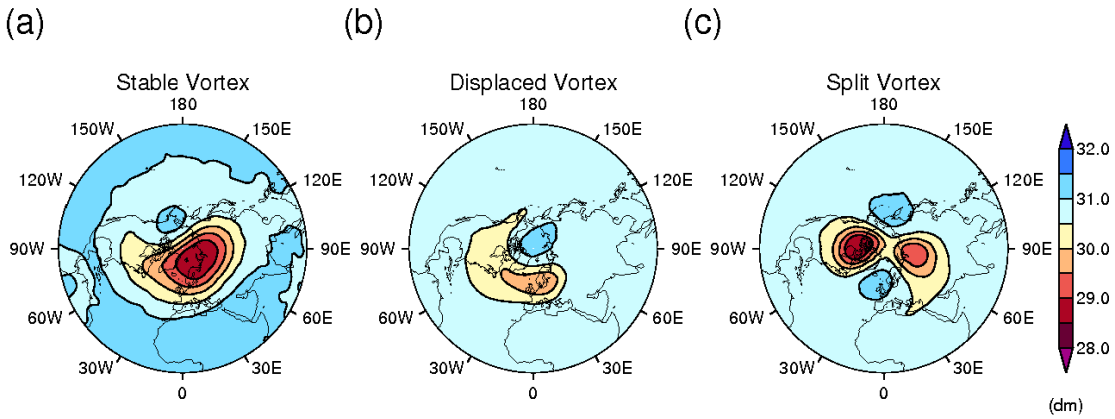

Figure 2.2: Daily averaged geopotential height fields for three different states of the NH polar vortex on the 10 hPa surface in the ERA-40 re-analysis. (a) A stable vortex (01/12/81), (b) a displaced vortex (10/12/87), (c) a vortex that has split into two daughter vortices (29/12/84). Red shows cyclonic motion and blue shows anticyclonic motion. The units are geopotential decameters. From Mitchell [2010]. Used with permission.

Figure 2.2 shows daily mean geopotential height (GPH) snapshots at 10 hPa. Figure 2.2(a) illustrates that typically low GPH, indicating cold air, is found over the pole, though planetary waves are present and the vortex is not completely circular.

2.2 Stratospheric waves

In the most general sense used in the stratospheric dynamics literature, a wave is a deviation from some mean state, normally taken to be the zonal mean. The most important type of extratropical stratospheric wave in wintertime is the large-scale Rossby wave, or planetary wave (section 2.2.1), whose action normally prevents the vortex state from tending towards the radiative-equilibrium state, reducing the ZMZW and raising the polar ZMT. The troposphere is a major source of planetary wave action, with wave structures forced by the flow of air around large-scale topography, planetary-scale heat sources and the statistically averaged effects of synoptic-scale eddies. For this reason the SH winter vortex is stronger and more stable than that in

the NH, since there is less planetary wave activity, due to there being less landmass [Shepherd, 2000; Plumb, 2010].

Another important type of stratospheric wave is the gravity wave [Shepherd, 2000; Fritts and Alexander, 2003]. This type of wave is generated by several processes, including air flow over rough topography, convection and frontogenesis. These make a net easterly contribution to ZMZW acceleration in the stratosphere and span a wide range of length and temporal scales, meaning they must usually be partly parameterised in stratospheric models [Geller *et al.*, 2013]. They make a larger net easterly contribution to ZMZW acceleration in the winter mesosphere, and this leads to driving of a meridional circulation (section 2.2.3) that warms the polar stratosphere. Gravity waves have also been suggested to affect variability of the wintertime polar stratosphere, including stratospheric sudden warming (SSW) events (section 2.4) [e.g. Rind *et al.*, 1988; Birner and Williams, 2008], and they contribute to wave driving of the QBO (section 2.5) and the equatorial stratospheric semi-annual oscillation (SAO; section 2.6).

2.2.1 Quasi-geostrophic Rossby waves

Large scale forced extratropical planetary waves obey quasi-geostrophic (QG) theory to a good approximation [Andrews *et al.*, 1987]. This theory applies to hydrostatically balanced, incompressible fluid flow in a rotating reference frame (taken here to be fixed with respect to the Earth's surface) when the typical horizontal velocity scale is much smaller than the product of the typical length scale of the flow and the Coriolis parameter, $f = 2\Omega \sin \phi$, and changes in the flow happen on time scales much longer than $1/f$. Here Ω is the angular velocity of the reference frame and ϕ the latitude. In what follows the fluid is assumed to be an ideal gas.

Following Andrews *et al.* [1987], the "geostrophic velocity" can be defined by the zonal component $u_g = -\Phi_y/f$ and the meridional component $v_g = \Phi_x/f$, where Φ is the geopotential height, x and y are the zonal and meridional coordinates on a local Cartesian grid and subscripts denote partial differentiation with respect to the subscripted variable. Under the quasi-geostrophic approximations in the absence of

external forcing, using beta-plane geometry such that $f = f_0 + \beta y$ where $f_0 = f(\phi_0)$ and $\beta = 2\Omega a^{-1}\cos\phi_0$, with a the Earth's radius and ϕ_0 a reference latitude, it can be shown that

$$D_g q = f_0 \rho_0^{-1}(\rho_0 Q/\theta_{0z})_z. \tag{2.1}$$

Here

$$D_g \equiv \frac{\partial}{\partial t} + u_g\frac{\partial}{\partial x} + v_g\frac{\partial}{\partial y}$$

where t is time and

$$q = f_0 + \beta y - u_{gy} + v_{gx} + f_0\rho_0^{-1}(\rho_0\theta_e/\theta_{0z})_z \tag{2.2}$$

is the "quasi-geostrophic (QG) potential vorticity (PV)". $z = -H\ln(p/p_s)$ is log-pressure height, p atmospheric pressure, $p_s = 1000\,\mathrm{hPa}$ and $H = RT_s/g$ where g is the acceleration due to gravity and T_s is a constant reference temperature. $\rho_0(z) = \rho_0(0)e^{-z/H} = p/RT_s$ and $\theta_0 = T_s(p_s/p)^\kappa$ are the corresponding reference density and potential temperature profiles respectively with $\kappa = R/c_p \approx 2/7$, where R is the specific gas constant and c_p the specific heat at constant pressure. Q is the diabatic heating rate and θ_e is the departure of the potential temperature from θ_0, which is assumed to be much smaller than θ_0.

Let the flow be approximately zonal with components $[\overline{u}(y,z), 0, 0]$. (Overbars denote zonal mean quantities in what follows.) Linearising equation 2.1 about this mean state gives

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)q' + v'\overline{q}_y = f_0\rho_0^{-1}(\rho_0 Q'/\theta_{0z})_z \tag{2.3}$$

where primed quantities are small deviations from the zonal mean such that, for example, $q = \overline{q} + q'$. For the case of constant $\overline{u}$ and $N^2 = H^{-1}R\theta_{0z}e^{-\kappa z/H}$ it can be shown this equation supports wave propagation, with vertical propagation dependent on the condition

$$0 < \overline{u} - c < \overline{u}_c = \beta(k^2 + l^2 + \epsilon/4H^2)^{-1} \tag{2.4}$$

where c is the wave's zonal phase velocity, k and l are the zonal and meridional wavenumbers respectively, and $\epsilon = f_0^2/N^2$. Equation 2.4 is the "Charney-Drazin criterion", after Charney and Drazin [1961]. For stationary waves forced by topography,

the condition is $0 < \overline{u} < \overline{u}_c$, meaning waves can only propagate vertically (from the troposphere to the stratosphere, for example) if the zonal wind is westerly and not too strong. Whilst the approximations that the flow is purely zonal, that $\overline{u}$ is constant and that the zonal asymmetries are much smaller than $\overline{u}$ are not realistic, the criterion captures qualitatively the fact that wave amplitudes in the stratosphere (meaning the size of departures from the zonal mean) are only substantial during winter when the winds are westerly, as expected if this is necessary for upwards wave propagation to take place, and are dominated by waves with small wavenumbers for which $\overline{u}_c$ is larger.

For more general mean states in which $\overline{u}$ varies with y and z with time-independent wave amplitudes, again following Andrews *et al.* [1987], the streamfunction of the zonally asymmetric part of the flow may have the form

$$\psi' = e^{z/2H}\mathrm{Re}[\Psi(y,z)e^{ik(x-ct)}]$$

where Ψ satisfies

$$\Psi_{yy} + \epsilon\Psi_{zz} + n_k^2\Psi = 0. \tag{2.5}$$

$$n_k^2(y,z) = (\overline{u}-c)^{-1}\overline{q}_y - k^2 - \epsilon/4H^2 \tag{2.6}$$

is the square of the "refractive index", by analogy with the equations that describe the propagation of light and sound rays. Under the assumption that the meridional and vertical length scales of the zonally asymmetric part of the flow are small compared to that of the zonal flow, such that n_k^2 varies relatively slowly spatially, locally sinusoidal wave solutions can be found such that equation 2.5 describes the propagation of these waves through space, with their group velocities tending to turn towards regions where n_k^2 is most positive, and with the waves becoming evanescent where $n_k^2 < 0$. If $\Psi(y,z)$ has the form $\mathrm{e}^{i(ly+mz)}$ then the Charney-Drazin criterion for vertical propagation is recovered by setting $m^2 > 0$ and $k^2 + l^2 + \epsilon/4H^2 > 0$ and approximating that $\overline{q}_y = \beta$.

This result can be generalised to spherical coordinates, with λ being longitude and $f = 2\Omega\sin\phi$, for which (neglecting heating and mechanical friction)

$$\left(\frac{\partial}{\partial t} + \frac{\overline{u}}{a\cos\phi}\frac{\partial}{\partial\lambda}\right)q'_{(M)} + a^{-1}v'\overline{q}_\phi = 0,$$

where

$$v' = (fa\cos\phi)^{-1}\Phi'_\lambda,$$

$$q'_{(M)} = \frac{1}{fa^2}\left[\frac{\Phi'_{\lambda\lambda}}{\cos^2\phi} + \frac{f^2}{\cos\phi}\left(\frac{\cos\phi}{f^2}\Phi'_\phi\right)_\phi + \frac{f^2a^2}{\rho_0}\left(\frac{\rho_0\Phi'_z}{N^2}\right)_z\right]$$

is a modified form of the QG PV, and

$$\overline{q}_\phi = 2\Omega\cos\phi - \left[\frac{(\overline{u}\cos\phi)_\phi}{a\cos\phi}\right]_\phi - \frac{a}{\rho_0}\left(\frac{\rho_0 f^2}{N^2}\overline{u}_z\right)_z. \tag{2.7}$$

Then substituting the wave solution

$$\Phi' = e^{z/2H}\mathrm{Re}[\Psi(\phi,z)e^{is[\lambda-(a\cos\phi)^{-1}ct]}]$$

where s is the zonal wave number, gives, for constant N,

$$\frac{f^2}{a^2\cos\phi}\left(\frac{\cos\phi}{f^2}\Psi'_\phi\right)_\phi + \frac{f^2}{N^2}\Psi_{zz} + n_s^2\Psi = 0.$$

$$n_s^2 = \frac{\overline{q}_\phi}{a(\overline{u}-c)} - \frac{s^2}{a^2\cos^2\phi} - \frac{f^2}{4N^2H^2} \tag{2.8}$$

is the refractive index squared, which has a similar form to that defined by equation 2.6. Equation 2.8 is obtained in a similar way to the derivation outlined by Andrews *et al.* [1987] for the case of stationary waves[1].

Despite being based on the assumption that zonal asymmetries are small and have small wavelengths, which is unrealistic in the stratosphere, this "ray theory" seems useful for understanding qualitatively how zonal asymmetries depend on the zonal mean flow. For example, the theory predicts that for a typical wintertime $\overline{u}$ profile, stationary wave propagation will be predominantly upwards at high latitudes and will turn towards the tropics with increasing height (figure 2.3). The waves turn towards the tropics largely due to the spherical geometry of the Earth [Karoly and Hoskins, 1982]. This is in agreement with calculations of the Eliassen-Palm (EP) flux (section 2.2.2), which indicates the flux of easterly zonal momentum associated with zonal asymmetry, and which is parallel to the wave group velocity under the

[1]Karoly and Hoskins [1982] give a derivation of a refractive index when N^2 is height-dependent, though this more general result does not strongly affect predicted wave behaviour in the stratosphere.

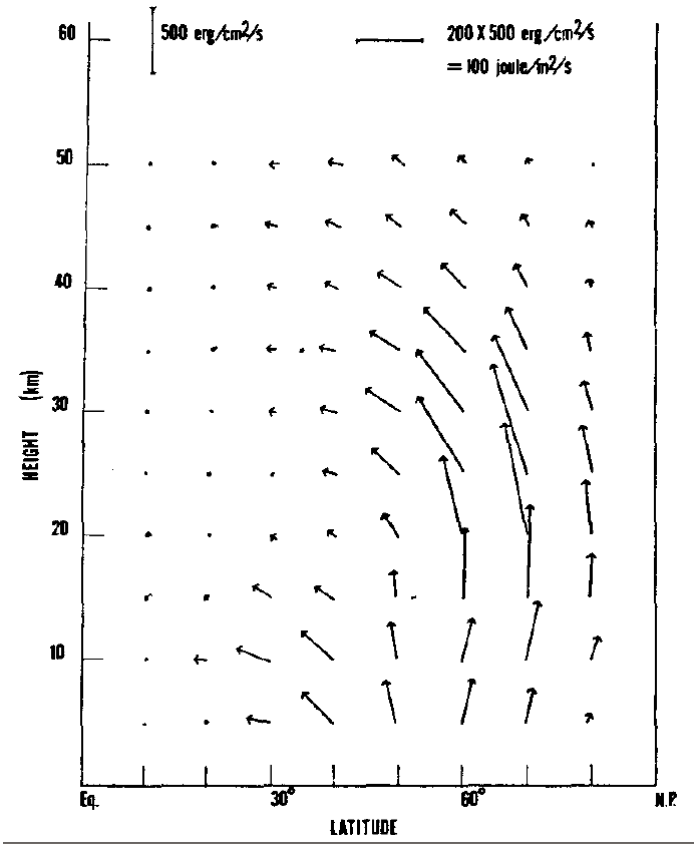

Figure 2.3: Energy flux in the meridional plane associated with zonal wavenumber 1 predicted by ray theory given the refractive index for planetary waves in a representative zonal wind profile for the NH during winter, from Matsuno [1970]. This compares well qualitatively with the climatological EP flux e.g. figure 3.2(b). © American Meteorological Society. Used with permission.

.

assumptions of ray theory. It is also argued in chapter 6 that the refractive index is useful for understanding some aspects of the way the stratosphere responds to an applied torque.

2.2.2 Eliassen-Palm flux

The influence of the wave components of the flow on the zonal mean component is better understood than the influence of the zonal mean component on the waves, which requires several simplifying assumptions to be made, not all of which are realistic, as described previously in section 2.2.1. Following Andrews *et al.* [1987], the

Eulerian-mean (EM) equations can be derived by taking the Eulerian zonal mean of the primitive equations of atmospheric flow. For example, the zonal mean zonal momentum equation is

$$\overline{u}_t+\overline{v}\left[(a\cos\phi)^{-1}(\overline{u}\cos\phi)_\phi - f\right]+\overline{w}\,\overline{u}_z = -(a\cos^2\phi)^{-1}(\overline{v'u'}\cos^2\phi)_\phi-\rho_0^{-1}(\rho_0\overline{w'u'})_z+\overline{X} \tag{2.9}$$

where u, v and w are the full zonal, meridional and vertical components of the wind and X is the zonal component of a mechanical forcing (such as friction). Note that the primed quantities can be large. Under the quasi-geostrophic (QG) approximation, this equation becomes

$$\overline{u}_t - f\overline{v}_a - \overline{X} = -(a\cos^2\phi)^{-1}(\overline{v'u'}\cos^2\phi)_\phi \tag{2.10}$$

where v_a is the ageostrophic part of the meridional flow.

The terms on the RHS of equation (2.9) are contributions to the zonal acceleration due to waves. Another EM equation relates the time derivative of potential temperature θ to the wave terms $\overline{v'\theta'}$ and $\overline{w'\theta'}$. The EM equations could be used to describe the zonal mean acceleration of the vortex. However, in their study of the 1976/7 major SSW, O'Neill and Taylor [1979] found that the terms $f\overline{v}_a$ and $(a\cos^2\phi)^{-1}(\overline{v'u'}\cos^2\phi)_\phi$ in equation 2.10 were much larger than $\overline{u}_t$ and mostly cancelled one another out. Physically, the convergence of wave zonal momentum flux is mostly balanced by the Coriolis force acting on an induced mean meridional circulation in this framework. This limits the usefulness of these equations for understanding wave-mean flow interaction.

This motivated development of the transformed Eulerian-mean (TEM) equations. The "residual mean meridional circulation" is defined as $(\overline{v}^*, \overline{w}^*)$, with

$$\overline{v}^* = \overline{v} - \rho_0^{-1}(\rho_0\overline{v'\theta'}/\overline{\theta}_z)_z \tag{2.11}$$

and

$$\overline{w}^* = \overline{w} + (a\cos\phi)^{-1}(\cos\phi\,\overline{v'\theta'}/\overline{\theta}_z)_\phi. \tag{2.12}$$

This circulation approximates the Lagrangian-mean circulation outside the stratospheric surf zone well, with the eddy terms approximating the Stokes drift of air parcels where there is wave activity [Dunkerton, 1978; Pendlebury and Shepherd, 2003]. (Other definitions of the residual circulation are also possible.) Substituting for $(\overline{v}, \overline{w})$ in the EM equations gives the TEM equations. The only TEM equations to contain wave terms are the zonal mean zonal momentum and thermodynamic equations:

$$\overline{u}_t + \overline{v}^* \left[(a\cos\phi)^{-1} (\overline{u}\cos\phi)_\phi - f \right] + \overline{w}^* \overline{u}_z = D_F + \overline{X} \tag{2.13}$$

and

$$\overline{\theta}_t + a^{-1}\overline{v}^*\overline{\theta}_\phi + \overline{w}^*\overline{\theta}_z - \overline{Q} = -\rho_0^{-1} \left[\rho_0 (\overline{v'\theta'}\,\overline{\theta}_\phi / a\overline{\theta}_z + \overline{w'\theta'}) \right]_z . \tag{2.14}$$

$D_F = (\rho_0 a\cos\phi)^{-1} \nabla \cdot \mathbf{F}$, where $\mathbf{F}$ is the EP flux, with components

$$F^\phi = \rho_0 a\cos\phi (\overline{u}_z \overline{v'\theta'}/\overline{\theta}_z - \overline{v'u'}) \tag{2.15}$$

and

$$F^z = \rho_0 a\cos\phi \left\{ \left[f - (a\cos\phi)^{-1} (\overline{u}\cos\phi)_\phi \right] \overline{v'\theta'}/\overline{\theta}_z - \overline{w'u'} \right\} . \tag{2.16}$$

The TEM equations describe the evolution of a zonal mean state $(\overline{u}, \overline{\theta}, \overline{v}^*, \overline{w}^*)$ for given wave forcing terms. In the QG approximation, the terms on the RHS of equation 2.14 are negligible, and so the wave forcing only enters in the zonal momentum equation as D_F.

The EP flux can be related to physical properties of the flow. If wave amplitudes are small, $\mathbf{F}$ satisfies an equation of the form

$$\frac{\partial A}{\partial t} + \nabla \cdot \mathbf{F} = D \tag{2.17}$$

where A is the "wave activity density" and D includes forcing and dissipation due to frictional and diabatic effects. A is in general a mean quadratic function of disturbance quantities. In the QG approximation $A = \frac{1}{2}\rho_0 \overline{q'^2}/\overline{q}_y$ where q is the QG PV, and $A > 0$ and may be taken as a measure of wave amplitude where $\overline{q}_y > 0$ [2]. $\mathbf{F}$ is the flux of

[2]Note that in the equatorial stratosphere, where the QG approximation is less accurate, A may be negative, for example where there are Kelvin waves [Andrews *et al.*, 1983].

wave activity density, and non-zero $\nabla \cdot \mathbf{F}$ indicates either dissipation or transience of wave activity. If the spatial and temporal scales of wave quantities are much less than those of variations in the background zonal flow, then $\mathbf{F} = \mathbf{c_g} A$, where $\mathbf{c_g}$ is a suitably defined group velocity. The EP flux of a small-amplitude QG Rossby wave is upward if the contours of constant phase tilt westward with increasing height, and is downward for the opposite tilt. The flux is poleward if the contours tilt westward with increasing latitude and equatorwards otherwise.

Additionally it can be shown that in the QG approximation, with no mechanical or diabatic forcing, that when $\nabla \cdot \mathbf{F} = 0$, $\overline{u}_t = \overline{\theta}_t = \overline{v}^* = \overline{w}^* = 0$ is a solution for the mean flow. This is a "nonacceleration theorem". Thus the wave component of the flow can be considered to force changes in $\overline{u}$ and $\overline{\theta}$ and to induce a residual mean meridional acceleration if and only if $\nabla \cdot \mathbf{F}$ is non-zero. Furthermore, in the QG approximation, $\nabla \cdot \mathbf{F}$ equals the zonal mean meridional transport of QG PV by eddies even if wave amplitudes are large [Edmon *et al.*, 1980].

Several early studies argued that the TEM equations are useful for understanding the zonal mean climatology and SSWs. Edmon *et al.* [1980] put forward that this framework is useful for understanding the zonal mean climatology of the troposphere. Palmer [1981a,b] argued that the TEM equations are useful for understanding major SSWs, noting that the EP flux convergence is quite well correlated with $\overline{u}_t$ in the 1979 and 1980 major SSWs, despite $f\overline{v}^*$ also being large, so that changes in $\overline{u}$ could be associated with changes in convergence of wave activity flux. The data were consistent with changes in EP flux being due to wave transience. Dunkerton *et al.* [1981] found similar results in a modelled major SSW, and Dunkerton and Baldwin [1991] found that the EP flux convergence is well correlated with the angular momentum tendency in data for ten winters. Andrews *et al.* [1983] argued that the TEM framework indicates that planetary waves act to decelerate the zonal mean flow, and that this is in better agreement than the EM picture with model simulations showing that damping waves in the lower stratosphere leads to the westerly zonal wind at high altitudes strengthening.

However, changes in EP flux may not always be best interpreted in terms of changes in the structure of planetary Rossby waves. For example, O'Neill and Pope [1988] argued that in their simulated major SSW, changes in EP flux reflected interactions between stratospheric vortices. They also stressed that the stratosphere is rarely in a zonally symmetric state as assumed for most of the above physical interpretation of EP flux. Even accepting this physical interpretation, it can be difficult to tell if a change in EP flux and its convergence is due to wave transience, dissipation, or non-linear effects [Andrews, 1987] – Rossby waves in the real atmosphere are large in amplitude, so non-linear effects can be important. With regard to understanding the influence of the QBO on the vortex (section 2.7), the QBO modulation of extratropical EP flux is only ∼20% of the climatological flux in the mid-stratosphere [e.g. Dunkerton and Baldwin, 1991]. This is much smaller than the changes that occur during SSWs, so non-linear and dissipative effects could be more important for the physical interpretation in this case, as they will not necessarily scale down by the same amount.

For the majority of this thesis the EP flux is interpreted simply as the flux of easterly zonal momentum associated with zonal asymmetry without using it to infer changes in wave propagation. However, in chapter 6 it is argued that some aspects of the EP flux response to a torque imposed in the stratosphere in a numerical model are consistent with a physical picture of changing planetary wave propagation.

2.2.3 Wave-driven meridional circulations

Convergence of stratospheric EP flux is important for "driving" large scale residual meridional circulations, including the Brewer-Dobson circulation in which air is transported from low to high latitudes [Holton *et al.*, 1995; Shepherd, 2000, 2002, 2007; Haynes, 2005; Plumb, 2010]. The residual circulation is driven in the sense that if the EP flux convergence (or more generally the total torque on the right hand side (RHS) of equation 2.13) and diabatic terms are specified, then the residual circulation can be calculated from the primitive equations, and in the absence of EP flux convergence a circulation can only be driven by diabatic effects, which are only effective if they

oscillate at high frequency. However it should be borne in mind that the EP flux also depends on the zonal mean circulation (section 2.2.1).

The residual circulation response to a torque has been studied by Eliassen [1951], Plumb [1982],Garcia [1987] and Haynes *et al.* [1991]. If there is convergent EP flux in the extratropics, as is generally the case in the wintertime stratosphere, the residual circulation at the position of the convergence is poleward. The Coriolis force acting on this circulation opposes the zonal mean deceleration associated with the EP flux convergence (since usually $|f| > |(a\cos\phi)^{-1}(\overline{u}\cos\phi)_\phi|$ in equation 2.13). The meridional mass flux is divergenceless, so in general, when the torque is time-dependent, there is a return equatorwards flow above and below the region of EP flux convergence and vertical motion at higher and lower latitudes, forming two overturning cells. The mass flux is generally greatest in the lower cell, where the air density is greater.

The structure of the residual circulation depends on the time-dependence of the EP flux convergence. In the steady state, the Coriolis force acting on the circulation completely balances the EP flux convergence. The circulation does not extend meridionally outside the region of the convergence, and extends downwards following a contour of constant angular momentum, with no upper circulation cell. Hence, the residual circulation at a given point is entirely determined by the EP flux convergence above that point [Haynes *et al.*, 1991]. This is known as "downward control". If the EP flux convergence is time-dependent then the circulation extends further meridionally and is less strong so the Coriolis force does not completely cancel the convergence, so $\overline{u}_t$ is non-zero. This is illustrated in figure 2.4.

The effect of a mechanical zonal torque ($\delta\overline{X}$) added to the RHS of equation 2.13, with the D_F perturbation (δD_F) associated with wave feedbacks also specified, can similarly be predicted: if $\delta\overline{X} + \delta D_F$ is negative, the effect is to weaken the ZMZW and drive a poleward residual circulation; if the sum is positive then the effect is the opposite. However, prediction of δD_F requires an understanding of the wave feedbacks. This has not been clearly elucidated in previous studies, but may be important for understanding the influence of the QBO and other natural forcings on

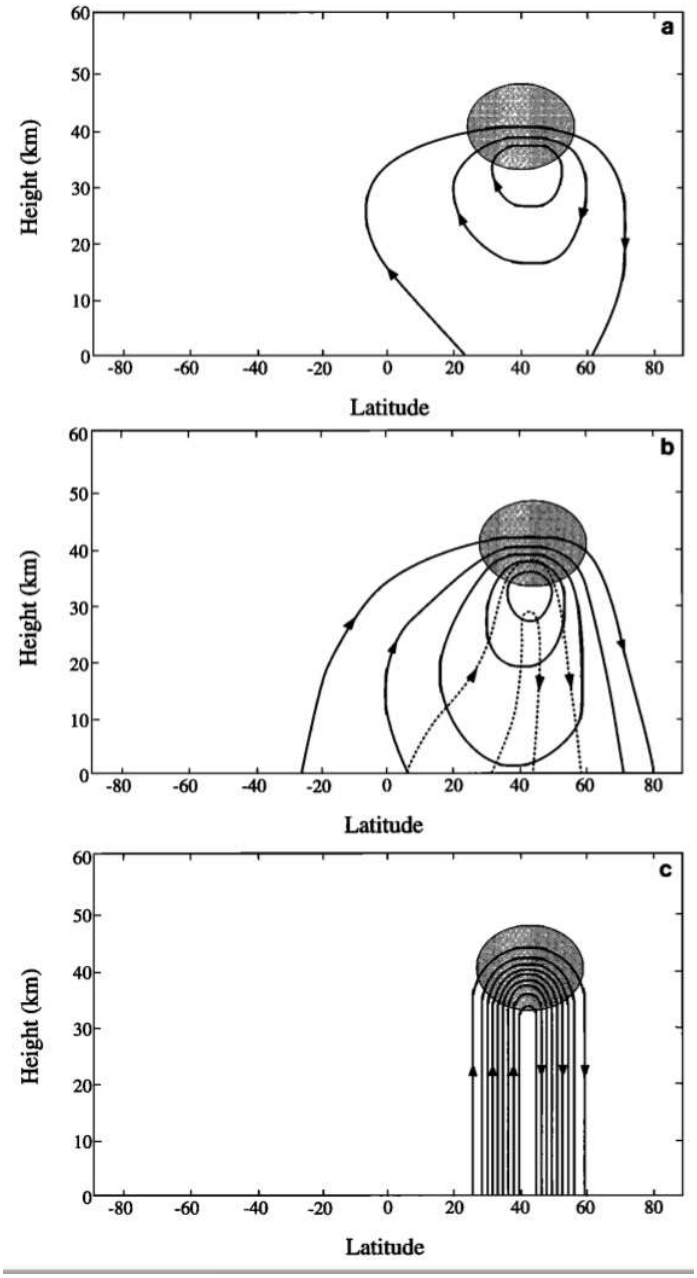

Figure 2.4: Idealised QG numerical experiments on the response of the mean stratospheric zonally symmetric circulation in the latitude-height plane to a westward force that is sinusoidal in time applied in the shaded region. Contours are streamlines, with the same contour interval used in each panel. (a) Adiabatic response for torque with period much shorter than radiative cooling time scale. (b) Response for torque with annual frequency if radiative time scale is 20 days; the solid and dashed contours show the response that is in phase and 90° out of phase with the forcing, respectively. (c) Steady state response. The model is unbounded below, so the ordinate has arbitary origin but can be regarded, for instance, as showing approximate vertical distance above the lower boundary of the "overworld" (the region where isentropes lie entirely in the stratosphere, above about 100 hPa). From Holton *et al.* [1995]. © Wiley. Used with permission.

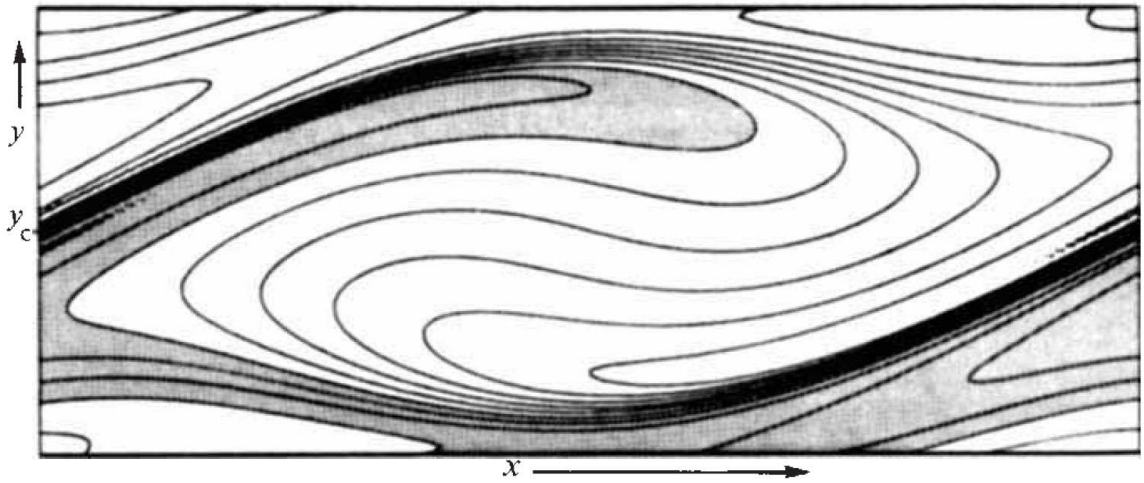

Figure 2.5: Analytical solution from the time-dependent theory of nonlinear critical layers exhibiting irreversible deformation of material contours due to advection by a two-dimensional (height-independent) 'Aleutian vortex' set up by a stationary Rossby wave on a shear flow. The shading picks out values of PV intermediate between the higher and lower values (unshaded) at bottom and at top/centre, respectively. The y scale is exaggerated. Initially the contours lie parallel to the x axis and represent a monotonic gradient of PV. The initial flow is in the x direction before the waves are excited, and its velocity is proportional to $(y - y_c)$ so that there is a 'critical line', where flow speed equals wave phase speed, at $y = y_c$. The time elapsed is 0.57 times the time taken for an air parcel to make one complete trip around the centre of the vortex or 'cat's eye'. From McIntyre and Palmer [1983]. © Elsevier. Used with permission.

the vortex, and also the effect of changing gravity wave parameterisations in models [e.g. Cohen *et al.*, 2013]. This is examined in chapters 5 and 6.

In the tropics f is small, with the result that wave-driven meridional circulations have a smaller meridional scale [Plumb, 1982]. An important example is the meridional circulation associated with the QBO (section 2.5).

2.2.4 Critical surfaces and layers

As waves propagate (using the language of linear wave theory) towards "critical surfaces" in the latitude-height plane where $\overline{u} = c$, for example the "zero wind line" in the tropics for stationary waves, linear analysis breaks down and nonlinear analysis becomes necessary [Shepherd, 2000; Plumb, 2010].

Understanding the behaviour of planetary waves that propagate near vertical critical surfaces may be very important for understanding the HT relationship, as discussed in section 2.7.2. Analytical solutions for inviscid, incompressible, approximately horizontal two-dimensional flow on a beta-plane in the vicinity of the critical surface were considered by Stewartson [1978] and Warn and Warn [1978], who showed a solution exists for which PV may be irreversibly stirred in a region of narrow meridional width called the "critical layer" near the critical surface (which is a 1D line in their 2D model). The PV contours take on a "cat's eye" structure (figure 2.5). Initially the critical layer absorbs the zonal momentum carried by waves incident upon it, but after some time it becomes perfectly reflecting. Killworth and McIntyre [1985] review this solution in detail and show it is numerically unstable, but that nevertheless in the absence of dissipation and forcing, advection of fluid parcels outside the layer and diffusion of PV, the layer must become reflecting in the long time average, even if wave amplitudes are large. Haynes [1985] and Haynes [1989] show by numerical integration that the aforementioned instability may severely disrupt the eye-like structure. Tung [1979] also showed for 3D motion that a vertically-oriented critical surface should be reflecting in the case that wave amplitudes are small (and it is this work that was referred to by Holton and Tan [1980] in their discussion of the possible mechanism by which the QBO affects the vortex).

Haynes [2003] gave a qualitative explanation of the time-averaged reflecting behaviour: if the layer sustained absorption of eddy momentum flux without dissipative processes cancelling the effects of this absorption then the critical layer would continually widen, which cannot happen indefinitely. A complementary viewpoint is that sustained absorption of wave momentum flux would mean the critical layer would experience zonal deceleration, and so would move continuously towards the wave source, which again cannot happen for all time. Plumb [2010] gives a qualitative explanation that for the "Stewartson-Warn-Warn" solution, stirring of the PV by the cat's eye structure removes meridional PV gradients, which implies that the mean meridional

transport of PV by eddies vanishes and then so does $\nabla \cdot \mathbf{F}$ in the QG approximation, such that wave activity density can no longer change with time within the layer (section 2.2.2).

In the real atmosphere there is dissipation of wave activity by diabatic and viscous processes, fluid parcels move in and out of the critical layer and motion is three dimensional. Therefore the result of Killworth and McIntyre [1985] can only be expected to hold approximately. The critical surface in the tropics must be at least partially absorbing of eddy zonal momentum flux to be consistent with the observed overall convergence of the flux in the tropics on monthly time scales [Killworth and McIntyre, 1985; Andrews *et al.*, 1987]. Structures similar to the cat's eye in figure 2.5 have been observed in the NH mid-latitude wintertime stratosphere, however, albeit with a much larger meridional scale [McIntyre and Palmer, 1983, 1984; Shepherd, 2000]. This indicates that the "breaking" of Rossby waves and the associated irreversible mixing of PV and flattening of meridional PV gradients in a "surf zone" does occur. Mid-latitude stirring of PV leads to sharp PV gradients forming at the vortex edge and in the sub-tropics. Juckes and McIntyre [1987] also captured this behaviour in a high resolution 2D model of the NH winter stratosphere.

In summary, the large-scale zonally asymmetric part of the extratropical stratospheric circulation seems consistent with the physical picture of planetary waves being generated by the troposphere with upward group velocities and entering the stratosphere. These waves' ray paths are steered according to the zonal mean circulation and eventually turn equatorwards (section 2.2.1). The waves break in the mid-latitude surf zone adjacent to critical surfaces in the tropics, as described above. On average the waves deposit easterly momentum throughout the extratropical stratosphere. Their effect on the zonal mean circulation can be understood using the TEM framework (section 2.2.2) – they cause the ZMZW to be weaker than it would be under radiative equilibrium conditions and drive the poleward stratospheric part of the Brewer-Dobson circulation (section 2.2.3).

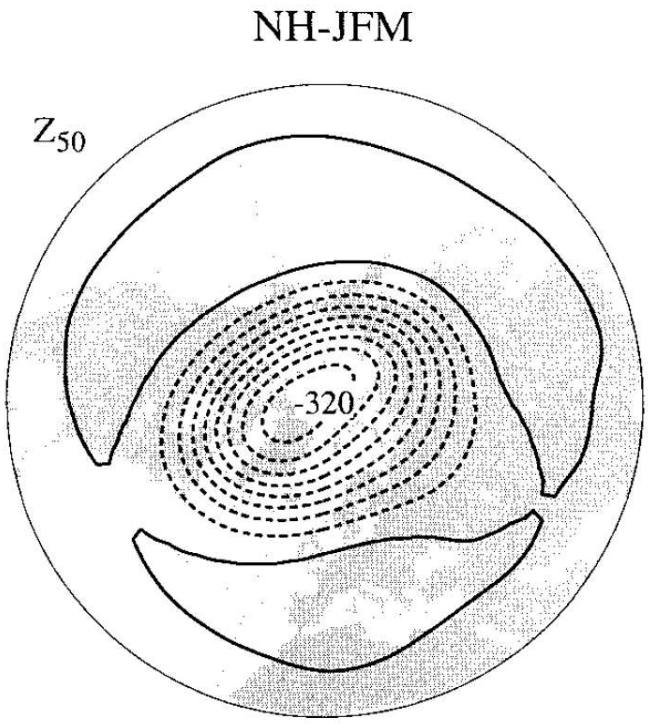

Figure 2.6: Leading modal structure of the monthly mean 50 hPa height anomaly field for January–March in the NH displayed as a regression map based on its standardised principal component time series. The contour spacing is 40 m with contours at (...,-60, -20, 20,...) m. From Thompson and Wallace [2000]. © American Meteorological Society. Used with permission.

2.3 Annular modes

The Northern and Southern annular modes (NAM and SAM respectively) are the leading modes of variability on monthly or seasonal time scales in each hemisphere, and are often analysed in studies of stratospheric change [e.g. Thompson and Wallace, 2000; Kushner, 2010]. They are often indexed by the leading principal component of extratropical GPH, either on one pressure surface or averaged over height and perhaps zonally averaged [e.g. Baldwin and Dunkerton, 1999; Thompson and Wallace, 2000; Baldwin and Thompson, 2009], and the resulting index is insensitive to the index definition [Kushner, 2010]. Variability in other quantities associated with changes in the annular mode indices can be found by linear regression of anomalies onto the indices.

In the wintertime stratosphere in both hemispheres, variability in the annular mode indices is associated with an approximately zonally symmetric change in the strength of the polar vortex (although as will be shown in chapter 4, the zonally asymmetric component is important for understanding the observed influence of the QBO on planetary waves). Figure 2.6 shows the regression of 50 hPa GPH onto the NAM index as defined by Thompson and Wallace [2000], with negative GPH at high latitudes and positive GPH in mid-latitudes. The monthly or seasonal mean response of the NH wintertime stratosphere to natural forcings such as the QBO, the El Niño-Southern Oscillation (ENSO), volcanic eruptions and the eleven-year solar cycle has been found to be quite similar to the NAM signature [Dunkerton and Baldwin, 1991; Kodera, 1995; Sassi *et al.*, 2004; Labitzke, 2005; Ruzmaikin *et al.*, 2005; Stenchikov *et al.*, 2006] and the importance of this for understanding how the QBO affects the vortex is discussed in chapter 4. The stratospheric NAM signature differs from the associated tropospheric variability, which is manifested as meridional shifts of zonal wind jets and is more zonally asymmetric, especially in the NH. The stratospheric signatures of the AMs also differ from those of the troposphere in that they only have a substantial magnitude in mid-winter in the NH and in spring in the SH and are partly forced by tropospheric variability, whereas the tropospheric signatures have a less pronounced seasonal cycle and their spatial structures are not likely to be strongly affected by the stratosphere, as they can be reproduced in simple atmospheric models without a stratosphere [Kushner, 2010]. Transport of momentum by waves also acts to support annular mode index anomalies in the troposphere [Limpasuvan and Hartmann, 2000] whereas in the stratosphere there is a complex interaction between the waves and the zonal mean flow that can give rise to oscillatory behaviour, as described in section 2.4.

2.4 Stratospheric sudden warmings

Sometimes major wave events result in SSWs, large disturbances of the vortex associated with a rapid increase in lower stratospheric polar cap temperature and a very zonally asymmetric flow. Such an event is defined by Andrews *et al.* [1987] as a reversal of the meridional ZMT gradient at 10 hPa or below between 60N and 90N. These are classified as "minor" if the ZMZW at (60N, 10 hPa) does not reverse and "major" if it does. Major SSWs are typically classified into two types: "vortex displacements" in which the vortex is displaced off the pole and distorted into a comma shape (figure 2.2(b)), and "vortex splits" during which the vortex separates into two separate cores of high PV air of comparable size (figure 2.2(c)). However, the definition of an SSW and the classification into minor warmings, displacements and splits is somewhat arbitrary. Major SSWs can simultaneously show characteristics of displacements and splits, and some minor SSWs can appear more disruptive to the vortex and be associated with greater polar cap temperature anomalies than some major SSWs [Mitchell *et al.*, 2013].

Charlton and Polvani [2007] defined major SSWs as occurring when the (60N, 10 hPa) ZMZW reverses and reported that including the ZMT gradient reversal criterion only excludes one of these major SSWs in ERA-40. They found that major SSWs occur with a frequency of about 6 per decade in ERA-40, most frequently during January and February, with displacements and splits occurring with a frequency ratio of about 1.2:1. Several minor SSWs occur in the NH each year.

Some work has focussed on explictly studying the vortex geometry during SSWs rather than zonal mean diagnostics in order to gain greater understanding [Waugh, 1997; Matthewman *et al.*, 2009; Mitchell *et al.*, 2011a; Hannachi *et al.*, 2011; Mitchell *et al.*, 2013]. Coughlin and Gray [2009] showed that about 10% of days in the October–March period display a disturbed vortex, including all minor and major SSWs.

It is generally accepted that SSWs are associated with greater upward EP flux entering the stratosphere from the troposphere and a greater convergence of this flux

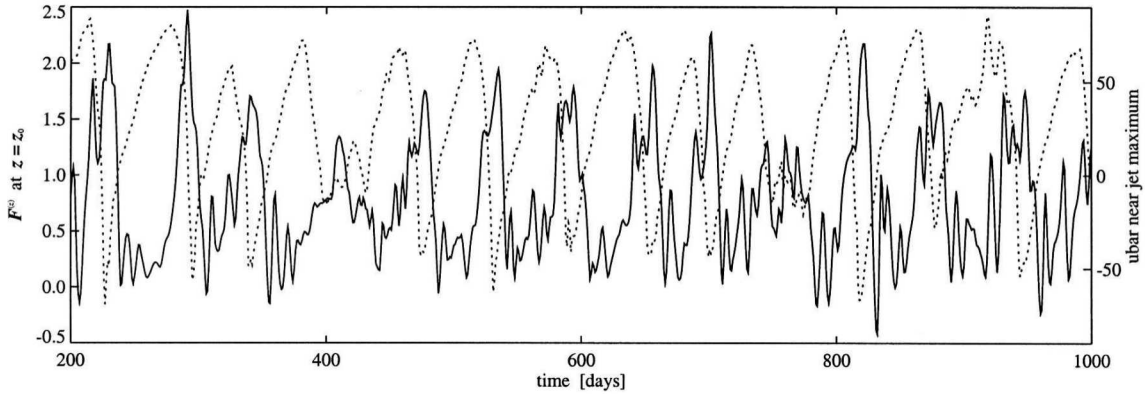

Figure 2.7: Vertical EP flux at the lower boundary averaged over latitude (solid) and $\overline{u}$ at 60N and $z = 41\,\mathrm{km}$ (dashed) in the primitive equation model of the stratosphere of Scott and Polvani [2006]. © American Meteorological Society. Used with permission.

in the high latitude stratosphere that causes deceleration of the vortex [e.g. Matsuno, 1971; Dunkerton *et al.*, 1981; Palmer, 1981a,b; Andrews *et al.*, 1987; Newman *et al.*, 2001]. This anomalous upward EP flux is dominated by a zonal wavenumber-1 component in the case of displacements and by a zonal wavenumber-2 component in the case of splits [Charlton and Polvani, 2007]. The warming process is approximately adiabatic and associated with descending air at high latitudes, with air rising and cooling at low latitudes to compensate [Andrews *et al.*, 1987]. The anomalous upward EP flux is normally associated with greater planetary wave amplitudes in the troposphere, but also depends on the stratospheric state, which may be "preconditioned" to be more receptive of upward EP flux from the troposphere owing to the dependence of wave propagation on the zonal mean flow (section 2.2.1), for example if the vortex is confined further polewards than usual [McIntyre, 1982; Charlton and Polvani, 2007]. Resonance may also play a role in creating large wave amplitudes [Plumb, 2010]. Birner and Williams [2008] suggested intermittent bursts of forcing by gravity waves may also trigger SSWs when the forcing by planetary waves alone is not quite sufficient. However, O'Neill and Pope [1988] and Scott and Dritschel [2006] argued the cause of SSWs may also sometimes be better thought of as an interaction between the polar vortex and the Aleutian High anticyclone.

The dynamics of SSWs bear some similarity to "vacillation cycles" observed in stratospheric models run under steady wintertime conditions with steady wave amplitudes at the tropopause [e.g. Holton and Mass, 1976; Yoden, 1987a,b; Christiansen, 1999; Scott and Haynes, 2000; Yoden *et al.*, 2002; Scott and Polvani, 2004, 2006; Scott *et al.*, 2008]. When wave forcing from the troposphere is sufficiently strong, the vortex cycles between a state with strong westerly winds and a weakened state with easterly winds. The cycles result from the non-linear interaction between the zonal mean and wave parts of the flow. As the vortex strengthens, upward EP flux into the stratosphere increases along with convergence of the flux, causing a rapid breakdown of the vortex and the ZMZW to become easterly. This causes a decrease in the upward EP flux in accordance with the theory of section 2.2.1 that allows the vortex to strengthen again due to radiative forcing, so the cycle repeats (figure 2.7). Yoden *et al.* [1996, 2002] and Christiansen [1999] identified similar dynamical behaviour in models with a temporally varying troposphere. This indicates that some portion of vortex variability arises from internal variability of the stratosphere, as well as from variability in the wave forcing from the troposphere.

Holton and Mass [1976] found, in their simple model representing wave-mean flow interaction in the stratosphere, that a sharp transition in the vortex strength and the magnitude of its variability occurs as wave forcing increases from zero, with the ZMZW strong and quiescent with low forcing and weaker and exhibiting vacillations when the forcing is above a certain threshold. O'Sullivan and Dunkerton [1994] found evidence of a similar sharp transition in the vortex state as wave forcing increases in a primitive equation model. Inclusion of a seasonal cycle smoothed this transition. Scaife and James [2000] and Gray *et al.* [2003] found a relatively smooth transition between the low-forcing low-variability and large-forcing large-variability vortex regimes in a primitive equation model. Christiansen [2003, 2010] argued that the distribution of the NH stratospheric October–March mean NAM index is bimodal, which may indicate the existence of preferred "strong" and "weak" vortex regimes in the real atmosphere.

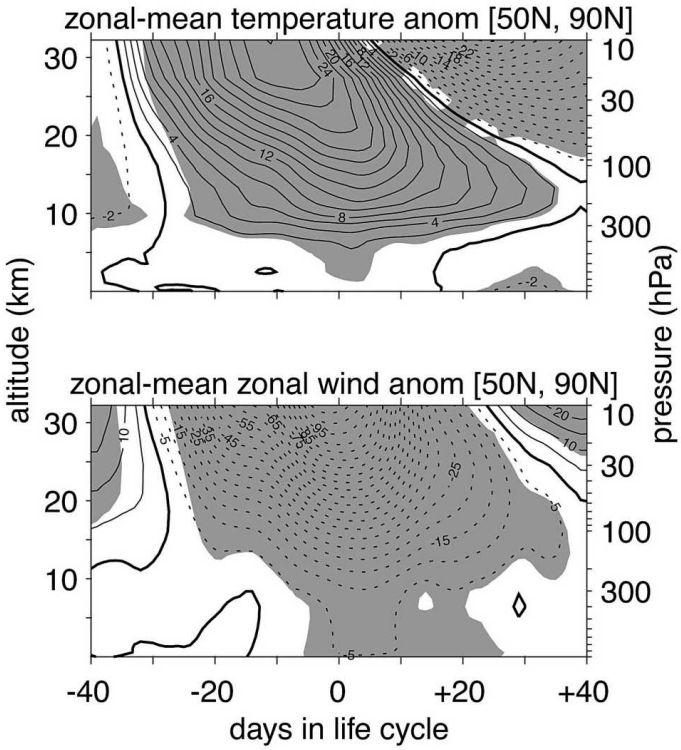

Figure 2.8: Zonal mean temperature anomalies and zonal mean zonal wind anomalies integrated poleward of 50N during the composite life cycle of SSWs. Negative contours are given as dashes. Zero contours are given as a bold solid line. Contour interval for temperature (zonal wind) is 2 K (5 ms^{-1}). Dark gray shading indicates areas with a 95% confidence level (based on t statistics). From Limpasuvan *et al.* [2004]. © American Meteorological Society. Used with permission.

SSWs evolve over a period lasting around two–three months. During the warming, the 50–90N 10 hPa polar cap ZMT rises by ∼30 K over ∼20 days and then falls again over roughly the next 30–40 days. Limpasuvan *et al.* [2004] show that the ZMT and ZMZW anomalies propagate downwards with time, decaying as the ZMT anomaly maximum moves below 50 hPa and into the upper troposphere (figure 2.8). A cold anomaly is found above the warm anomaly and descends, which means that at 10 hPa the ZMT averaged over the whole winter is not much affected by SSWs [Charlton and Polvani, 2007].

Matthewman *et al.* [2009] studied the evolution of vortex displacement and splitting events and found that individual events shared many features in common. Prior

to the ZMZW reversal in displacements, the vortex moves away from the pole in the direction of Europe to a latitude of around 65N, with its major axis approximately parallel to circles of constant latitude. The westward tilt of the vortex with increasing height becomes larger so that at the time of the ZMZW reversal the vortex spans nearly half the globe in longitude. Splits start with the vortex elongating and rotating cyclonically, with the elongation aligning in most cases with the 80E–100W longitude circle, set by the structure of the stationary wave forcing. Then the vortex separates into two daughter vortices almost simultaneously throughout the stratosphere, usually over Siberia and Canada, with the Siberian vortex normally being the larger of the two. Some SSWs combine the characteristics of displacements and splits, with the vortex being displaced towards Europe and then splitting into two, with daughter vortices over Siberia and Europe. Following both types of events, the vortex recovers over a period of several weeks, with much variability in the details of this between individual events.

2.5 The quasi-biennial oscillation

The QBO is a phenomenon in the equatorial tropical stratosphere whereby the ZMZW direction on a given pressure level alternates between being easterly and westerly, with the easterly and westerly wind regimes descending with time from the upper to the lower stratosphere. As each descends, a new wind regime of opposite sign forms above and begins its descent. The average period is 28 months, which varies between 22–34 months for individual cycles [Baldwin *et al.*, 2001]. The QBO dominates variability of the equatorial stratosphere, as can be seen in figure 2.9 which shows a time series of the 10S–10N mean ZMZW on all pressure levels in ERA-40, in which the QBO is clearly visible between ∼5–90 hPa. Pascoe *et al.* [2005] and Crooks and Gray [2005] found a QBO signal up to 1 hPa, although at this level variability is mostly due to the SAO (section 2.6). Below about 50 hPa, the QBO winds are close to being zonally symmetric, but at greater altitudes during the westerly phase, the propagation of

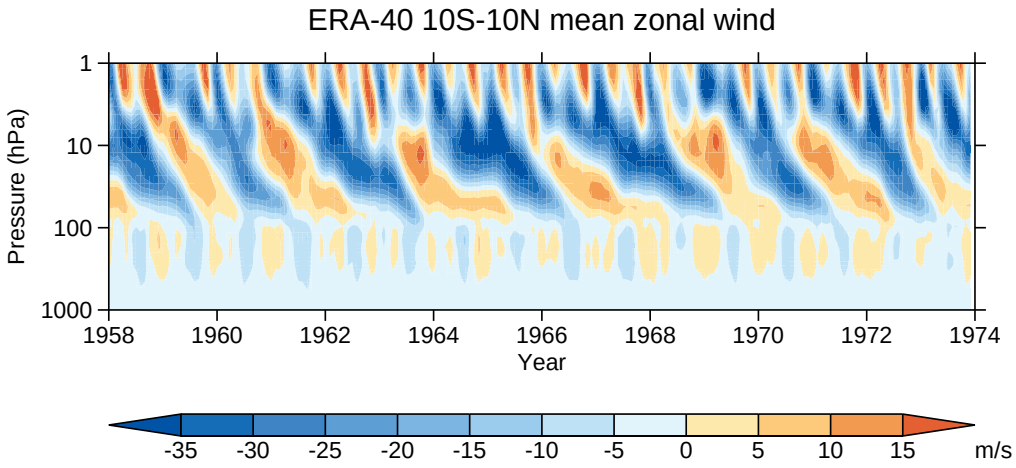

Figure 2.9: The 10S–10N ZMZW as a function of time and pressure in ERA-40, displaying the QBO.

planetary waves across the Equator can cause the zonal wind to vary with longitude [Hamilton, 1998]. The meridional half-width of the QBO is about 12°.

The state with easterly winds in the equatorial lower stratosphere and westerly winds above is termed QBO-E and that with westerly winds below and easterly winds above is termed QBO-W. Various quantities have been used to define the QBO phase, such as the sign of the zonal wind on a particular level [e.g. Holton and Tan, 1980] (the definition primarily used in this thesis), a phase angle defined using the leading two principal components of the equatorial wind [e.g. Wallace *et al.*, 1993] or the vertical shear of the zonal wind between two levels [e.g. Hitchman and Huesmann, 2009].

The easterly and westerly wind regimes are not symmetrical. The easterly winds are stronger than the westerlies, reaching about 20–25 ms^{-1} at 44 hPa compared to about 10–15 ms^{-1} for the westerly winds. The easterly wind regime descends more slowly, at a rate of about 2 hPa/month, than the westerly winds, which descend at about 4 hPa/month [Pascoe *et al.*, 2005].

Lindzen and Holton [1968] provided the first explanation of the QBO that is close to the current understanding. Additional important contributions were made by

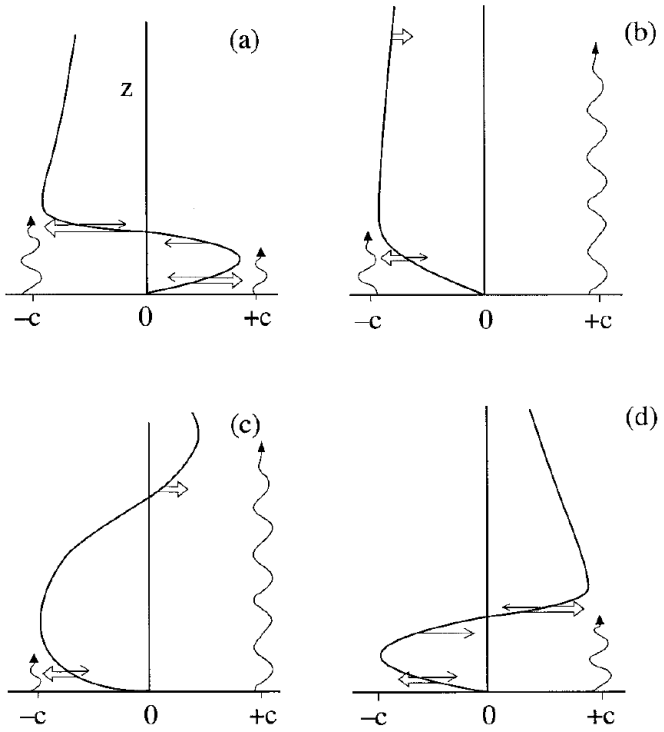

Figure 2.10: Schematic representation of the evolution of the mean flow in Plumb's [1984] analog of the QBO. Four stages of a half cycle are shown. Double arrows show wave-driven acceleration, and single arrows show viscously driven accelerations. Wavy lines indicate relative penetration of eastward and westward waves. After Plumb [1984]. © Terrapub. Used with permission.

Holton and Lindzen [1972], Plumb [1977] and Plumb [1984]. Briefly, waves generated in the tropical troposphere propagate upwards and are strongly dissipated when they approach a critical layer where the zonal wind is close to the wave's zonal phase speed. Thus if easterly winds are present in the lower stratosphere, an easterly-propagating wave whose phase velocity is below that of the maximum easterly wind causes easterly acceleration just below the layer, schematically illustrated in figure 2.10. Meanwhile westerly-propagating waves pass through the easterly wind layer until they meet a layer of westerly winds greater than the waves' phase speed, causing westerly acceleration below this layer and above the layer of easterly winds. In this way the layer of easterly winds moves downwards, until it reaches the layer where wave forcing is generated. Here it is eroded by viscous dissipation as the westerly winds descend from above and increase the magnitude of $\partial^2 u/\partial z^2$. The process repeats with westerly

winds in place of the easterlies until easterly winds are again present in the lower stratosphere, completing one QBO cycle. In the absense of very strong vertical diffusion, a state with no oscillation where the eastward and westward momentum deposited by waves balance is not stable and would not be expected to occur naturally [Plumb, 1977].

According to this model, the period of the QBO is determined by the total upward momentum flux and the distance between the wave source and the region of the semi-annual oscillation, and the amplitude is limited by the maximum wave phase speed. Dunkerton [1997] argued that known fluxes of Kelvin, Rossby-gravity and gravity waves from the tropical troposphere are sufficient to explain the observed QBO period, and that meridionally-propagating Rossby waves may also be important. Kawatani *et al.* [2010] showed that a high-resolution atmospheric model can simulate a fairly realistic QBO and that many wave types are important but argued there are insufficient observations to verify the realism of the model's wave forcing. Haynes [1998] argued that the QBO's width is confined due to divergence of wave momentum flux causing a much smaller zonal acceleration response in the extratropics than in the tropics as the torque is more effectively cancelled by the Coriolis force acting on an induced meridional circulation (section 2.2.3), although meridional gradients in wave sources may also be important. The Brewer-Dobson circulation, annual cycle and ozone advection are also important considerations for fully explaining the QBO [Dunkerton, 1997; Baldwin *et al.*, 2001; Gray, 2010].

The asymmetry in the descent rate in the easterly and westerly wind regimes is due in part to the presence of a meridional circulation associated with the QBO. A steady near-zonal circulation that satisfies hydrostatic balance and for which $\partial T/\partial y|_{\phi=0} = 0$ can be shown to satisfy

$$\frac{\partial u}{\partial z} = \frac{-R}{H\beta}\frac{\partial^2 T}{\partial y^2}$$

near the Equator in a similar way to the derivation of the thermal wind equation, without requiring f to be large [Andrews *et al.*, 1987] (this relationship also follows if the circulation is time-evolving but v is antisymmetric about the Equator, so $v = 0$

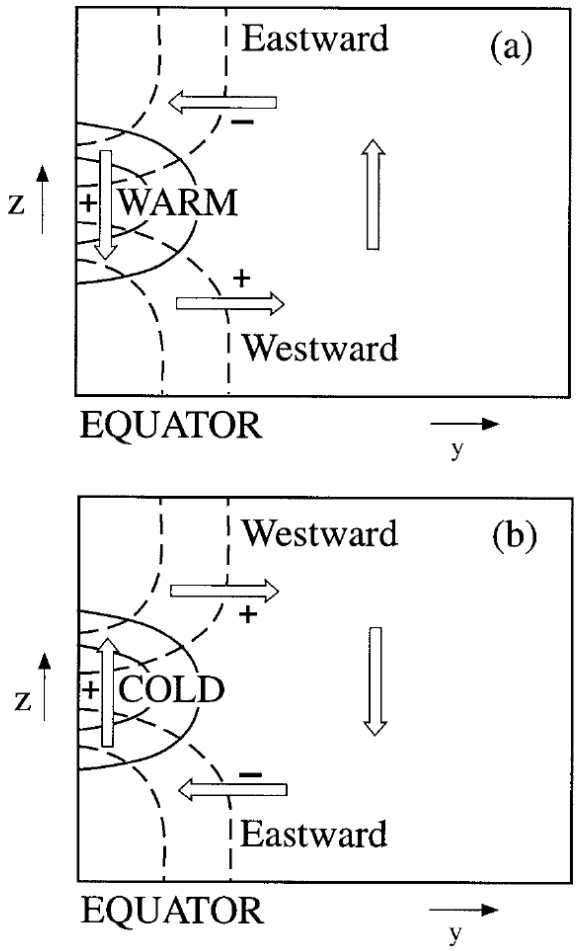

Figure 2.11: Schematic latitude-height sections showing the mean meridional circulation associated with the equatorial temperature anomaly of the QBO. Solid contours show temperature anomaly isotherms, and dashed contours are zonal wind isopleths. Plus and minus signs designate signs of zonal wind accelerations driven by the mean meridional circulation. (a) Westerly shear zone. (b) Easterly shear zone. After Plumb and Bell [1982]. © Royal Meteorological Society. Used with permission.

and $\partial v/\partial t + u\partial v/\partial x + v\partial v/\partial y + w\partial v/\partial z = 0$ on the Equator). This relationship implies the existence of a warm spot beneath the westerly winds and a cold spot beneath the easterlies, in agreement with observations [Randel *et al.*, 1999]. These are maintained by a combination of adiabatic heating/cooling by vertically-moving air and diabatic heating/cooling due to ozone variations brought about by this vertical motion [Plumb and Bell, 1982; Hasebe, 1994; Li *et al.*, 1995]. Thus the warm spot beneath the westerly winds is associated with descending air and the cold spot beneath the easterlies with ascending air, causing the westerly winds to descend more quickly. This is illustrated in figure 2.11. The easterly wind descent is slowed so much in some years that it stalls around 30 hPa during northern winter, due to the equatorial upwelling associated with the Brewer-Dobson circulation sometimes being strong enough at that time to fully resist the downward descent of the winds (for example, in 1965 in figure 2.9) [Pascoe *et al.*, 2005]. This helps bring about partial seasonal locking of the westerly-to-easterly transition, which occurs more frequently during April–July near 50 hPa than in other months [Dunkerton, 1990; Hampson and Haynes, 2004; Pascoe *et al.*, 2005]. The easterly-to-westerly transition exhibits seasonal locking to a lesser degree. The vertical motion at the Equator is part of a set of cells with vertical motion between latitudes of about 20–40° of opposite sign to that at the Equator, and hence there are also temperature anomalies of opposite sign at these latitudes [Baldwin *et al.*, 2001]. The meridional circulation is quite symmetric about the Equator around the equinoxes but not at the solstices, when the circulation in the winter hemisphere is much stronger than that in the summer hemisphere and extends into the extratropics [Jones *et al.*, 1998; Kinnersley and Tung, 1999; Kinnersley, 1999].

The QBO properties may exhibit variability over interannual time scales and longer. The locking to the seasonal cycle was weaker between the late 1970s and late 1990s than before and after this time [Anstey and Shepherd, 2008; Anstey *et al.*, 2010; Christiansen, 2010]. It has also been suggested that the QBO is influenced by the solar cycle [Crooks and Gray, 2005; Pascoe *et al.*, 2005; Frame and Gray, 2010] and Taguchi [2010] found that the QBO phase propagation is faster and its

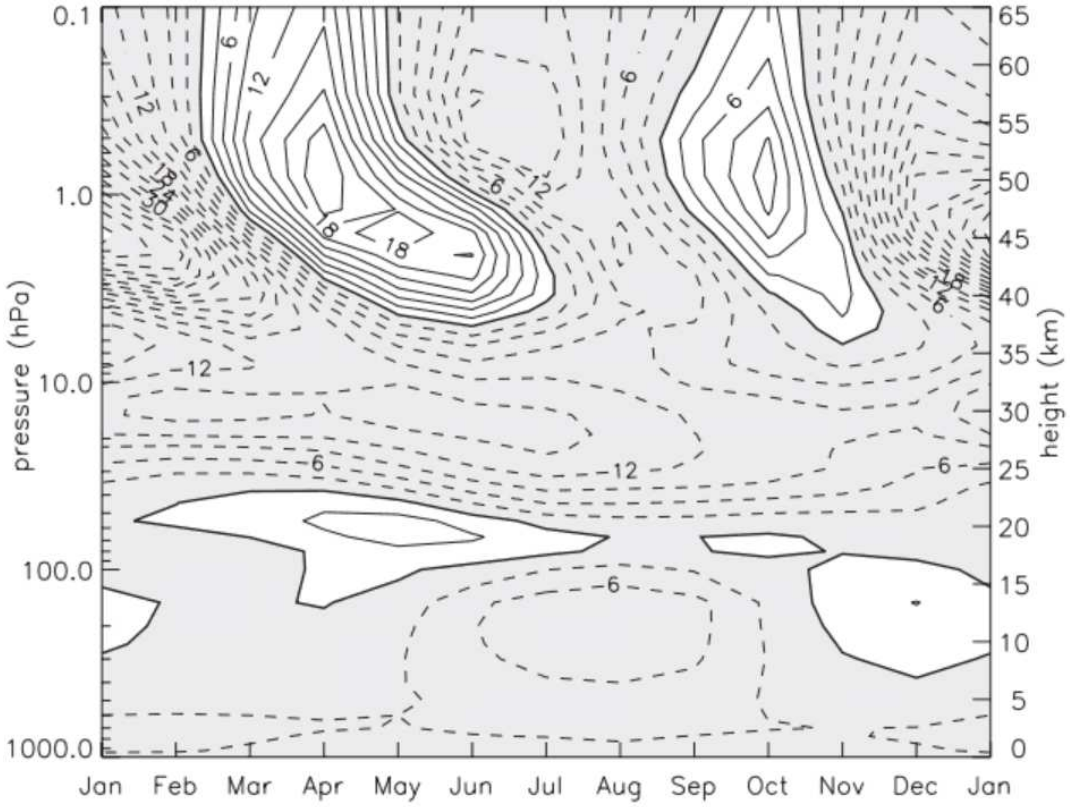

Figure 2.12: Annual cycle of equatorial ZMZW (1979–2008). Contour interval is $3\,\mathrm{ms}^{-1}$. From Gray [2010]. © Wiley. Used with permission.

amplitude is smaller during El Niño than during La Niña. Kuai *et al.* [2009] and Read and Castrejón-Pita [2012] have suggested there is intermittent synchronisation between the QBO and SAO.

Not all atmospheric models spontaneously produce a QBO. Important factors are having a fine vertical resolution in the stratosphere to better resolve short-wavelength gravity waves or parameterising the zonal momentum flux associated with such waves, a small diffusion coefficient to prevent smoothing of meridional gradients in the zonal wind, and a sufficient upward wave momentum flux produced by convection [Baldwin *et al.*, 2001].

2.6 The semi-annual oscillation

The SAO is an important component of variability above about 5 hPa at the Equator, recently reviewed by Gray [2010]. Figure 2.12 illustrates the seasonal cycle of equatorial ZMZW, showing easterly and westerly phases that descend with time, similarly to the QBO phases but occurring regularly in a 6-month cycle. The ZMZW time series shown in figure 2.9 illustrates that whilst the raw ZMZW time series shows the SAO, there is also substantial irregular variability in the uppermost stratosphere. The cycles in different halves of the year are not symmetrical, with the maximum easterly and westerly wind speeds being greatest in the cycle between December and June. The SAO is forced by upwards propagating equatorial waves that have a vertical group velocity fast enough that they avoid being damped below 5 hPa and which have zonal phase speeds outside the range of the QBO winds, so they do not encounter critical layers. There are also important contributions from meridionally propagating Rossby waves and from the advection of easterly winds from the summer hemisphere by the Brewer-Dobson circulation. These have large annual and semi-annual components in their temporal evolution due to the seasonal cycle and hemispheric asymmetry, and are likely to partly explain the locking of the SAO to the seasonal cycle and the annual variation in the SAO amplitude.

2.7 The Holton-Tan relationship

2.7.1 Observations and modelling studies

Numerous studies have found that the vortex is stronger during QBO-W than in QBO-E in observations dating from the late 1950s (figure 2.13) [Holton and Tan, 1980, 1982; Labitzke, 1982; Baldwin and Dunkerton, 1991, 1998; Dunkerton and Baldwin, 1991; Baldwin and O'Sullivan, 1995; Naito and Hirota, 1997; Gray *et al.*, 2001b, 2004; Pascoe *et al.*, 2005; Ruzmaikin *et al.*, 2005; Camp and Tung, 2007b; Lu *et al.*, 2008; Christiansen, 2010; Yamashita *et al.*, 2011; Mitchell *et al.*, 2011b]. The relationship is strongest when equatorial winds near 40–50 hPa are used to define the QBO phase.

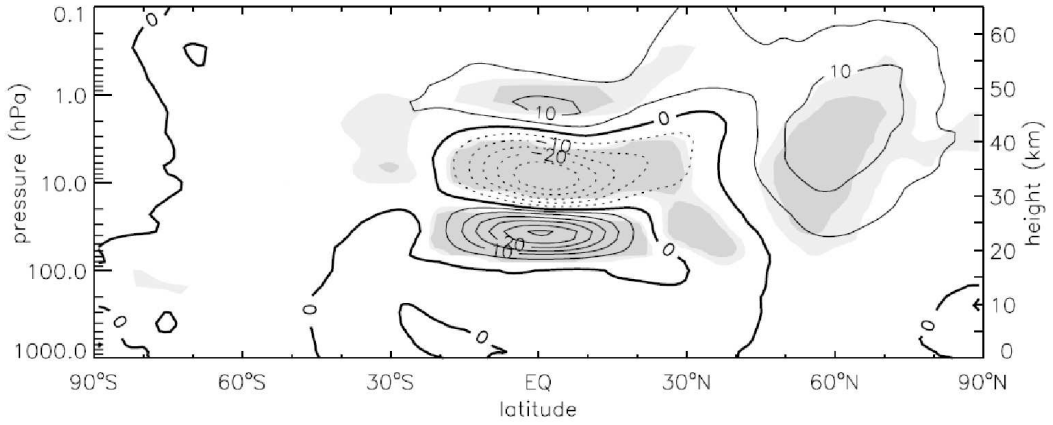

Figure 2.13: December–January zonal mean zonal wind difference for QBO-W minus QBO-E years with t test confidence shading shown at 95% and 99%; contours are at 5 ms^{-1} intervals. From Pascoe *et al.* [2005]. © Wiley. Used with permission.

The different observational studies generally agree on the magnitude of the QBO influence. Up to the 1990s the QBO and ENSO phase were correlated which made it difficult to clearly distinguish their effects [Wallace and Chang, 1982; Baldwin and O'Sullivan, 1995], but in later studies that used datasets that cover a longer time period the effects of the QBO and ENSO are more clearly separable [Garfinkel and Hartmann, 2007].

Lu *et al.* [2008] conducted one of the most recent analyses and studied a relatively long period, 1958–2006, in ERA-40 and ECMWF operational analyses, and found that the mean Nov–Jan ZMZW at (60N, 10 hPa) differs between QBO-W and QBO-E years by around 9 ms^{-1} and the ZMT of the polar lower stratosphere differs by up to about 4 K. The differences are largest in December and January, reaching $\sim$15 ms^{-1} and $\sim$5 K in (60N, 10 hPa) ZMZW and lower stratospheric polar ZMT respectively. February also exhibits smaller differences. The ZMZW and ZMT differences descend as winter advances, and in February a negative ZMT difference in the polar stratosphere above 10 hPa is present. This descent is somewhat similar to what happens during SSWs (section 2.4). The correlation between November–March mean 50 hPa equatorial ZMZW and (54N, 10 hPa) ZMZW between 1958–2006 is 0.64. Mitchell

et al. [2011b] found that the vortex-integrated PV is less during QBO-E than during QBO-W from November through February and that the vortex centroid is displaced more southwards in November.

In the latitude-height plane, the ZMZW difference in the northern extratropics between QBO-W and QBO-E is an unequal dipole, with small negative ZMZW differences south of ∼40–45N [Dunkerton and Baldwin, 1991; Pascoe *et al.*, 2005, figure 2.13]. Dunkerton and Baldwin [1991] found that the latitude where the differences change sign moves northwards through the winter, and Ruzmaikin *et al.* [2005] showed the GPH difference is very similar to the NAM signature.

The influence of the QBO on the SH vortex is less well studied. Baldwin and Dunkerton [1998] show the difference between QBO-W and QBO-E extratropical ZMZW is greatest in November near (70S, 5 hPa) where it reaches up to $14\,\mathrm{ms}^{-1}$, with the QBO phases defined using EOF analysis, roughly corresponding to using the sign of equatorial ZMZW at 25 hPa. This is the period of breakdown of the SH vortex, when it is more sensitive to wave activity. Naito [2002] and Anstey *et al.* [2010] found similar results in longer datasets. The QBO influence in the high-latitude SH seems relatively short-lived compared to that in the NH.

The HT relationship has been replicated in a variety of models, including a barotropic model [O'Sullivan and Salby, 1990], primitive equation models of the stratosphere and mesosphere with imposed wavenumber-1 GPH forcing at the bottom boundary and with QBO winds represented by a simple analytical expression [Holton and Austin, 1991; O'Sullivan and Young, 1992; O'Sullivan and Dunkerton, 1994; Hampson and Haynes, 2006] and with equatorial winds relaxed to their observed values [Gray *et al.*, 2001a, 2003, 2004], an atmospheric primitive equation model with a simplified representation of physical processes [Naito *et al.*, 2003], atmospheric GCMs with an imposed QBO [Hamilton, 1998] and with a spontaneous QBO [Niwano and Takahashi, 1998; Calvo *et al.*, 2007; Marshall and Scaife, 2009; Anstey *et al.*, 2010] and in chemistry climate models with an imposed QBO [Yamashita *et al.*, 2011] and a spontaneous QBO [Naoe and Shibata, 2010], to give some examples. Therefore the presence of the HT relationship is robust and insensitive to the choice of model, the

nature of the planetary wave forcing and the representation of the QBO. Few modelling studies have reported obtaining a seasonal evolution of the HT relationship that matches observations, however – indeed only Anstey *et al.* [2010] reported a clear HT relationship in November, when the correlation between equatorial and vortex ZMZW is greatest in observations.

There is evidence from modelling studies that the strength of the HT relationship is sensitive to the tropospheric wave forcing. Holton and Austin [1991] and O'Sullivan and Young [1992] found in their models of the stratosphere and mesosphere that sensitivity of the vortex to changes in the equatorial winds is small when the prescribed wave forcing at the bottom boundary is very small or very large, and greatest when the forcing is intermediate. Gray *et al.* [2003] found that imposing easterly equatorial winds in their primitive equation model had a similar effect on vortex variability to raising the wave forcing. This is consistent with the discussion of the different regimes of vortex variability that exist for different values of wave forcing in section 2.4, as at low levels of wave forcing the influence of the QBO-E phase is not great enough to cause disturbances in the vortex, and at large levels of forcing the vortex is very variable even during QBO-W, so that in each case the QBO has little influence. It is only when the wave forcing is near the strength at which the transition between these regimes occurs that sensitivity of the vortex to the QBO influence is large. Holton and Austin [1991] and Gray *et al.* [2003] suggested that the NH tends to be in the intermediate forcing regime, and that this allows external factors such as the QBO to have an influence. This may also explain why the HT relationship is only present in the SH during spring, when wave activity is greater than in mid-winter.

This non-linear behaviour means that observations and different models cannot always be easily compared because, for example, a model may simulate the equatorial wind influence on propagating planetary waves correctly but predict an incorrect strength of the HT relationship due to a climatological bias in wave forcing or the vortex. Various differences between models or between models and the real atmosphere

may combine to produce differences between the modelled and observed HT relationships, or cancel out to give good agreement despite differences in the underlying processes.

Despite there being a clear influence of the QBO on the wintertime mean vortex strength, there is not a clear influence on the observed SSW frequency. Lu *et al.* [2008] report that there were 15 SSWs in QBO-E and 16 in QBO-W between 1957–2006. However Mitchell *et al.* [2011b] show that 14 SSWs occurred during QBO-E and 6 during QBO-W in ERA-40 between 1957–2002. Other than choosing slightly different time periods, these authors use slightly different definitions of the QBO phases, with Lu *et al.* [2008] defining a winter to be in QBO-E/W according to whether the magnitude of the ZMZW anomaly at (0.56N, 50 hPa) exceeds $2\,\mathrm{ms}^{-1}$ and Mitchell *et al.* [2011b] defining the QBO-E/W winters according to whether the magnitude of 5S–5N mean 50 hPa ZMZW anomaly exceeds $5\,\mathrm{ms}^{-1}$. This illustrates that small changes in how the data are analysed can have a large effect on the apparent relative frequency of SSWs in each QBO phase, a point also discussed by Baldwin *et al.* [2001]. Naoe and Shibata [2010] found that, between 1980–2004 in ERA-40 and ERA-Interim (section 3.2), the distributions of daily mean ZMT in QBO-E and QBO-W differ most in the middle rather than in the high-ZMT tail associated with SSWs. Some modelling studies have examined the effect of the QBO on SSWs and generally agree that SSWs are more frequent or tend to happen earlier in QBO-E, suggesting that this influence may become apparent in observations as more data becomes available [O'Sullivan and Salby, 1990; Holton and Austin, 1991; O'Sullivan and Young, 1992; O'Sullivan and Dunkerton, 1994; Hamilton, 1998; Gray *et al.*, 2003, 2004].

The strength of the HT relationship may vary on decadal time scales. Lu *et al.* [2008] found the correlation between equatorial winds and the vortex strength averaged over February–March to be much weaker during the period 1977–1997 than before and after. They did not explain this change but noted that it coincided with a shift in north Pacific climate. Anstey and Shepherd [2008] and Christiansen [2010] suggested the change may be related to the QBO phase transitions becoming less

strongly locked to the seasonal cycle in this period. Christiansen [2010] argued this could have occurred by chance. Other natural forcings on the vortex may also have affected the strength of the QBO's influence (section 2.7.3). However, it is not yet clear if a true change in the relationship between the vortex strength and equatorial winds occurred in the observed period.

2.7.2 Mechanism of the Holton-Tan relationship

Whilst the existence of the HT relationship in observations and models is very robust, understanding the mechanism is still important for having confidence in observations of apparent interactions with other forcings, such as with the eleven-year solar cycle and ENSO (section 2.7.3) and in the seasonal timing of the effect, which is still not well reproduced by models. It is also important for knowing what models must represent well in order to reproduce the HT relationship and exhibit realistic vortex variability.

Most previous studies of the mechanism have worked in the framework of a particular conceptual model of how the QBO and the vortex interact, namely that the QBO and the vortex can be considered as separate phenomena and that the behaviour of the QBO is largely independent of the behaviour of the vortex, so that it can be considered as an external forcing acting on the vortex. No studies have suggested that the QBO properties are strongly dependent on the vortex state, consistent with this model. It also seems sensible, given that the primary reason for the vortex existing is the radiative cooling of the wintertime polar stratosphere (section 2.1), which is not very dependent on the phase of the QBO, and that the primary influence on QBO behaviour is the interaction of upward-propagating equatorial waves with the mean flow in the tropical stratosphere (section 2.5), which does not depend strongly on the state of the vortex. Moreover, all the mechanisms that have been proposed essentially agree that the QBO causes a zonal torque to be exerted on the high-latitude stratospheric flow in winter, but differ regarding how this torque is produced. However, given that the QBO and the vortex both exist in the same atmosphere, other conceptual models which consider both to be more fundamentally inter-dependent may also lead to fruitful insights, though this is not developed further in this thesis.

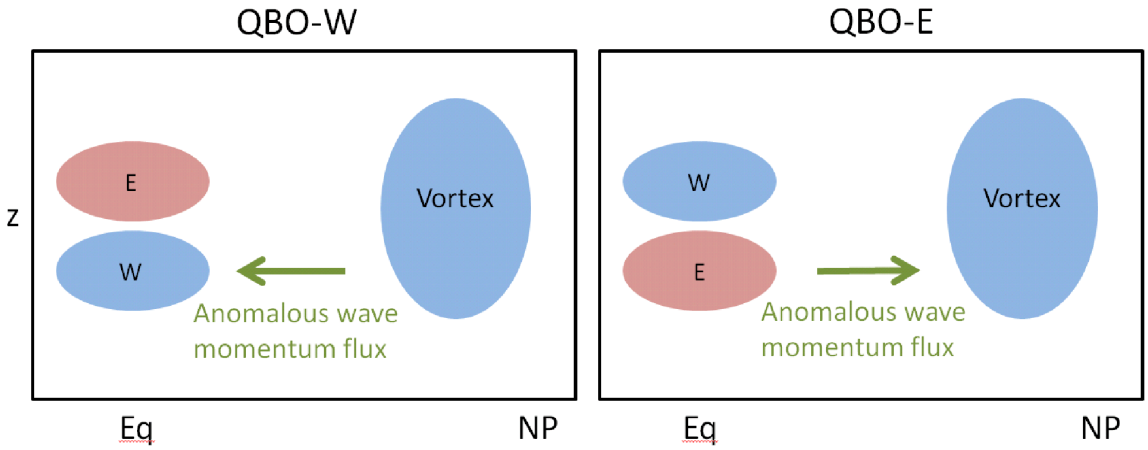

Figure 2.14: A schematic illustration of the Holton-Tan mechanism. NH wintertime zonal wind is shown in colour (blue is westerly and red is easterly) and the anomalous momentum flux associated with zonal asymmetry is shown by green arrows for QBO-W (left) and QBO-E (right) in the latitude-height plane between the Equator and the North Pole (NP). The anomalous flux is poleward (corresponding to the absolute flux being less equatorward) during QBO-E.

The explanation for the HT relationship put forward by Holton and Tan [1980] involved the equatorial winds influencing the waveguide for extratropical planetary waves. Low-wavenumber stationary planetary waves dominate wave forcing of the extratropical NH stratosphere, and the tropical zero wind line is a critical surface for these waves. Holton and Tan [1980] referred to the work of Tung [1979] who argued that this surface ought to reflect small-amplitude planetary waves back towards the pole – an extended discussion of the physics of this process was given in section 2.2.4 which suggests this may also be true for large-amplitude waves. In QBO-E, the critical surface in the lower stratosphere is moved polewards into the NH subtropics, so Holton and Tan [1980] suggested that this would concentrate wave activity in the NH polar region. This would weaken the vortex through convergence of EP flux. The opposite would occur in QBO-W. Henceforth this will be referred to as the "Holton-Tan mechanism", illustrated schematically in figure 2.14 (based on the arguments of Naoe and Shibata [2010] – Holton and Tan [1980] did not themselves explicitly predict the direct influence of the QBO on the EP flux). McIntyre [1982] suggested that movement of the lower stratospheric critical surface could affect resonance of

planetary waves in the extratropical stratosphere as well. In three dimensions, the effect of the shift of the critical surface may be to enhance the breaking of Rossby waves in the extratropical stratosphere, thereby weakening the vortex [e.g. O'Sullivan and Salby, 1990].

This mechanism correctly predicts that the HT relationship is strongest in winter months in the NH, when wave activity is strongest, and it can explain why the QBO influence on the SH vortex is only apparent during southern spring [Baldwin and Dunkerton, 1998], as wave activity in the SH is weaker and less able to affect the strong SH vortex in mid-winter. However, it is not clear from theory how reflective critical surfaces in the real atmosphere are, or whether any increased reflection of wave momentum flux would directly affect polar latitudes. Holton and Tan [1980] were careful to say that their proposal was speculative.

One other possible mechanism by which the QBO may influence the vortex was proposed by Kodera [1991], who argued that the ZMZW anomalies associated with the QBO meridional circulation (figure 2.11) may also affect planetary wave propagation. Ruzmaikin *et al.* [2005] suggested the meridional circulation may directly affect the vortex.

Observations of the modulation of extratropical planetary waves by the QBO appear to provide only partial support for the HT mechanism. The GPH wavenumber-1 amplitude and associated upwards EP flux (section 2.2.2), are greater in November and December in QBO-E, but in January and February these values are greater during QBO-W (although the difference in these months is not highly statistically significant) [Holton and Tan, 1980, 1982; Dunkerton and Baldwin, 1991; Hu and Tung, 2002; Ruzmaikin *et al.*, 2005]. Given that the vortex and planetary wave structures are present throughout all these months, at first sight it seems difficult to explain the difference in the modulation of planetary waves between the November–December and January–February periods. For example, Hu and Tung [2002] concluded "it appears likely that the Holton-Tan mechanism is applicable only in early winter". In section 4.2 it will be argued that the observations *are* reconcilable with the HT mechanism.

Modelling studies also only appear to provide partial direct support for the HT mechanism. Studies in models that use simpler dynamics than that of the primitive equations include that of O'Sullivan and Salby [1990], who show in an equivalent barotropic model of the stratosphere, initialised from a January climatological state, that the flow develops stronger low-wavenumber components when equatorial winds are easterly. Their interpretation was that QBO-E is associated with greater horizontal mixing by planetary waves. The zonally symmetric linear quasi-geostrophic model with parameterised planetary wave breaking of Hauck and Wirth [2001] showed that QBO-E gives greater wavenumber-1 and 2 GPH amplitudes and a much larger stratospheric EP flux convergence, which may indicate greater deceleration of the vortex by planetary waves (section 2.2.2), albeit with the largest differences in mid-latitudes.

Holton and Austin [1991], O'Sullivan and Young [1992], O'Sullivan and Dunkerton [1994] and Gray *et al.* [2003] examined the HT relationship in primitive equation models of the stratosphere. Holton and Austin [1991] found that for intermediate planetary wave forcing, wavenumber-1 GPH amplitudes peak faster and reach a greater maximum in QBO-E compared to QBO-W. However, EP flux differences near the equator were found to be small initially, so any influence of the equatorial QBO on planetary waves could not be understood in terms of ray theory. O'Sullivan and Young [1992] observed that EP flux convergence was greater at high latitudes in QBO-E for moderate imposed wave amplitudes at the tropopause, but with very small amplitudes there is only a difference in the sub-tropics, implying there is a non-linear dependence on the wave amplitude. O'Sullivan and Dunkerton [1994] found that wavenumber-1 amplitudes are greater in QBO-E than in QBO-W in the polar lower stratosphere. Gray *et al.* [2003] observed that QBO-E favoured upwards EP flux in the upper stratosphere two to three months into their integrations, but that in the lower stratosphere the EP flux was more downwards, and EP flux convergence was only enhanced in the mid-latitude upper stratosphere.

The aforementioned studies all used models which have simplified dynamics and which suffered from one or more of the following deficiencies: that the forcing from the troposphere was imposed as a low-wavenumber perturbation which does not react

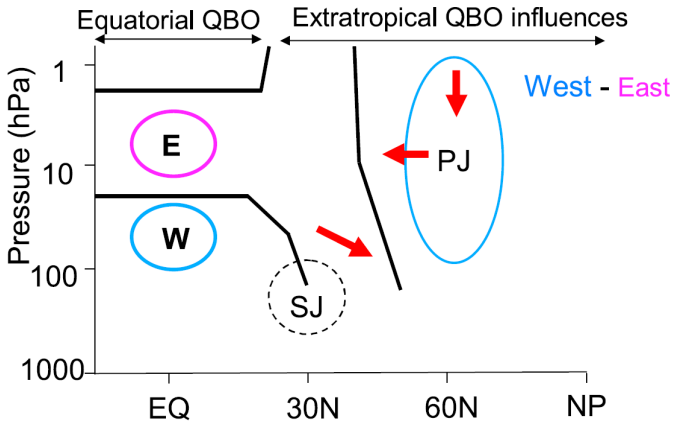

Figure 2.15: A schematic illustration of the EP flux difference (red arrows) between QBO-W and QBO-E by Naoe and Shibata [2010] that is consistent with their proposed mechanism for the HT relationship. Black contours indicate a zero difference in zonal wind between the easterly and westerly phases of the QBO. W is in the westerly phase of the QBO, E is in the easterly phase of the QBO, SJ is the subtropical jet, and PJ is the polar jet. Westerly anomalies are blue, and easterly anomalies are pink. © Wiley. Used with permission.

to the stratospheric state; the QBO structure was not realistic; or the initial conditions were zonally symmetric. Using a GCM could improve the representation of the dynamics and the atmospheric state, and examining the mechanism behind the HT relationship in GCMs is important for checking the robustness of the above results.

Several studies have examined the role of planetary wave forcing for the HT relationship in GCMs. Balachandran and Rind [1995] found that EP flux convergence in the high-latitude stratosphere was greater in QBO-E and the EP flux was directed less equatorward. Hamilton [1998] found that in QBO-E on the whole there is greater EP flux convergence in the polar stratosphere. Niwano and Takahashi [1998] observed in a 15-year GCM integration that there is greater upwards EP flux in QBO-E than in QBO-W at high-latitudes but less in mid-latitudes. Calvo *et al.* [2007] found more upwards EP flux in QBO-E in the months their model displayed an HT relationship, although EP flux convergence was only substantially greater in the upper stratosphere.

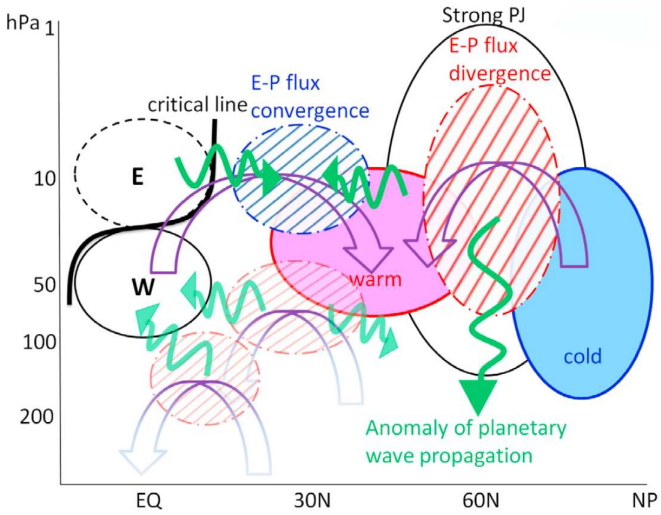

Figure 2.16: A schematic illustration of the EP flux and residual circulation differences between QBO-W and QBO-E from Yamashita *et al.* [2011] that is implied by their proposed mechanism for the HT relationship. Thin black lines represent the zonal wind (solid lines, westerly wind; broken lines, easterly wind), and the thick black line is the critical line. Red and blue areas represent warm and cold anomalies, respectively. Hatched red and blue areas represent anomalies of the divergence and convergence of the EP flux, respectively. Purple vectors denote the anomaly of the residual mean meridional circulation. Green arrows denote the anomaly in propagation of planetary waves. The faint colors in the low latitude lower stratosphere indicate that the effect on the polar vortex is unclear. © Wiley. Used with permission.

Naoe and Shibata [2010] showed the composite differences of lower stratospheric extratropical EP flux between QBO-W and QBO-E in ERA-40 and in a chemistry-climate model were not more equatorward, and argued that this contradicts the HT mechanism. This led them to argue that the QBO meridional circulation has an important role in the HT relationship and that the shift of the critical surface in the lower stratosphere is not important. Garfinkel *et al.* [2012b] reached a similar conclusion by examining the transient response to nudging equatorial winds towards QBO-E in a general circulation model (GCM) without coupled chemistry.

Anstey *et al.* [2010] suggested on the basis of GCM results that the QBO equatorial ZMZW anomalies in the middle stratosphere may be more important than the lower stratospheric anomalies for influencing the vortex. They noted that Polvani *et al.* [1995] found that increased planetary wave breaking led to a strengthened vortex

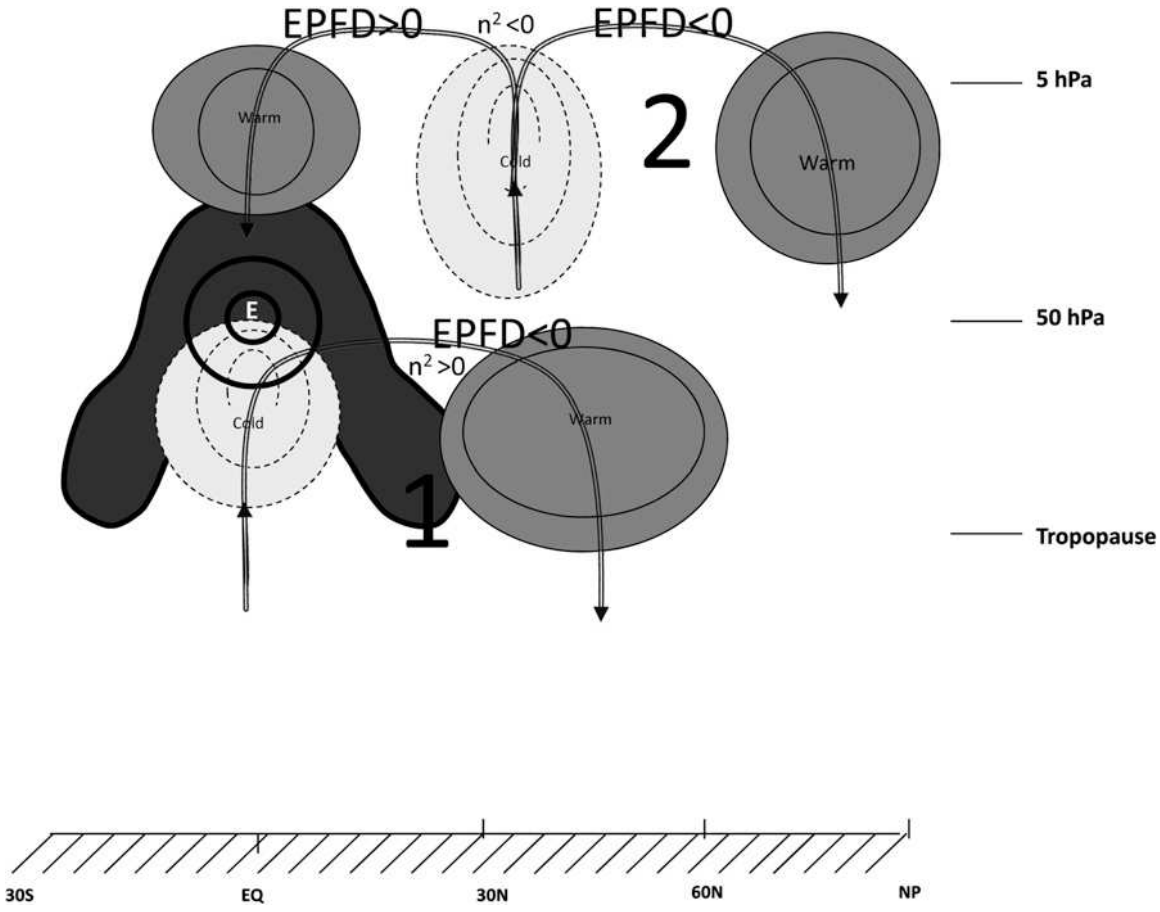

Figure 2.17: Schematic of the response of the extratropics to the QBO 16–30 days after applying QBO-E forcing, from Garfinkel *et al.* [2012b]. Arrows denote the mass-weighted meridional circulation. Thin contours are for temperature. All features, except for the easterly maxima at the equator (denoted with a thick contour), are a response to the QBO-E winds rather than being due directly to the externally imposed torque. Garfinkel *et al.* [2012b] claim that changes in EP flux divergence (EPFD) are directly related to changes in the index of refraction. In region 1 (subtropical lower stratosphere), the axisymmetric meridional circulation of the QBO in thermal wind balance dominates, although the poleward boundary of the easterlies and the latitude of the subsidence is modulated by eddies. That is, eddies propagating from the troposphere are influenced by the subtropical critical line and break in the lower stratosphere throughout the midlatitudes, forcing a TEM circulation that warms the midlatitude lower stratosphere. In region 2 (midlatitude upper stratosphere), subpolar Rossby waves are restricted from propagating into the subtropics because of a decrease in midlatitude index of refraction and therefore break closer to the pole. These Rossby waves lead to a TEM circulation that warms the pole. The warm anomaly reaches the lower stratosphere in the third and fourth month after branching. The effects in regions 1 and 2 are mechanistically distinct according to Garfinkel *et al.* [2012b].

in their shallow water model when radiative relaxation was included, which means it may be the case that during QBO-W it is the middle stratospheric equatorial easterlies that give rise to the stronger vortex through this mechanism. Yamashita *et al.* [2011] similarly proposed that the northwards critical surface shift in the middle stratosphere in QBO-W is more important than the southwards shift in the lower stratosphere, and argued that this seemed more consistent with composite differences of EP flux between QBO-W and QBO-E in their chemistry-climate model than the HT mechanism.

Figures 2.15, 2.16 and 2.17 show the predictions of the proposed mechanisms of Naoe and Shibata [2010], Yamashita *et al.* [2011] and Garfinkel *et al.* [2012b] respectively for the QBO influence on the EP flux, which can be compared to the prediction of the HT mechanism shown in figure 2.14. The suggestion of Anstey *et al.* [2010] would seem to predict a QBO influence on the EP flux similar to that of the HT mechanism, but shifted up to the middle stratosphere, adjacent to the equatorial easterlies in QBO-W and the equatorial westerlies in QBO-E. Most importantly, these alternatives to the HT mechanism do not indicate that any influence of tropical lower stratospheric ZMZW on the EP flux should directly affect the polar stratosphere.

However, experiments with primitive equation models have shown that the vortex is more disturbed when equatorial winds are relaxed towards a constant easterly value at all heights, which would not be expected to produce a strong meridional circulation due to the lack of vertical wind shear in the tropics [Gray *et al.*, 2003][3], or towards easterly jets [Naito *et al.*, 2003; Gray *et al.*, 2004]. O'Sullivan and Young [1992] studied the effect of adding equatorial stratospheric easterlies and westerlies to wintertime initial conditions on the subsequent flow in a primitive equation model. Temperatures were relaxed to a profile in balance with the equatorial anomalies, so that no QBO meridional circulation existed, and easterly winds were found to cause vortex weakening. Together these results suggest that neither the meridional

[3]Note that there may have been vertical wind shear at the base and top of the tropical stratosphere in these experiments, which could have contributed to driving a meridional circulation, though one probably unlike that associated with the QBO.

circulation nor a southwards shift of the middle stratospheric critical surface are necessary to produce the HT relationship. This raises the questions of whether GCMs behave differently to the primitive equation models, or whether the effects of shifts in the critical surface at different heights do not combine linearly.

It has also been proposed that changes in upper stratospheric winds associated with the QBO are important for producing the HT relationship. Gray *et al.* [2001b] found in rocketsonde data that the vortex strength is well correlated with upper stratospheric vertical wind shear near 1 hPa, although only a 26-year time series of winds in this region could be constructed and the statistical significance of the result was not high. Gray *et al.* [2001a] found in a stratosphere-mesosphere primitive equation model that it was necessary to include equatorial upper stratospheric winds along with the lower stratospheric winds to reproduce the HT relationship. Although many other studies have reproduced the HT relationship by imposing only the lower stratospheric QBO wind anomalies, this indicates that upper equatorial winds may play a role too. Gray [2003] demonstrated in the same model that equatorial lower stratospheric winds influenced the vortex in early winter and upper stratospheric winds had an influence in late winter. Matthes *et al.* [2004] found that equatorial upper stratospheric winds needed to be taken into account to reproduce the effect of the eleven-year solar cycle on the vortex in an atmospheric GCM.

Overall, regarding evidence for the different proposed mechanisms by which the QBO may affect the vortex, observational and modelling studies fail to provide a clear and consistent picture. The modelling studies indicate that on average there is greater upwards EP flux into the high-latitude stratosphere from the troposphere and greater EP flux convergence in the stratosphere during QBO-E than during QBO-W, although the locations of these effects differ between models. These results have been interpreted as being broadly consistent with the HT mechanism, although it has been argued that the effects at low latitudes do not conform to predictions. It will also be argued in section 4.2, however, that many possible mechanisms may be consistent with these results, and that it is likely to be more useful to examine the transient response of the vortex to an imposed QBO forcing on short time scales, perhaps as

short as several days. All the aforementioned studies examined either the transient response on much longer time scales or just differences in the mean state between the different QBO phases in long integrations.

No study has clearly been able to separate the influences of different aspects of the QBO structure, such as the meridional circulation or the lower and upper stratospheric winds. Nor has any study explained why there is not greater planetary wave activity in QBO-E in January and February.

2.7.3 Non-linearity in the combined influence of the QBO and other natural forcings on the vortex

There is some observational and modelling evidence that the QBO influence on the vortex interacts non-linearly with the influences of the eleven-year solar cycle and ENSO. Understanding these non-linearities, if they are real, could be helpful for predicting vortex variability.

Overall in observations, the vortex is stronger in solar maximum than in solar minimum [e.g. Kodera, 1995; Kodera and Kuroda, 2002; Lu *et al.*, 2009; Frame and Gray, 2010], although the effect of the solar cycle is not highly statistically significant. This is generally consistent with the results of modelling studies [e.g. Balachandran and Rind, 1995; Gray *et al.*, 2004; Ineson *et al.*, 2011; Cnossen *et al.*, 2011].

However, it has been suggested based on observational data that the solar cycle influence in February reverses in different QBO phases, with the vortex strengthening with increasing solar activity in QBO-E and weakening in QBO-W. Correspondingly, the vortex is weaker in QBO-E and stronger in QBO-W at solar minimum (the "normal" HT relationship) but the relationship reverses in solar maximum [Labitzke, 1987, 2005; Labitzke and Loon, 1988; Kodera, 1993; Gray, 2010]. Naito and Hirota [1997] found that the HT relationship is only statistically significant in solar minimum when considering January–March mean ZMZW, and Lu *et al.* [2008] found the HT relationship weakens but does not reverse in solar maximum when considering extratropical January–February mean ZMZW and ZMT. Camp and Tung [2007b] argued that the February–March mean vortex strength is weaker during either QBO-E

or solar maximum but the vortex is not weakened much more by the combination of both of these forcings than by either one alone, which is why the HT relationship seems weaker in solar maximum. However, no study has shown that differences between solar maximum and minimum of the HT relationship sign and magnitude are highly statistically significantly different from one another, and it does not seem clear that the substantial differences seen in February are unlikely to have arisen by chance. Baldwin and Dunkerton [1989] questioned the statistical significance on the grounds that replacing the solar cycle index by artificial harmonic time series can give similar results in the data that was available at that time.

Relevant modelling studies include that of Kodera [1991], who found there was an HT relationship only when the specified solar UV radiation was reduced by 20%. Balachandran and Rind [1995] also found the HT relationship was stronger in their model when UV input was reduced and Matthes *et al.* [2004] found the HT relationship is weaker in solar maximum in a GCM. Again, however, the statistical significance of the differences is not clear and it is unknown how many models do not show this effect since studies where the effect is not found may not end up being published[4]. No modelling studies have reported that the HT relationship reverses when solar input is high. Overall, therefore, there is an indication in the observations and some models that the HT relationship is weaker during periods of high solar activity, but the evidence is not yet convincing. Given that there is no clear theoretical reason to think there would be a strong non-linear interaction between the effects of the QBO and solar cycle on the vortex, it seems possible that none exists.

There is also some observational and modelling evidence that the QBO and ENSO influences on the vortex interact non-linearly. The vortex is weaker during El Niño than during La Niña in observations [e.g. Wallace and Chang, 1982; Baldwin and O'Sullivan, 1995; Camp and Tung, 2007a; Garfinkel and Hartmann, 2007, 2008; Mitchell *et al.*, 2011b] (although the results of the earliest studies were not robust

[4]This is also true for other statistical relationships in models, but seems particularly problematic in the case of the QBO-solar cycle interaction, due to the lack of statistical significance of this effect in observations and the lack of any theoretical explanation.

due to the short length of the observational record and the difficulty of separating out the effects of other forcings [Hamilton, 1993a]). This relationship has also been found in most modelling studies [e.g. Hamilton, 1993b; Sassi *et al.*, 2004; Garcia-Herrera *et al.*, 2006; Manzini *et al.*, 2006; Taguchi and Hartmann, 2006] (an exception is that of Lahoz [2000]). These studies indicate that the GPH zonal wavenumber-1 amplitude is greater in El Niño, due to enhancement of the Pacific-North American pattern, and this may explain the weakening of the vortex as this would be expected to cause greater upward EP flux into the high-latitude stratosphere.

Garfinkel and Hartmann [2007] found in observations that the influence of ENSO is diminished during QBO-E (as did Free and Seidel [2009]) and that of the QBO is diminished during El Niño, and Wei *et al.* [2007] found that the HT relationship is only statistically significant during La Niña. Garfinkel and Hartmann [2010] suggested that the QBO may alter the tropospheric ENSO teleconnection patterns and hence its influence on the vortex, although a portion of the observed effect could be due to sampling variability. Calvo *et al.* [2009] found in an atmospheric GCM that during the QBO-E phase, the warming of the vortex by El Niño is larger during December and January than during QBO-W, and that during El Niño periods, the warming of the vortex by QBO-E is greater and propagates downwards more quickly than during La Niña periods. Richter *et al.* [2011] found in a chemistry climate model that the SSW frequency is similar when either or both of QBO and ENSO variability are represented, but decreases considerably when neither is present. In addition, Hurwitz *et al.* [2011] found in observations that "warm pool" El Niño events warm the SH vortex in November–December only during QBO-E, suggesting this non-linearity is present in both hemispheres.

This non-linear interaction between the QBO and ENSO influences on the vortex seems consistent with the modelling results that show the vortex response to the QBO phase is dependent on the tropospheric wave forcing (section 2.7.1). During El Niño the vortex may move more into a high-forcing regime which makes the vortex less sensitive to the influence of the QBO, with the opposite occurring in La Niña, though this is speculative. However, the statistical significance of the difference in

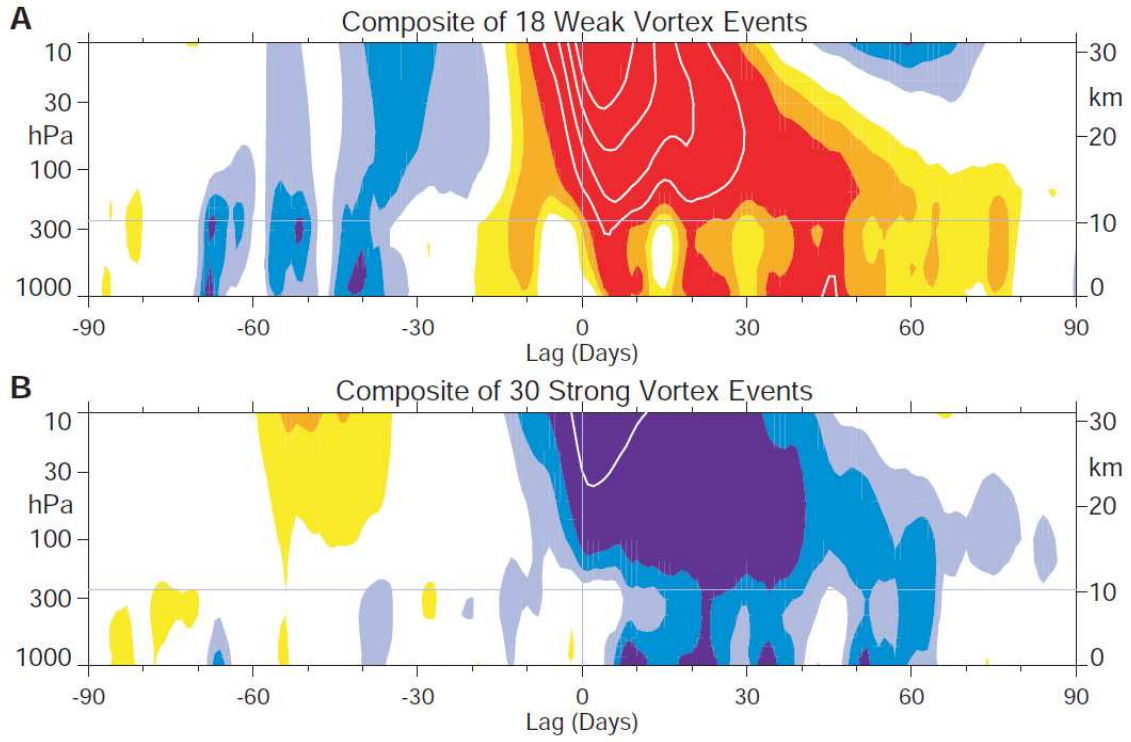

Figure 2.18: Composites of time-height development of the northern annular mode for (A) 18 weak vortex events and (B) 30 strong vortex events. The events are determined by the dates on which the 10 hPa annular mode values cross −3.0 and +1.5, respectively. The indices are nondimensional; the contour interval for the color shading is 0.25, and 0.5 for the white contours. Values between −0.25 and 0.25 are unshaded. The thin horizontal lines indicate the approximate boundary between the troposphere and the stratosphere. From Baldwin and Dunkerton [2001]. © The American Association for the Advancement of Science. Used with permission.

the observed HT relationship strength in different ENSO phases and the robustness of the interaction in models is again not clear at this time.

2.8 Stratospheric influence on the extratropical troposphere

In recent years evidence has accumulated indicating that vortex variability may affect the troposphere, which provides additional motivation for understanding the HT relationship. Baldwin and Dunkerton [1999] and Baldwin and Dunkerton [2001] found that weakenings of the vortex are associated with the tropospheric NAM index be-

coming more negative subsequently on average for a period of two to three months (figure 2.18), much longer than the usual time scales of tropospheric variability but comparable to the time taken for the vortex to recover its strength following an SSW. This relationship is reproduced by general circulation models (GCMs) [Charlton-Perez *et al.*, 2013]. Other examples of studies that have shown a link between stratospheric and tropospheric variability include Thompson *et al.* [2002], who showed that weakenings of the vortex are associated with extreme cold events in eastern North America, northern Europe and eastern Asia, and Scaife and Knight [2008] who found that the cold European weather of January 2006 was predicted better in a model with a better stratospheric resolution. Kolstad *et al.* [2010] and Thompson *et al.* [2010] also concluded that the evolution of SSWs is connected with cold air outbreaks in the NH and Sigmond *et al.* [2013] found that seasonal prediction of the NAM index is more skilful shortly after the onset of SSWs. Hitchcock *et al.* [2013] reported that SSWs followed by particularly persistent lower stratospheric anomalies are associated with stronger tropospheric anomalies.

The observed association between stratospheric and tropospheric anomalies is not proof that the stratosphere causes changes in the troposphere, but a number of modelling studies have shown that perturbing the stratosphere alone leads to changes in the troposphere. For example Polvani and Kushner [2002] found that cooling the high-latitude stratosphere in a primitive equation model led to the NH tropospheric jet shifting polewards and strengthening at the surface and Norton [2003] and Simpson *et al.* [2011] found that damping stratospheric variability reduced the time scale of the tropospheric annular modes. Charlton *et al.* [2004] showed that changing stratospheric initial conditions in a numerical weather-prediction model affects the Arctic oscillation at the surface and Jung and Barkmeijer [2006] found that applying a forcing to bring about a weakening of the vortex results in a negative North Atlantic Oscillation response.

Stratospheric variability may be important for the way the northern extratropical troposphere responds to external forcings. Anstey and Shepherd [2013] show that there is a marginally statistically significant difference in the NAM signature

between QBO-W and QBO-E in observations corresponding to a difference of about one standard deviation in the NAM index, which may be due to the HT relationship. This is also seen in the modelling results of Marshall and Scaife [2009]. Ineson and Scaife [2009] found in a GCM that the wintertime ENSO teleconnection to Europe is only present in winters with SSWs, the frequency of which may also be affected by ENSO (section 2.7.3). Cagnazzo and Manzini [2009] found that the teleconnection existed in a model with a well-resolved stratosphere and not in an equivalent model without one, suggesting stratospheric variability is related to this teleconnection and Bell *et al.* [2009] found that the ENSO teleconnection to Europe is modified if stratospheric variability is suppressed. Ineson *et al.* [2011] argue their GCM simulations indicate that maxima in the eleven-year solar cycle cause the polar vortex to be stronger which leads to a more positive NAM index.

Various mechanisms have been proposed to explain the stratospheric influence on the troposphere. For example it has been proposed that stratospheric PV anomalies would be expected to be related to tropospheric anomalies through the PV inversion relationship [e.g. Hartley *et al.*, 1998; Ambaum and Hoskins, 2002], that vortex disturbances cause planetary waves to be reflected down to the troposphere [e.g. Perlwitz and Harnik, 2003] or that changes in stratospheric EP flux convergence induce changes in the residual meridional circulation in the troposphere [Haynes *et al.*, 1991; Shepherd, 2002]. Additionally, eddy feedbacks within the troposphere are likely to be important for setting the magnitude and structure of the tropospheric response [e.g. Kushner and Polvani, 2004; Song and Robinson, 2004].

Thus understanding NH stratospheric variability may improve seasonal forecasts of NH weather [Baldwin *et al.*, 2003a; Charlton *et al.*, 2003; Shaw and Shepherd, 2008; Marshall and Scaife, 2009; Thompson *et al.*, 2010; Maycock *et al.*, 2011].

Chapter 3

Data, models and statistical methods

3.1 The ERA-40 reanalysis

The ECMWF Re-Analysis (ERA-40; Uppala *et al.* [2005]) is used in this thesis. ERA-40 is a re-analysis produced using 3D-Var data assimilation to output the most likely state of the atmosphere at each 6 hour interval from September 1957 to August 2002 up to 1 hPa. The assimilating model has 24 levels above 100 hPa.

Most of the observational data for the stratosphere is from radiosondes and satellites. In 2004 about 800 radiosonde soundings per day were reaching 30hPa and 350 per day reached 10hPa. Most radiosonde coverage is over the NH landmasses. Satellites have provided near-global coverage since 1979, although their vertical resolution in the stratosphere is limited to about 10km [Randel *et al.*, 2004].

Randel *et al.* [2004] found that ERA-40 matches measurements of the zonal mean circulation derived from radiosonde, rocketsonde and lidar measurements quite closely and performs quite well compared to other analyses and re-analyses. Errors in the upper stratosphere are larger, with ERA-40 being biased cold by up to 5K between 2–5hPa, although Uppala *et al.* [2005] report they are biased warm in the early part of the reanalysis. ERA-40 also displays an oscillatory vertical structure in temperature near the poles during winter, which is especially pronounced over Antarctica. Interannual variability of the ZMZW at 60N in ERA-40 agrees closely with that in

other analyses and re-analyses ["SPARC intercomparison of middle-atmosphere climatologies" report, 2002]. ERA-40 agrees well with the NCEP/NCAR re-analysis in its representation of the polar vortices [Karpetchko *et al.*, 2005] and SSWs [Charlton and Polvani, 2007]. The occurrence of SSWs also agrees with studies of individual SSWs [Charlton and Polvani, 2007]. The ERA-40 estimate of the QBO amplitude is realistic and falls between that of rocketsonde and radiosonde measurements [Randel *et al.*, 2004; Baldwin and Gray, 2005].

There is not much literature which compares the zonally asymmetric component of the flow structure in ERA-40 to that in raw data and other analyses. The differences in zonal asymmetries in ERA-40 between QBO-E and QBO-W are examined in chapter 4 where it is explained that the results are in agreement with those obtained using other atmospheric datasets.

Overall ERA-40 seems to represent large-scale features in the NH stratosphere well, although it is probably not reliable above 10 hPa prior to 1979 due to the small amount of data available at these levels before the satellite era.

3.2 The ERA-Interim reanalysis

In chapter 8, the ERA-Interim reanalysis data [Simmons *et al.*, 2007a,b; Dee *et al.*, 2011] is used to produce target equatorial zonal wind profiles to which the Met Office Unified Model (section 3.4) is nudged. ERA-Interim data between January 1979 and December 2010 are used. ERA-Interim improves upon ERA-40 by using 4D-Var data assimilation in a model with greater resolution and improved model physics. It is markedly more consistent with direct observations over Antarctica, where the oscillatory structure in temperature is less pronounced. In the tropical stratosphere, the strength of the upwelling is in better agreement with observations, whilst in ERA-40 it is too strong. The upper stratospheric cold bias is reduced compared to that in ERA-40. ERA-Interim is likely to have a better representation of the stratosphere above 10 hPa than ERA-40, so it is used in preference to ERA-40 to create target equatorial zonal wind profiles.

3.3 The UKMO Stratosphere-Mesosphere Model (SMM)

In chapters 5, 6 and 7 of this thesis, the SMM is used to study qualitative aspects of how the stratosphere responds to an applied torque. The SMM is a global primitive equation model of the stratosphere and mesosphere, described in detail by Fisher [1987]. In the configuration used here it has 5° resolution in latitude and longitude and 2 km resolution in height over the domain 16–80 km. The height coordinate is log-pressure with scale height 6.95 km, so this corresponds to the pressure domain 0.01–100 hPa, with every 16 km increase in altitude corresponding to a factor 10 decrease in pressure. The model integrates the primitive equations with prognostic variables held on an Arakawa A-grid, using finite difference schemes for spatial derivatives and a leapfrog time-stepping scheme with an Asselin time filter and a time step of 4 minutes. In order to ensure stability, truncation of zonal Fourier modes is applied at or poleward of latitudes 72.5N nd 72.5S and a Shapiro filter is applied to remove waves with wavelengths smaller than three grid boxes, with complete removal of waves with wavelengths smaller than two grid boxes. The vertical velocity is set to zero half a grid space above the top level. At the bottom boundary the GPH is specified and enters as a forcing term in the zonal and meridional momentum equations.

The radiative contribution to the diabatic heating rate is computed by the MIDRAD scheme [Shine, 1987], which calculates absorption of solar radiation from above and of radiation reflected from the troposphere below by ozone and diatomic oxygen, assuming a climatological ozone distribution based on satellite and balloon data collected in the early 1980s, and absorption of thermal radiation by the 15 μm band of carbon dioxide, the 9.6 μm band of ozone and the vibration-rotation and rotation bands of water vapour, with carbon dioxide and water vapour assumed to be well mixed. This has the advantage over relaxing towards a prescribed temperature field, as has been done in many previous modelling studies of stratospheric dynamics, that temperature variability is damped in a more realistic way.

Gravity wave drag is parameterised simply by Rayleigh friction terms $-u/\tau(z)$ and $-v/\tau(z)$ in the zonal and meridional momentum equations respectively. $\tau(z)$ is 10^7s $\approx$ 116 days below 50 km, and decreases towards 1.4 days at 80 km. The exact equation above 50 km is

$$\tau(z) = (10^{-7} + 1.5 \times 10^{-5}[1 - \exp\{(50 - z/\mathrm{km})/40\}])^{-1}\,\mathrm{s},$$

where this friction also acts to create a "sponge layer" which reduces spurious wave reflection from the model top.

The SMM has been shown to capture stratospheric variability reasonably well when initiated in a realistic state and forced at the lower boundary with observed 100 hPa GPH. For example, Butchart *et al.* [1982] showed that an older model version with lower horizontal resolution could simulate the 1979 vortex splitting event reasonably well, and O'Neill and Pope [1993] found that the model is able to simulate well the 1980 vortex displacement event. The model has also been used in many studies to examine stratospheric dynamics [e.g. Fairlie *et al.*, 1990; Fisher *et al.*, 1993; Scaife and James, 2000; Gray *et al.*, 2004]. The model's simulation of the mesosphere is less well validated, but it is its simulation of the stratosphere which is the main focus in chapters 5, 6 and 7. The model's resolution is sufficient for simulating large scale planetary waves but it is not able to simulate accurately the response of gravity wave drag to changes in the large scale circulation. Scott and Polvani [2006] showed that the internal stratospheric variability of a primitive equation model starting at a similar resolution to that of the SMM does not change much qualitatively as horizontal resolution is increased by a factor of 4 in each direction.

The Rayleigh friction does not conserve zonal momentum and this can give rise to physically unrealistic downward influence of processes internal to the upper atmosphere and affect the response to applied forcings [Shepherd *et al.*, 1996; Shepherd and Shaw, 2004]. The model version with Rayleigh friction applied as above is used since this version has been shown to represent stratospheric variability reasonably well. It is the qualitative nature of how the extratropical stratosphere responds to an

applied torque that is of interest in chapters 5, 6 and 7 and it is shown that this is not strongly affected by the presence of Rayleigh friction in appendices 5.A and 6.A.

The details of the model runs used in this thesis and comparisons of simulated and observed stratospheric variability are presented in chapter 5.

3.4 The UKMO Unified Model (HadGEM2-CCS)

Analysis of simulations by the UKMO HadGEM2-CCS GCM is presented in chapter 8 and output from this model is used as a bottom boundary condition of the SMM in chapters 5, 6 and 7. HadGEM2-CCS is a coupled ocean-troposphere-stratosphere model with 60 atmospheric levels in the vertical up to 84 km altitude (corresponding to a pressure of approximately 0.01 hPa), 32 levels above 16 km (corresponding approximately to a pressure of 100 hPa) and atmospheric horizontal resolution 1.25° in latitude and 1.875° in longitude. The oceanic resolution is 1° polewards of 30° latitude and increases smoothly to 0.33° at the Equator. The model includes parameterised orographic gravity wave drag up to 40 km height, using the scheme of Webster *et al.* [2003], and parameterised non-orographic gravity wave drag (NOGWD), using the scheme of Warner and McIntyre [1999] as implemented by Scaife *et al.* [2002]. The NOGWD causes the model to exhibit a spontaneous QBO. The radiation scheme is described by Edwards *et al.* [2004] and the model does not include a chemistry scheme, apart from parameterised methane oxidation, although it includes a coupled carbon cycle. For full model details see Martin *et al.* [2011].

Data are presented from a 240-year control run of HadGEM2-CCS[1], which has a constant CO_2 mass mixing ratio of 4.35×10^{-4}, prescribed zonally symmetric monthly mean climatological ozone which is the 1850–1860 mean seasonal cycle of the dataset of Cionni *et al.* [2011], volcanic aerosol optical depth 0.0097 and solar constant 1365 W/m^2. Data on 20 pressure levels at or above 100 hPa are used.

Osprey *et al.* [2013] found that HadGEM2-CCS exhibits a realistic stratospheric climatology and realistic variability. The DJF mean ZMZW is similar to that in

[1]The control run was performed by Steven Hardiman at the UKMO.

ERA-40. The model has a warm bias in the tropical lower stratosphere, a cold bias in the tropical upper stratosphere of 2–6 K and a warm bias of up to 4 K in the DJF high latitude NH stratosphere. The (60N, 10 hPa) ZMZW shows a realistic seasonal cycle and variability through the northern winter [Hardiman *et al.*, 2012]. The area-integrated PV of the vortex is smaller than in ERA-40, but the vortex has a realistic aspect ratio and centroid latitude [Mitchell *et al.*, 2012].

The overall SSW frequency is consistent with that in ERA-40, although SSWs occur less frequently in November–January and more frequently in February–March in the model [Osprey *et al.*, 2013]. This may be because the tropospheric planetary wave forcing is slightly weaker than that in ERA-40 (section 3.4.2). The distribution in time and height of model final warming dates is similar to that in observations. The model produces a realistic "tropical tape recorder" of equatorial stratospheric water vapour concentrations and realistic stratospheric "age of air", indicating that the model Brewer-Dobson circulation and stratospheric extratropical variability are realistic [Osprey *et al.*, 2013].

3.4.1 The QBO in HadGEM2-CCS

Figure 3.1(a) shows the 10S–10N ZMZW as a function of height and pressure in HadGEM2-CCS. The modelled QBO closely resembles that in observations (figure 2.9), with comparable maximum easterly and westerly ZMZW. The mean QBO period (defined by the time between successive westerly to easterly transitions at 50 hPa with a five month running mean applied to the ZMZW time series) in the model is 26.4 months, which is close to the ERA-40 mean period of 28.0 months. The modelled QBO is more regular than that in ERA-40, as can be seen by comparing figures 2.9 and 3.1(a), and also from the fact that the standard deviation of the model QBO period is 1.8 months, compared to 3.9 months in ERA-40.

Figures 3.1(b) and (c) show the modelled QBO amplitude as a function of height and latitude, which can be compared to the similar analysis of Pascoe *et al.* [2005] of the QBO in ERA-40. They defined the QBO component of equatorial winds to be that associated with the Fourier harmonics of the ZMZW with periods between 22–40

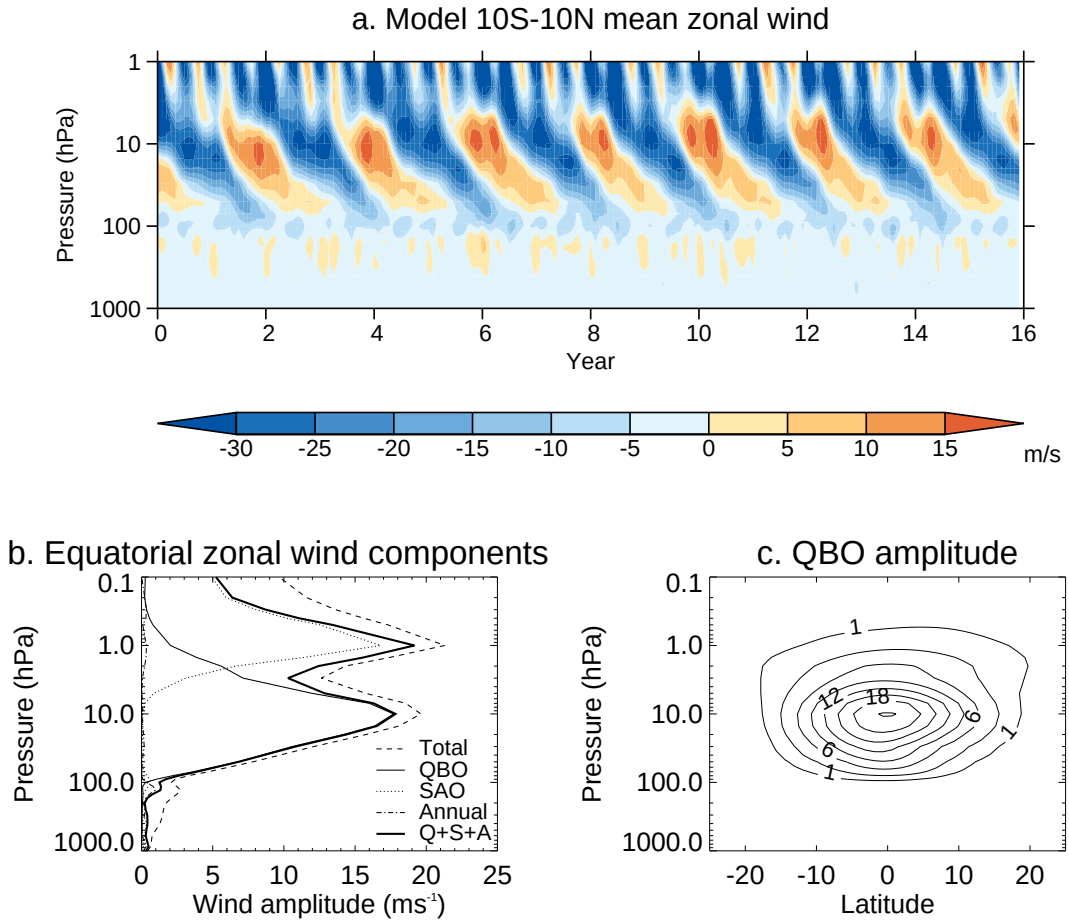

Figure 3.1: The QBO in HadGEM2-CCS. (a) shows the monthly, 10S–10N and zonal mean zonal wind as a function of time and pressure. (b) shows the amplitude of the QBO, SAO and annual cycle components of the 10S–10N and zonal mean zonal wind (see text). (c) shows the amplitude of the QBO as a function of latitude and pressure in ms^{-1}.

months, which dominate in the equatorial lower stratosphere. For the model data, all harmonics of 10S–10N ZMZW in the spectral band centred near the mean period with amplitudes greater than 1m around the 10 hPa level were included, which corresponds approximately to the amplitude cut-off used by Pascoe *et al.* [2005], which included periods in the range 22–34 months. Other important components of the equatorial winds are the SAO and annual cycle, which are taken to correspond to the Fourier components with frequencies of exactly six and twelve months respectively (although

note that since these oscillations are not pure sine waves, they will contribute some power to other frequencies as well).

Figure 3.1(b) shows the amplitude of all three components as a function of pressure. The model QBO amplitude peaks at approximately 18 ms^{-1} at 10 hPa, which is slightly higher than the peak amplitude in ERA-40, which is approximately 16 ms^{-1} at 15 hPa. The model QBO has a slightly greater amplitude at the uppermost levels than ERA-40 and a similar amplitude in the lower stratosphere, penetrating down to about 100 hPa. The model SAO also has a larger peak than that in ERA-40, at about 17 ms^{-1} compared to 11 ms^{-1} in ERA-40 at 1 hPa, although it falls away with height more quickly so that both the model and ERA-40 SAO amplitudes are zero at 10 hPa and 5 ms^{-1} at 0.1 hPa. ERA-40 shows an annual cycle signal on levels 1–5 hPa peaking at about 4 ms^{-1} and on levels 100–500 hPa peaking at about 3 ms^{-1} which is not captured by the model.

Figure 3.1(c) shows the amplitude of the model QBO as a function of latitude and height. The amplitude peaks at 21 ms^{-1} at 10 hPa at the Equator and decreases with latitude to be about 1 ms^{-1} at 20S and 20N, being fairly symmetric about the Equator. The corresponding figure for ERA-40 [Pascoe *et al.*, 2005, fig. 4a] appears similar, but with a smaller peak amplitude of about 15 ms^{-1} and a smaller amplitude at higher levels.

Overall, therefore, the HadGEM2-CCS QBO has a realistic meridional structure and mean period, with a slightly greater amplitude in the middle and upper stratosphere and a more regular descent than in ERA-40.

3.4.2 HadGEM2-CCS representation of planetary waves

For the model to be appropriate for examining the mechanism behind the HT relationship, it is necessary that it simulates planetary wave activity realistically. Figure 3.2(a) shows the 50–80N area-weighted meridional and temporal mean of the DJF GPH, with zonal mean subtracted to show the eddy component, in the model and in ERA-40. 50–80N is the approximate latitude range where planetary wave amplitudes are largest. There is very close agreement between the model and ERA-40,

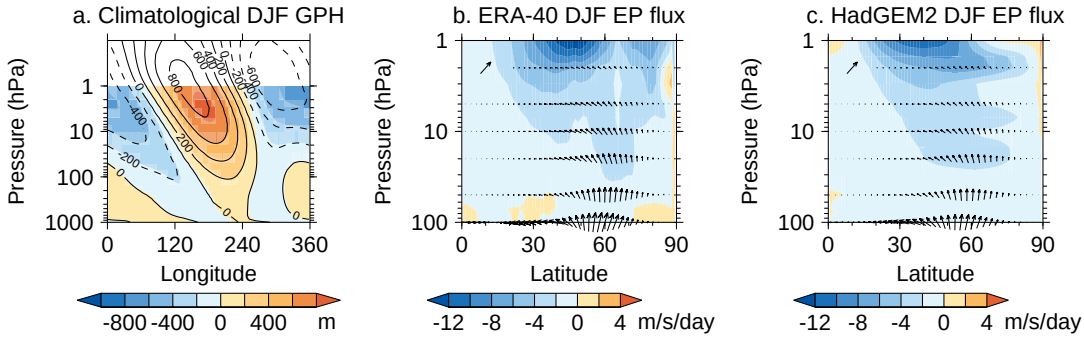

Figure 3.2: (a) shows the 50–80N area-weighted meridional and temporal mean of the DJF geopotential height, with zonal mean subtracted to show the eddy component, in HadGEM2-CCS (contours) and in ERA-40 (colours), with contours every 200m. (b) shows the ERA-40 DJF mean stratospheric EP flux (arrows) and D_F (colours), and (c) shows the same for the model. A reference arrow in the top left corner has a magnitude of $(10^7, 10^5)$ kg/s^2. The model data closely resemble the reanalysis.

particularly below 10 hPa where the quality of observations is better. Figures 3.2(b) and 3.2(c) show the DJF mean stratospheric EP flux and D_F (section 2.2.2) in the model and in ERA-40 respectively. The ERA-40 flux and D_F were calculated from 6-hourly wind and temperature data according to the equations of Andrews *et al.* [1987][2]. The model shows good qualitative agreement with ERA-40, exhibiting convergent upwards flux throughout most of the NH stratosphere, which is polewards below about 50 hPa and equatorwards above. Fluxes in the model are somewhat weaker, however – the upwards component is weaker by ∼20%.

3.5 Statistical methods

In general in this thesis statistical significance of the data is calculated according to "non-parametric" methods which only assume that the data are representative of the population they are drawn from. These methods are more robust than "parametric" methods which require making assumptions about the shape of the population distribution, such as that it is normal [Efron and Tibshirani, 1993]. The methods used

[2]The physical constants used are the same as those used by the Alfred-Wegener-Institut, shown at http://www.awi.de/en/research/research_divisions/climate_science/atmospheric_circulations/projects/candidoz/ep_flux_data/.

for estimating the probability of obtaining a value for the measured quantities with a magnitude at least as large as that observed under a given null hypothesis (the "p-value") are described in detail in the sections where results are presented. The p-value indicates only the probability of obtaining the measured signal under the null hypothesis and does not on its own specify whether there is likely to be a true relationship, since this also depends on prior information [Ambaum, 2010] – theoretical reasoning and combinations of different data sources can also be used to assess if the measured signal indicates a true physical process. All significance tests in this thesis are two-tailed.

Chapter 4

The Holton-Tan relationship on monthly time scales

The purpose of this chapter is to examine the HT relationship in ERA-40 (section 3.1) on monthly time scales, paying particular attention to the influence of the QBO on the zonally asymmetric structure in the extratropics, which has been used in some previous studies to make inferences about the mechanism behind the QBO's influence. It is argued that the mean differences between the high-latitude stratosphere in QBO-E and QBO-W are very like the NAM structure. Studies of the simple system of Lorenz [1963] and of tropospheric models indicate that many different forcings can give long-term mean responses that are very like the systems' leading EOFs. This is analagous to the extratropical influence of the QBO presented here being NAM-like and indicates that QBO-E minus QBO-W composite differences may not contain much information about the mechanism behind the HT relationship, and could be consistent with many different hypotheses for the mechanism. It is argued that the short-term transient response to an applied forcing like the QBO can be expected to be much more informative about the forcing mechanism. This chapter includes and expands upon part of the analysis of Watson and Gray [2014].

4.1 Observed HT relationship

Figure 4.1 shows the QBO-E minus QBO-W composite difference in the monthly mean NH EP flux and D_F (section 2.2.2) from November to February in ERA-40. The

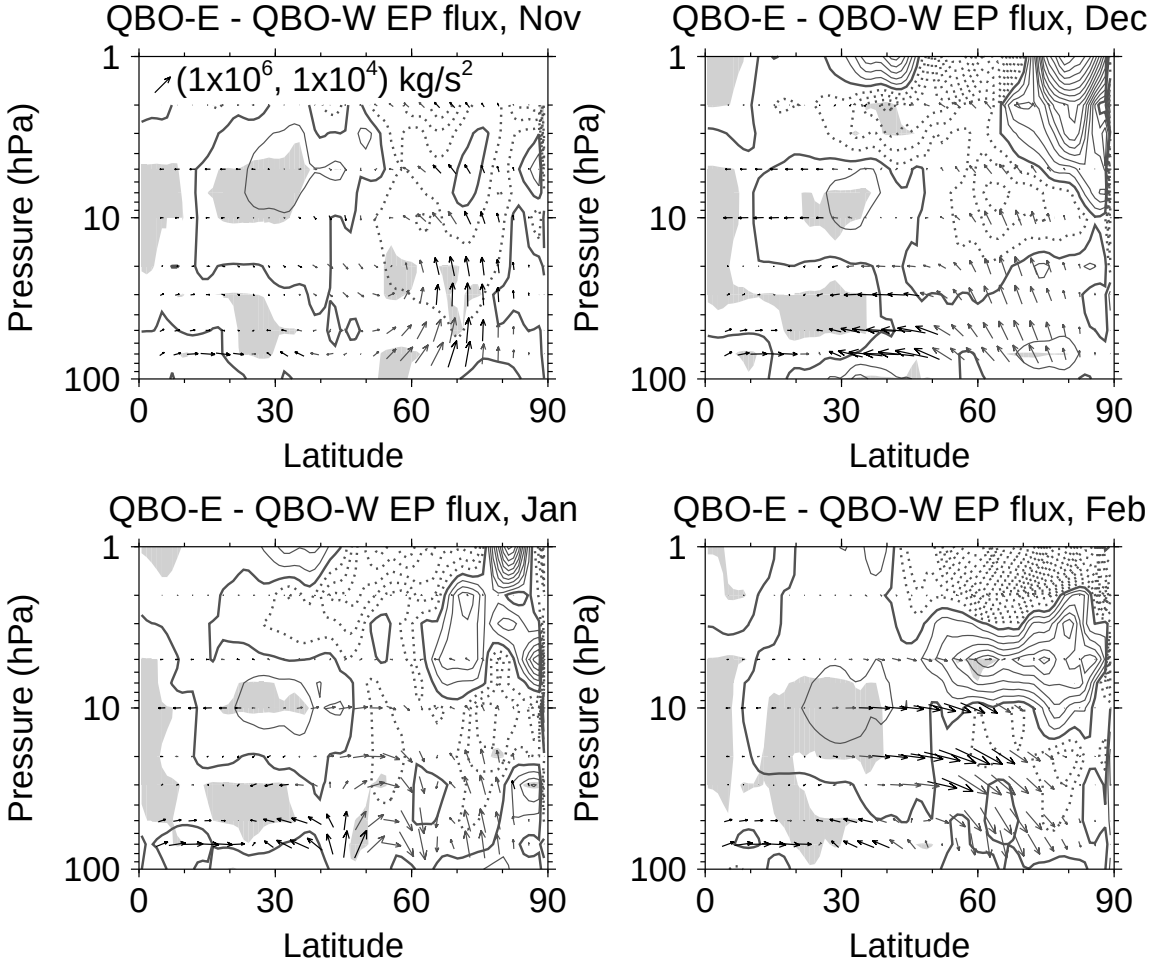

Figure 4.1: QBO-E minus QBO-W composite differences in ERA-40 of the winter-time EP flux (arrows, black where either the F^{ϕ} or F^{z} differences are statistically significant above the 95% level and grey otherwise) and D_F (contours, plotted at 0.5 m/s/day intervals, with negative contours dotted and a thickened zero contour). The EP flux is shown at pressures 2, 5, 10, 20, 30, 50 and 70 hPa and every 3.75° in latitude. A reference arrow is shown in the top left plot along with its (F^{ϕ}, F^{z}) values. Shading shows where D_F differences are statistically significant above the 95% level.

QBO is defined as being in its easterly or westerly phase when the 5S–5N November–February mean ZMZW is easterly or westerly respectively at 50 hPa, the pressure at which the correlation between the November–February mean ZMZW averaged over 5S–5N and that at (60N, 10 hPa) is greatest, at 0.58. In agreement with the findings of previous studies [e.g. Dunkerton and Baldwin, 1991; Ruzmaikin *et al.*, 2005], the EP flux is more upward in November and December north of 55N, but this signal is weak in January, when the 30–90N mean upward EP flux difference at 70 hPa is positive but not statistically significant, and the EP flux is more downward at high latitudes in February. In all months there is a poleward EP flux difference in the tropical lower stratosphere, indicating the equatorward flux is less in QBO-E. This is restricted to latitudes south of about 25N with increased equatorward flux to the north, and Naoe and Shibata [2010] and Yamashita *et al.* [2011] argued that this means reflection of eddy zonal momentum flux from the lower stratospheric easterlies in QBO-E cannot be directly influencing the vortex.

Statistical significances of the QBO-E minus QBO-W composite differences were calculated according to a Monte Carlo (MC) permutation test. Each year was assigned to a surrogate QBO-E or QBO-W group at random, and the composite difference between these random groups was calculated. This was repeated 1,000 times to find the probability that the magnitude of the difference would exceed that of the difference in the data at each grid-point under the null hypothesis that there is no dependence of these variables on the QBO. The same test is used to find the statistical significance of composite GPH differences below.

In order to compare the differences in EP flux and other quantities with the NAM, the NAM is first indexed by the leading principal component (PC) of the monthly-mean 3D GPH north of 20N between 1–100 hPa, calculated using the method of Baldwin *et al.* [2009]. Prior to performing the decomposition into PCs, the GPH was multiplied by a mass-weighting matrix $\mathbf{W}$, such that $\mathbf{x}_{\mathrm{GPH}} \rightarrow W^{1/2}\mathbf{x}_{\mathrm{GPH}}$, where $\mathbf{x}_{\mathrm{GPH}}$ is the 1D vector containing the GPH values at all grid points. $\mathbf{W}$ is diagonal and has, for each grid point, an element equal to the product of cosine of the latitude and the integral of pressure over the interval between mid-points between the grid point

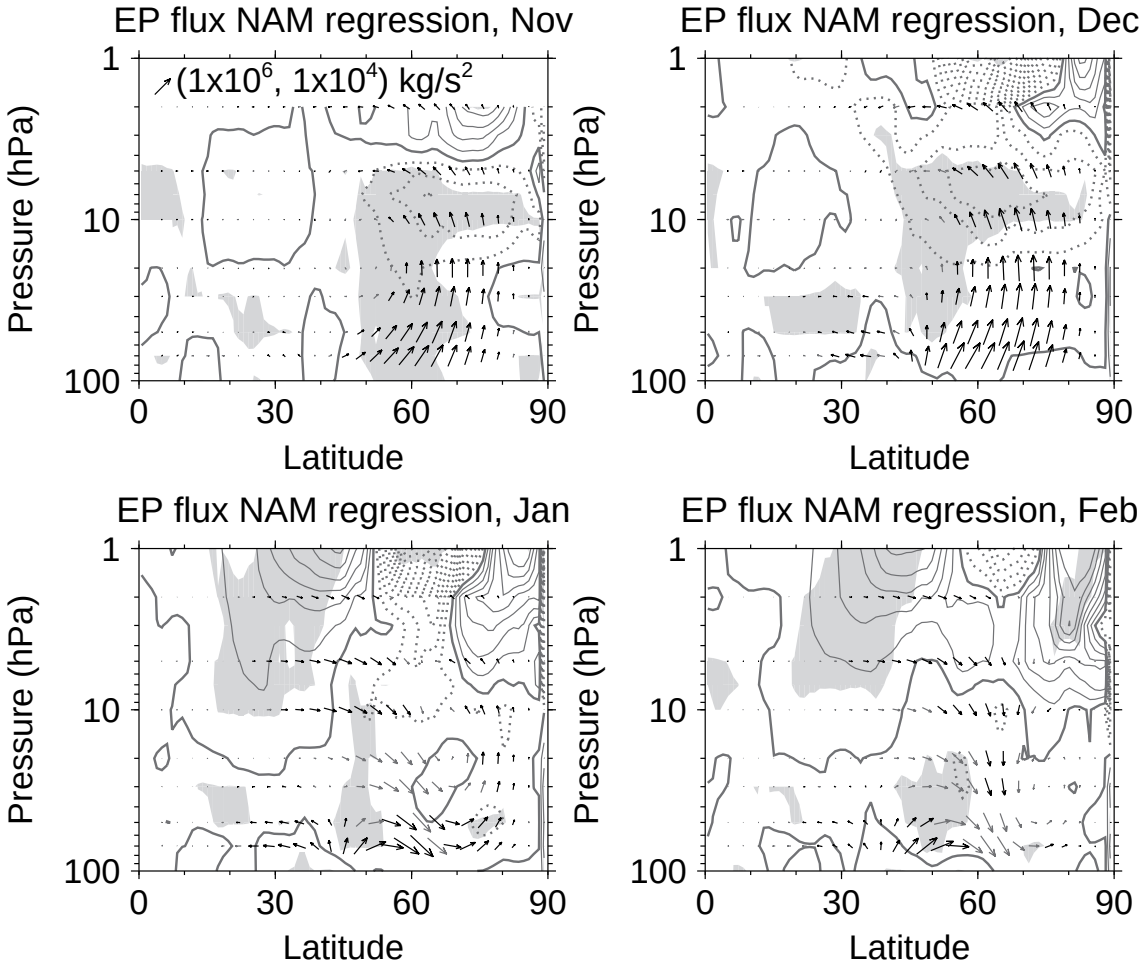

Figure 4.2: Regression of the EP flux and D_F onto the NAM index in ERA-40, showing the anomaly associated with a weaker vortex with my choice of sign of the index, plotted as for the QBO-E minus QBO-W composite differences in figure 4.1. There is a good correspondence between these regression patterns and the composite differences.

and neighbouring grid points on the pressure axis. This is similar to the NAM index of Thompson and Wallace [2000], but restricted to the stratosphere. The usual sign convention is reversed, so that when the index is positive the vortex is weaker, to more easily compare the NAM signature with the QBO-E minus QBO-W differences. Then the "NAM signature" of a given quantity is found by linearly regressing the time series of monthly mean values of that quantity against the index separately at each grid point. The presented signatures then correspond to a one standard deviation increase in the NAM index. This is done separately for each calendar month. The correlation between this index and the leading principal components of GPH on individual pressure levels 10 hPa and 50 hPa is 0.97 or greater in each calendar month November–February and the correlation with that at 5 hPa is 0.75 or above, so this index captures variability throughout the NH stratosphere well.

Figure 4.2 shows the regression of the EP flux against the NAM index, showing the anomaly associated with a weaker vortex. There is a good resemblance between the main qualitative features of this and the QBO-E minus QBO-W composite differences (figure 4.1) in each calendar month in the extratropics. The flux is more upward north of 55N in November and December, and near 45N and 80N in January. It is more downward near 60N in January and February (though the signal to noise ratio in January and February is not very high). The NAM regression does not show a poleward EP flux difference in the tropics, indicating that this feature in the differences is not strongly associated with extratropical variability.

Statistical significances of the regression coefficients of EP flux onto the NAM index were also calculated according to an MC permutation test. The time series of the NAM index for each calendar month was shuffled, and the regression coefficients against this random time series were calculated. This was repeated 1,000 times to find the probability at each grid-point that the magnitude of the regression coefficient would exceed that of the coefficient in the data under the null hypothesis that the variables have zero correlation with the NAM index. The same test is used to find the statistical significance of the regression coefficients of GPH below.

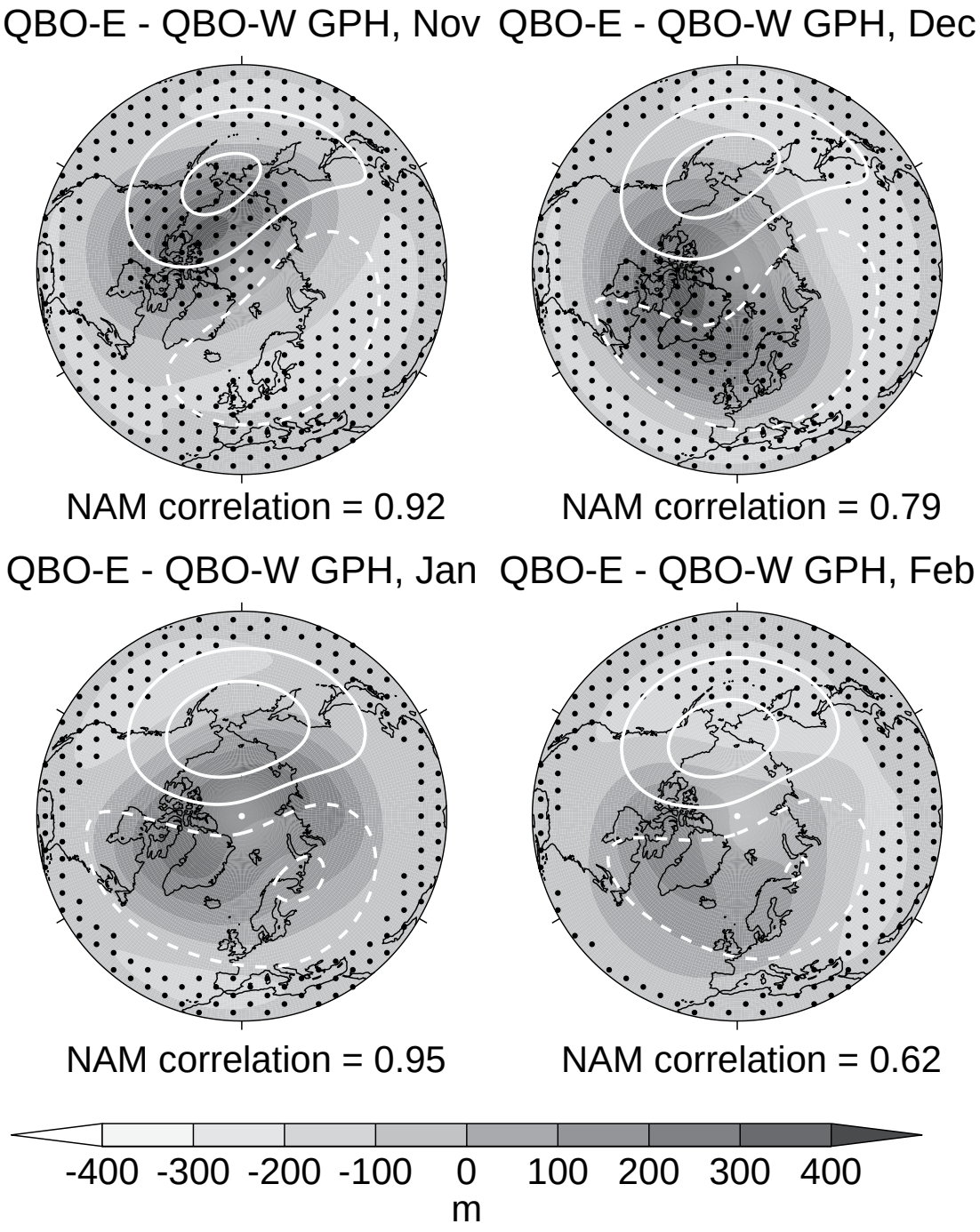

Figure 4.3: QBO-E minus QBO-W composite differences in ERA-40 of the wintertime GPH at 10 hPa north of 20N (greyscale). White contours show the climatological zonally asymmetric component of GPH with contour values ±200 and ±600 m with negative contours dashed. NAM correlation values indicate the anomaly correlation of the composite differences with the NAM signature north of 20N shown in figure 4.4. Stippling shows where GPH differences are statistically significant above the 95% level.

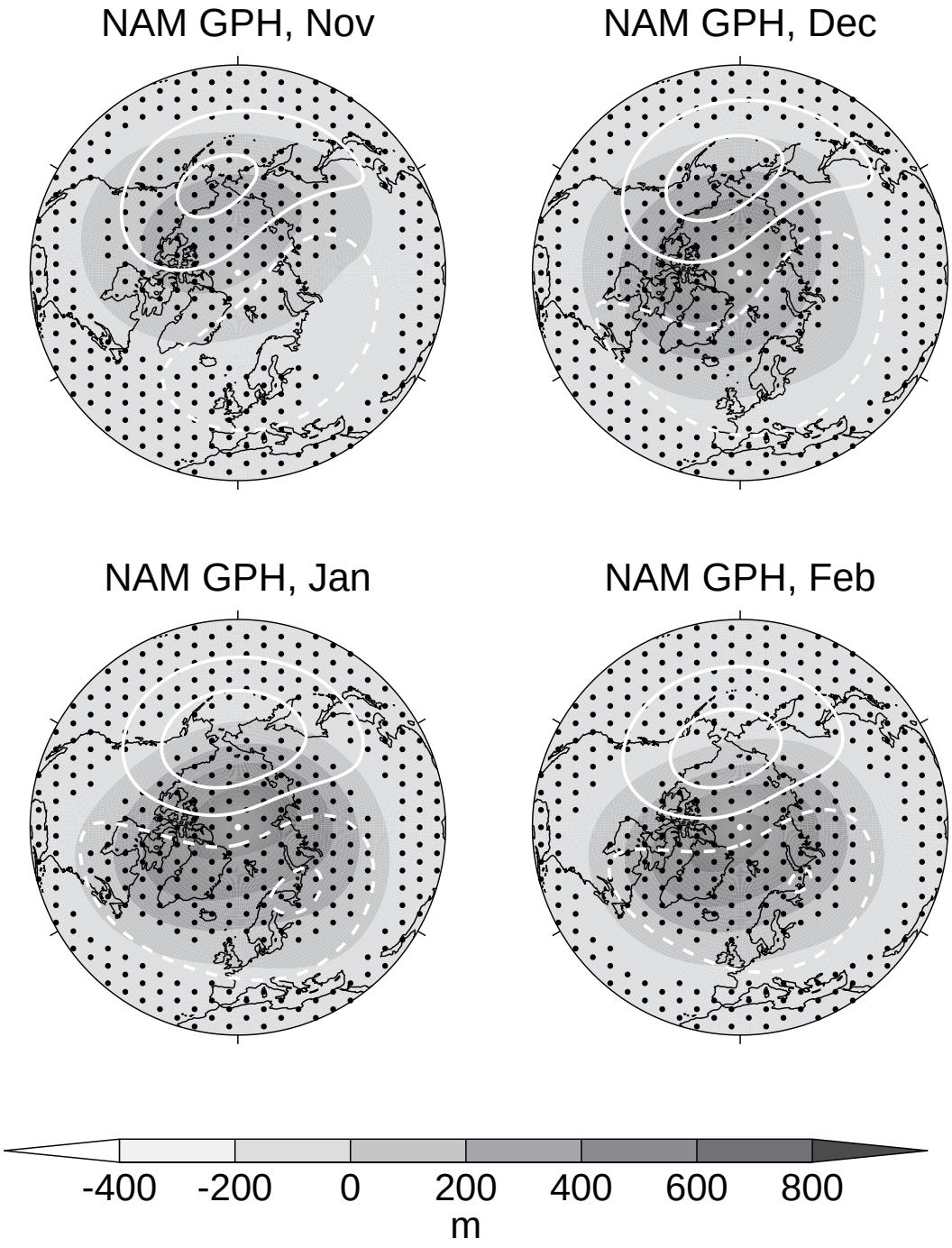

Figure 4.4: Regression of 10 hPa GPH north of 20N onto the NAM index in ERA-40 (greyscale), showing a close resemblance to the QBO-E minus QBO-W composite differences in figure 4.3. White contours show the climatological zonally asymmetric component of GPH with contour values ±200 and ±600 m with negative contours dashed. Stippling shows where the regression coefficients are statistically significant above the 95% level.

There is a similar correspondence between the QBO-E minus QBO-W composite difference and the NAM signature of GPH. Figure 4.3 shows the QBO-E minus QBO-W composite difference in the monthly mean NH 10 hPa GPH from November to February in ERA-40. The climatological eddy component, defined as the climatological GPH with the zonal mean subtracted, is also shown. Figure 4.4 shows the signature of the NAM in NH 10 hPa GPH, which bears a very good resemblance to the GPH differences between QBO-E and QBO-W. Anomaly correlations[1] with the NAM signature north of 20N are indicated below the composite GPH differences in figure 4.3. The correlations are between 0.79 and 0.95 in November–January, with a lower correlation in February when the differences are not highly statistically significant. The correlations are all greater if only anomalies north of 60N are considered, so this is not simply arising from a direct influence of the QBO on the subtropics.

In November, the top left panel of figure 4.3 shows GPH is greater in QBO-E over the Canadian Arctic and less over northern Europe. This represents positive interference with the wavenumber-1 part of the climatological wave pattern, and so the wavenumber-1 amplitude is greater, whilst the wavenumber-2 amplitude is slightly less (figure 4.5). Over the course of winter, however, the pattern shifts so that more positive GPH is found over the Arctic and north Atlantic with lower GPH over the North Pacific in January and February. This gives weak destructive interference with the climatological waves, so wavenumber-1 and 2 amplitudes are both slightly reduced. Hu and Tung [2002] argued that this indicates that the HT mechanism is not important in late winter.

Figure 4.5 also shows the change in the wavenumber-1 and 2 amplitudes at (60N, 10 hPa) associated with the QBO influence on the NAM only. To calculate this, the QBO-E minus QBO-W composite difference in the NAM index is multiplied by the

[1]The anomaly correlation between 2D anomaly patterns $\mathbf{x}$ and $\mathbf{y}$ is defined here as

$$\frac{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j} w_i x_{i,j} y_{i,j}}{\sqrt{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j} w_i x_{i,j}^2}\sqrt{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j} w_i y_{i,j}^2}},$$

where n_i and n_j are the number of gridpoints along the latitude and longitude dimensions and w_i is the cosine of the latitude.

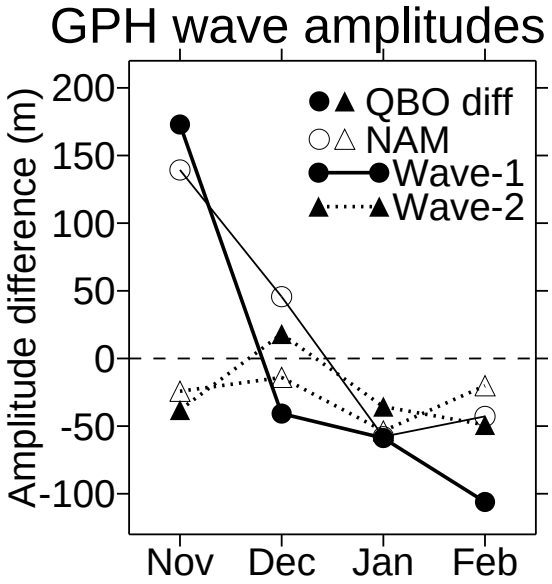

Figure 4.5: Lines with filled symbols show the QBO-E minus QBO-W composite differences of GPH wavenumber-1 amplitude (solid lines and circles) and wavenumber-2 amplitude (dotted lines and triangles) at 60N and 10 hPa for months November–February. Unfilled symbols show the change in wave amplitudes associated just with the QBO-E minus QBO-W difference in the NAM index. This shows that the main qualitative features of the QBO-E minus QBO-W differences in GPH wave amplitudes are largely explained just by QBO modulation of the NAM.

NAM GPH signature in figure 4.4. This is then added to the climatological GPH and the change in wave amplitudes computed. This shows that the observed changes in the wave amplitudes in QBO-E versus QBO-W correspond closely to the seasonal evolution of the NAM signature.

Previous work has noted that the QBO-E minus QBO-W composite difference of extratropical ZMZW and zonal mean GPH is very similar to the NAM signature [Dunkerton and Baldwin, 1991; Kodera, 1995; Ruzmaikin *et al.*, 2005], but here it has been shown that this is true for the EP flux and GPH wave component differences as well, which is important given that these differences have been used to try to understand the mechanism of the QBO's influence.

To determine if this behaviour may occur because the QBO is playing a role in setting the structure of the NAM pattern, the NAM signatures of EP flux and GPH were calculated using data only in the third of years in ERA-40 when equatorial winds

at 50 hPa were closest to their mean. These NAM signatures are very similar to those shown in figures 4.2 and 4.4. Thus the strong correlation between the QBO-E minus QBO-W composite differences and the NAM is not due to the QBO affecting the NAM structure.

4.2 Comparison with the behaviour of other dynamical systems

A similarity between the leading EOF of a system and its long-term mean response to an applied forcing has been observed in studies of other dynamical systems. Palmer and Weisheimer [2011] illustrate this in the simple system of Lorenz [1963] (that which gives rise to the famous Lorenz butterfly attractor). This system is described by the equations (with commonly used values of constants):

$$\begin{aligned} \dot{x} &= 10(y - x); \\ \dot{y} &= x(28 - z) - y; \\ \dot{z} &= xy - \tfrac{8}{3}z. \end{aligned} \tag{4.1}$$

Applying a steady force in any direction in the xy-plane gives rise to a shift in the system's mean state which is very nearly aligned with the system's leading EOF. Therefore the spatial pattern of the response closely resembles that of the leading EOF, which lies in the xy-plane at an angle of about 50° to the x-axis (figure 4.6).

Studies that have examined this in atmospheric models include that of Son and Lee [2006], who found in a tropospheric GCM that their model's steady-state response to tropical heating or high-latitude cooling often closely resembles the tropospheric NAM, and those of Ring and Plumb [2007, 2008], who found in another tropospheric GCM that the steady-state response to various mechanical and thermal forcings applied in the extratropics is also very NAM-like. Palmer [1999] argued that this behaviour could be a result of the atmosphere having preferred "regimes". Branstator and Selten [2009] examined the reasons why the response to greenhouse gas forcing in a tropospheric GCM is NAM-like. They conclude that it largely results from a linear effect whereby anomalies in the NAM tend to persist for a long time, so the NAM is a prominent pattern in natural variability and in the response to a forcing after

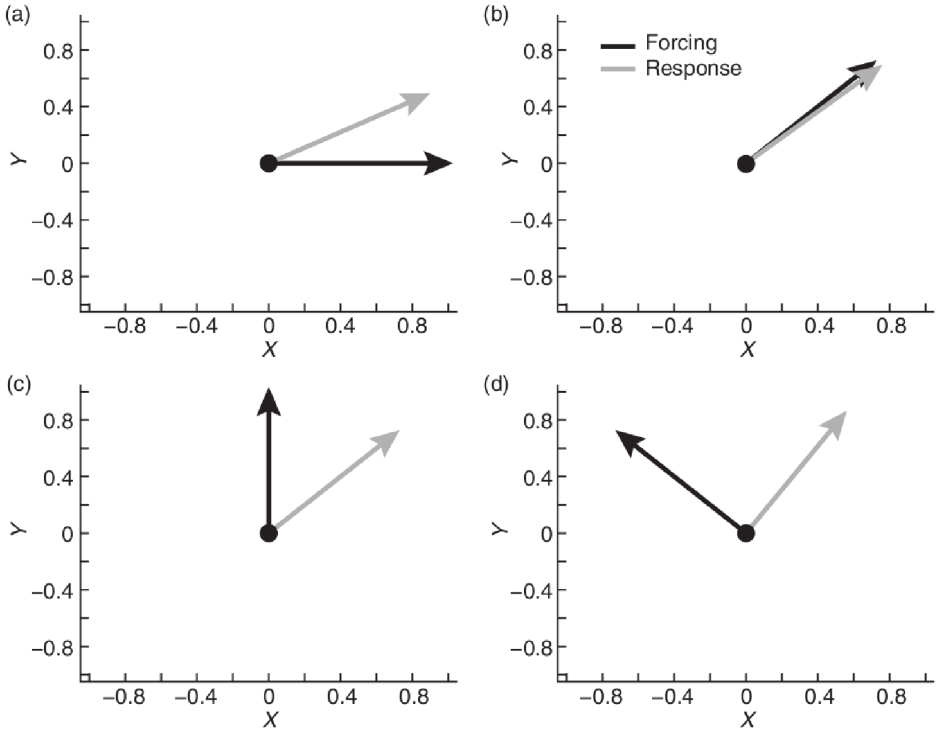

Figure 4.6: Time-mean response to forced Lorenz system (grey arrow) in relation to direction of forcing (black arrow) in the x-y plane: (a) $\theta = 0$, (b) $\theta = \pi/4$, (c) $\theta = \pi/2$ and (d) $\theta = 3\pi/4$, where θ is the angle of the forcing to the x-axis. The response is very nearly aligned with the system's leading EOF. From Palmer and Weisheimer [2011]. © Taylor & Francis. Used with permission.

time-averaging. The SH response to anthropogenic and natural forcings in GCMs is also SAM-like in some studies as well [e.g. Arblaster and Meehl, 2006; Son *et al.*, 2008, 2010]. This behaviour is also predicted by the fluctuation-dissipation theorem [Gritsun and Dymnikov, 1999], which has been argued to apply approximately to the atmosphere [e.g. Leith, 1975; Gritsun and Branstator, 2007].

Invoking this behaviour, which is present in atmospheric models and also in simpler systems like that of Lorenz [1963], may then explain the seasonal evolution of the pattern of EP flux and GPH differences in QBO-E and QBO-W, and supports the suggestion of Dunkerton and Baldwin [1991] that the QBO "excited a fundamental...mode of variability in the extratropical atmosphere". It may also explain the findings of previous studies that identify the stratospheric NAM as a prominent

pattern that appears not just in response to the QBO but also to other important natural influences on the vortex, including volcanic eruptions [e.g. Kodera, 1995; Stenchikov *et al.*, 2006], ENSO [e.g. Sassi *et al.*, 2004] and the solar cycle [e.g. Kodera, 1995; Labitzke, 2005]. It should be noted, however, that this behaviour is not fully general – for example, Son and Lee [2006] found that the tropospheric ZMZW response to tropical heating and high-latitude cooling in their GCM is not always NAM-like, Woollings [2008] showed that the vertical structure of tropospheric NH climate change projected by GCMs in the Coupled Model Intercomparison Project phase 3 differs from that of the NAM, and projected SH climate change was not found to be SAM-like by Shindell and Schmidt [2004]. In chapters 5 and 6 it is shown that the stratospheric response to an applied torque has much in common with annular mode dynamics, though there are differences between the simulated responses and the NAM.

Importantly, this behaviour implies that examining the long-term mean response to a forcing or using compositing does not in general yield much information about the forcing mechanism. If an anomaly in response to an unknown forcing resembles the leading EOF, it does not allow the nature of the forcing to be deduced since it could be the case that many different forcings give rise to this response. Therefore, the HT mechanism seems consistent with the observed vortex response to the QBO on monthly time scales, as it simply predicts that the vortex will be weaker in QBO-E, and the above discussion indicates this may manifest itself as a modulation of the NAM in composite analysis. However, these observations could also be consistent with other mechanisms that predict a weakening of the vortex during QBO-E. Furthermore, structure in the extratropical QBO-E minus QBO-W composite differences seems to be primarily related to the NAM signature rather than the forcing mechanism, and is not a reliable indicator of the mechanism.

Another way to understand the difficulty in using these composite differences to infer the forcing mechanism behind the HT relationship is that the differences in the wave components of the flow have contributions not only from the QBO-E/W equatorial wind pattern but also from the effect of the weaker vortex, as changes in the

zonally symmetric component of the flow will cause changes in the wave components. (This was also noted by Holton and Tan [1980] and Dunkerton and Baldwin [1991].) What is required is a way of computing the changes due to QBO-E/W forcing whilst the vortex state is close to constant.

4.3 Motivating examination of the short-term transient response

Here it is argued that whilst the QBO-E minus QBO-W composite vortex differences cannot be relied on to show the mechanism of the QBO's influence, the full time-dependent transient response of the vortex to QBO forcing should be much more useful.

If the forcing mechanism is simple (meaning it does not influence the system in question in a series of intermediate steps), then the response on short time scales following application of the forcing can be expected to show the mechanism clearly. Consider a system described by state vector $\mathbf{x}(t)$ that evolves according to equations

$$\dot{\mathbf{x}}(t) = \boldsymbol{\mathcal{L}}(\mathbf{x}(t), t), \tag{4.2}$$

where the equations may be non-linear and are explicitly time-dependent for generality. The equations governing atmospheric motion may be expressed in this form [Palmer, 1999]. Consider also a forced variant of this system described by state vector $\mathbf{x}'(t)$ which evolves according to similar equations with the addition of a state- and time-dependent forcing term,

$$\dot{\mathbf{x}}'(t) = \boldsymbol{\mathcal{L}}(\mathbf{x}'(t), t) + \mathbf{f}(\mathbf{x}'(t), t) \tag{4.3}$$

with $\mathbf{x}(0) = \mathbf{x}'(0) = \mathbf{x}_0$. In the context of the HT relationship, $\mathbf{x}$ and $\mathbf{x}'$ would represent the vortex state and $\mathbf{f}$ the influence of the QBO. Then as long as the difference between the state vectors, $\delta\mathbf{x} = \mathbf{x}'(t) - \mathbf{x}(t)$, is analytic, which will be the

case if both $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ are analytic, $\delta\mathbf{x}(t)$ can be evaluated for small t by writing its Taylor series (in index notation using summation convention) to give

$$\delta x_a(t) \;=\; f_a(\mathbf{x}_0, 0)\, t + \left(\left.\frac{\partial \mathcal{L}_a}{\partial x_b}\right|_{(\mathbf{x},\, t)=(\mathbf{x}_0,\, 0)} \delta\dot{x}_b(0) + \left.\frac{\mathrm{d} f_a}{\mathrm{d}t}\right|_{(\mathbf{x},t)=(\mathbf{x}_0,\, 0)} \right) \frac{t^2}{2} + \mathcal{O}(t^3). \quad (4.4)$$

Thus for short times, the difference between the systems is nearly proportional to the applied forcing at $t = 0$, when this is non-zero. State-dependence of the equations of motion (non-zero $\partial\mathcal{L}_a/\partial x_b$) acts to complicate the relationship between $\delta\mathbf{x}$ and the forcing as t increases and may be expected to become important on a time scale of the order of the system's dynamical time scale. (If $\mathbf{f}(\mathbf{x}_0, 0) = 0$, it can similarly be shown that $\delta\mathbf{x}(t)$ is proportional to $\mathrm{d}\mathbf{f}/\mathrm{d}t|_{(\mathbf{x},\, t)=(\mathbf{x}_0,\, 0)}$ to $\mathcal{O}(t^2)$ when this is non-zero, which is relevant for the results of chapters 5, 6 and 8.) Examining the system's response on short time scales shows the direct effect of the forcing before feedbacks due to the change in the state of the system become large.

Equation 4.4 can be re-written

$$\delta x_a(t) = \int_0^t f_a(\mathbf{x}, t')\, \mathrm{d}t' + \left.\frac{\partial \mathcal{L}_a}{\partial x_b}\right|_{(\mathbf{x},\, t)=(\mathbf{x}_0,\, 0)} \delta\dot{x}_b(0) \frac{t^2}{2} + \mathcal{O}(t^3).$$

Therefore, for forcing mechanisms that unfold in several steps or are otherwise time-dependent, the transient response would be expected to clearly show the cumulative forcing up to a time scale of the order of the dynamical time scale, which may still give useful information for testing hypotheses.

The system described by $\mathbf{x}'$ may be perturbed from that described by $\mathbf{x}$ in other ways, for example by perturbing the system state directly or by changing the form of $\mathcal{L}$, and evaluating the Taylor series of $\delta\mathbf{x}$ shows again that on short time scales $\delta\mathbf{x}$ is quite simply related to the way in which the system is perturbed at $t = 0$. Any two sets of equations of motion for which $\mathbf{x}$ and $\mathbf{x}'$ are analytic can be written in the form of equations (4.2) and (4.3) for a suitable choice of "effective forcing" in place of $\mathbf{f}$ and, to first order in t, $\mathbf{x}$ and $\mathbf{x}'$ will diverge as if the state $\mathbf{x}'$ is affected

by a steady forcing[2]. Therefore this discussion applies quite generally. So if $\mathbf{x}$ were considered to be the atmospheric state, including both the QBO and the vortex, so that the influence of equatorial winds on the vortex could be thought of as the vortex response to perturbations in the equatorial wind part of $\mathbf{x}$, then the short-term transient vortex response would again show clearly the direct way in which the vortex state changes in response to changes in equatorial winds.

No study has argued that any feedback of vortex changes onto the QBO is important for the mechanism behind the HT relationship, but even if it were it can be shown that the short-term transient response of the vortex to perturbations of equatorial winds would show clearly the one-way influence of these winds on the vortex. The effect of feedbacks would only become important at later times[3].

Hence the transient response of a system to a forcing can generally be expected to contain detailed information on the direct effect of a forcing, particularly on short time scales following application of the forcing. Thus the transient response can be used more effectively to discriminate between different hypotheses for a forcing mechanism

[2]To make this explicit, for the effect of a perturbation to the system state, set $\dot{\mathbf{x}} = \boldsymbol{\mathcal{L}}(\mathbf{x}, t)$ and $\dot{\mathbf{x}}' = \boldsymbol{\mathcal{L}}(\mathbf{x}', t)$ with $\mathbf{x}'(0) = \mathbf{x}(0) + \delta\mathbf{x}(0) = \mathbf{x}_0 + \delta\mathbf{x}(0)$. These equations may be written in the form of equations (4.2) and (4.3) by substituting the effective forcing $\boldsymbol{\mathcal{L}}(\mathbf{x}', t) - \boldsymbol{\mathcal{L}}(\mathbf{x}, t)$ for $\mathbf{f}$. Similarly, for a change in the equations of motion, set $\dot{\mathbf{x}} = \boldsymbol{\mathcal{L}}(\mathbf{x}, t)$ and $\dot{\mathbf{x}}' = \boldsymbol{\mathcal{L}}'(\mathbf{x}', t)$ with $\mathbf{x}'(0) = \mathbf{x}_0$, and the effective forcing is $\boldsymbol{\mathcal{L}}'(\mathbf{x}', t) - \boldsymbol{\mathcal{L}}(\mathbf{x}, t)$. An effective forcing can be defined for any analytic $\mathbf{x}$ and $\mathbf{x}'$ in a similar way. Whilst the effective forcing evaluated for all t could be very complicated, the $\mathcal{O}(t)$ part of $\delta\mathbf{x}$ is proportional to this effective forcing at $t = 0$, and so on short time scales the shifted difference vector $\delta\mathbf{x} - \delta\mathbf{x}(0)$ will behave as if it arises due to a steady applied forcing in accordance with equation (4.3), equal to $\boldsymbol{\mathcal{L}}(\mathbf{x}_0 + \delta\mathbf{x}(0), 0) - \boldsymbol{\mathcal{L}}(\mathbf{x}_0, 0)$ and $\boldsymbol{\mathcal{L}}'(\mathbf{x}_0, 0) - \boldsymbol{\mathcal{L}}(\mathbf{x}_0, 0)$ for the system state perturbation and change in the equations of motion respectively. This can be thought of as the instantaneous effect of the perturbation of the system on the evolution of the state vector.

[3]Let the state vector of the system $\mathbf{x}$ be divided into two interdependent parts such that $\mathbf{x} = (\mathbf{u}, \mathbf{v})$. Then the equations of motion become, without loss of generality,

$$\dot{\mathbf{u}}(t) = \boldsymbol{\mathcal{J}}(\mathbf{u}, \mathbf{v}, t); \;\; \dot{\mathbf{v}}(t) = \boldsymbol{\mathcal{K}}(\mathbf{u}, \mathbf{v}, t);$$

with $\boldsymbol{\mathcal{L}} = (\boldsymbol{\mathcal{J}}, \boldsymbol{\mathcal{K}})$. Then suppose $\mathbf{v}' = \mathbf{v} + \delta\mathbf{v}$ and $\mathbf{u}' = \mathbf{u} + \delta\mathbf{u}$ together represent a perturbed state of the system. In the context of the HT relationship, $\mathbf{v}$ may represent the equatorial wind and $\mathbf{u}$ the extratropical wind. Let $\mathbf{u}_0 = \mathbf{u}(t = 0) = \mathbf{u}'(t = 0)$ and $\mathbf{v}_0 = \mathbf{v}(t = 0)$ with $\delta\mathbf{v}(0) \neq 0$, representing a perturbation in the equatorial wind state. Then the time evolution of $\delta\mathbf{u}$ written as a Taylor expansion becomes

$$\delta\mathbf{u} = \left[\boldsymbol{\mathcal{J}}(\mathbf{u}_0, \mathbf{v}_0 + \delta\mathbf{v}(0), 0) - \boldsymbol{\mathcal{J}}(\mathbf{u}_0, \mathbf{v}_0, 0)\right] t + \mathcal{O}(t^2),$$

so on short time scales any dependence of $\dot{\mathbf{v}}$ on $\mathbf{u}$ does not have a big influence on the evolution of $\delta\mathbf{u}$.

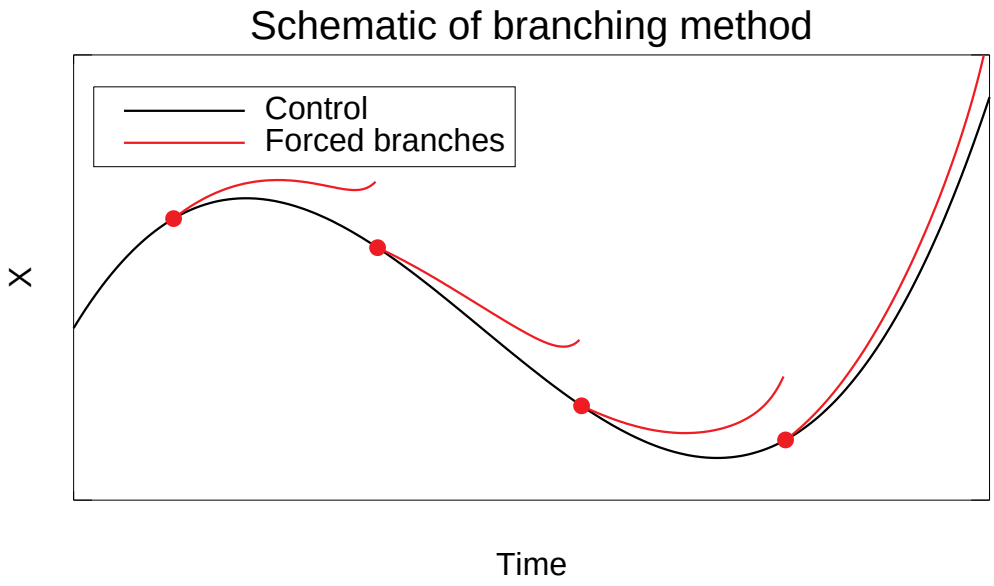

Figure 4.7: A schematic illustration of the method of branching runs with an applied forcing from an unforced control run so that the transient response to the forcing of the state vector "X" on short time scales following branching can be examined for a range of initial conditions.

than the long-term mean response in general. Figure 4.7 illustrates schematically how the transient response of a system to a forcing may be measured, by taking initial conditions from the control run of a numerical model of the system and branching off runs with a forcing applied, and examining the difference in the evolution of the model state. The basic idea is somewhat similar to that discussed by Palmer and Weisheimer [2011] to diagnose the causes of model errors by examining error growth on short time scales, but here it is applied to a different problem, namely understanding a forcing mechanism. For studies of the atmosphere, it is probably necessary to use a numerical model to evaluate the transient response. Using observations would require identifying two near-identical atmospheric states with different values of the forcing in question, which is practically impossible.

The principle is illustrated here in the Lorenz [1963] system. The set of equations (4.1) was integrated numerically using the fourth-order Runge-Kutta method[4] for 1 time unit starting from the point $(10, 10, 10)$ so that the system's trajectory

[4]The code was written in IDL and taken from http://www.exelisvis.com/Support/HelpArticleDetail/ArticleId/3743.aspx .

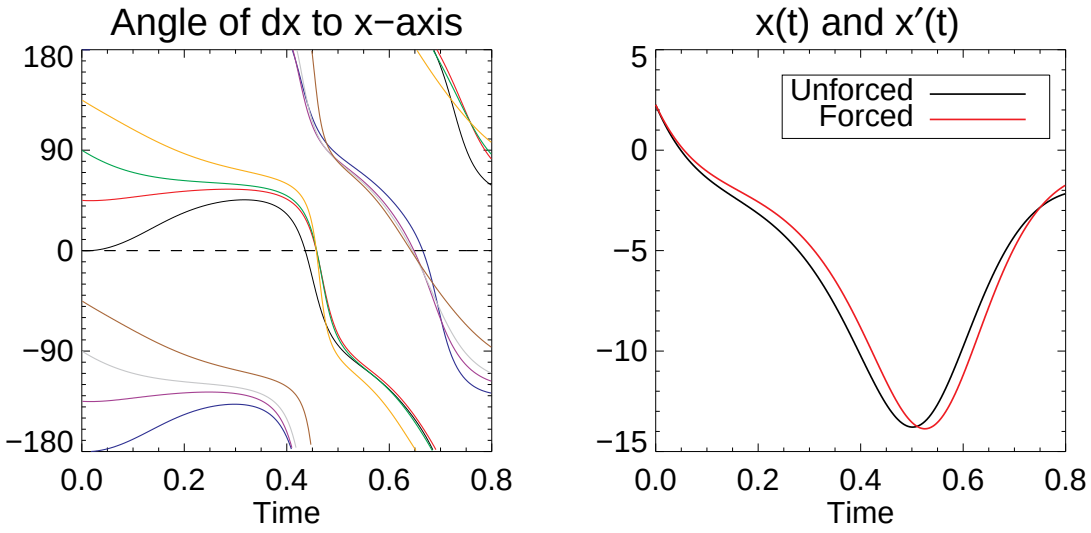

Figure 4.8: (a) shows the time evolution of the angle between the x-axis and the vector defined as the difference between the position of a particle following equations (4.1) with the addition of a constant forcing in the xy-plane and a particle with the same initial position that follows equations (4.1) with no forcing. Each line corresponds to a different forcing, with forcings applied in the x-y plane at angles $-180°$ to $145°$ to the x-axis in $45°$ increments. The angle of the difference vector matches the forcing angle for small times, and then takes on similar values for the different forcings. (b) shows time series of the x-coordinates of the unforced system and the system whose forcing is parallel to the x-axis. The deviations between the systems are small for the time interval shown.

was very close to the attractor after this time. The system was integrated for a further 9 time units, and a point on this trajectory was randomly selected (the results are not sensitive to the choice of this point). 8 forced trajectories were integrated from this point, where the equations were modified to include a constant force in the xy-plane of magnitude 5 velocity units, at angles $-180°$ to $145°$ to the x-axis in $45°$ increments. For the force at $45°$, which is almost along the line joining the two centres of the "wings" of the butterfly attractor, this force causes the trajectory to be in the domain $x > 0$ for about 70% of the time in a long integration (compared to 50% of the time in the unforced integration), corresponding to the forcing causing the trajectory to spend more time in the region of the fixed point in $x > 0$.

Figure 4.8(a) shows the time evolution of the angle between the x-axis and the difference vector between the state vectors of the forced and unforced systems for

each forcing angle. It can be seen that at small times, the effect of each forcing is clearly distinguished, with the angle of the difference vector very nearly the angle of the applied force, as expected from equation (4.4). The angle of the difference vector then drifts until it becomes very similar for each of the applied forcings after about 0.4 time units. There is also an indication of a preference for the difference vector to align at angles around $50°$ and $-130°$, which are approximately the angles the system's leading EOF makes with the x-axis. This corresponds to the phenomenon illustrated by Palmer and Weisheimer [2011] which is that the steady-state response to these applied forcings is the same (apart from the sign). (The z-component of the difference vector also evolves in a qualitatively similar way for each forcing angle, so adding this information does not help distinguish between the forcing angles.)

Figure 4.8(b) shows time series of the x-coordinates of the control trajectory and of the trajectory of the system with a force applied along the x-axis. This illustrates that the time scale over which information about the forcing angle is lost is very rapid and occurs long before the two trajectories have substantially diverged, so it is necessary to examine the difference vector a very short time after the forcing starts being applied in order to be able to infer the angle of the applied forcing.

Dynamical time scales in the wintertime polar mid-stratosphere are of the order of a week, so it could be anticipated that the relationship between the vortex response and the QBO forcing will be complicated by state-dependence of the equations of motion on this time scale. Previous studies of the transient vortex response to an imposed equatorial wind forcing in models [O'Sullivan and Salby, 1990; Balachandran *et al.*, 1991; Holton and Austin, 1991; O'Sullivan and Young, 1992; O'Sullivan and Dunkerton, 1994; Gray *et al.*, 2003, 2004; Gray, 2003; Garfinkel *et al.*, 2012b] have not clearly shown the response on time scales shorter than this. Therefore the direct effect on the vortex of changing the equatorial winds is not clear from these studies, and the short-term transient response is examined in chapter 8. Interpretation of the transient response could still be complicated, however: there is likely to be some time delay of unknown duration between the QBO being imposed and the vortex responding due to the spatial separation of the QBO and the vortex; the QBO may

have several effects on high latitudes which may not all be important for the HT relationship; and important effects may take longer than the dynamical time scale to substantially influence the response. The transient response provides information that is complementary to that from other experimental designs, for example using various idealised equatorial wind profiles to test mechanisms, and a full understanding of the mechanism will require combining results using various experimental approaches.

Chapter 5

Stratospheric response to an extratropical zonal torque: long-term mean response

In this chapter the UKMO Stratosphere-Mesosphere Model (SMM; section 3.3) is used to examine the stratospheric response to applied zonally symmetric extratropical zonal torques, with a focus on the mean response on time scales longer than a month. There are two main objectives to this work.

Firstly, in chapter 4 it was suggested that the long-term mean response of the NH extratropical stratosphere to an arbitrary forcing may be like the NAM, due to "feedbacks" arising from state-dependence of the atmospheric equations of motion dominating the "direct effect" of the torque after the torque has been applied for a sufficiently long time (section 4.3). It was also suggested that this may help to explain the spatial structure and seasonal evolution of the observed vortex response to external forcings, including the QBO. If this is correct then the long-term mean response to a torque will also be NAM-like. This suggestion was based on results from experiments on the Lorenz [1963] system [Palmer and Weisheimer, 2011] and models of the troposphere [Son and Lee, 2006; Ring and Plumb, 2007, 2008; Branstator and Selten, 2009], but the dynamics of these systems is quite different to the dynamics of the stratosphere and the experiments performed here test whether the large-scale dynamics of the stratosphere displays the same behaviour. It was noted in chapter 4 that this behaviour is not fully general and it is possible that it is not manifested in the

stratosphere. These experiments are similar to those performed by Ring and Plumb [2007] to study the tropospheric response to an applied torque in perpetual January conditions, but a wider variety of applied torques is considered and the response with a seasonal cycle is examined as well as that under perpetual January conditions.

The second objective of this work is to gain a better understanding of the interaction between the planetary wave and zonal mean parts of the flow. It was explained in section 2.2.1 that it is difficult to understand this using analytical methods, and in section 2.4 it was illustrated that wave-mean flow interaction can give rise to complicated behaviour such as vacillation cycles. This interaction has been explored in depth with regard to explaining the stratospheric climatology and the dynamics of SSWs (section 2.2). However, there are not many studies, to my knowledge, which consider the physics of the interaction associated with smaller perturbations to the vortex caused by external forcings such as the QBO. It is not clear, for example, that associated changes in the EP flux can generally be thought of as being due to changes in wave propagation. The results presented here and in chapter 6 provide another situation to aid conceptual understanding of the physics which is relevant to understanding such responses. Furthermore, whilst the stratospheric response to an imposed torque δX is well understood when the associated EP flux convergence feedback δD_F is specified, it is not understood how to predict δD_F (section 2.2.3). The experiments presented here aid understanding of the δD_F response.

These experiments may also have direct relevance to understanding the HT relationship. The QBO may act directly by affecting extratropical EP flux convergence (section 2.7.2), which is the same as applying a zonally symmetric torque as far as the extratropical zonal mean circulation is concerned. Then the resulting wave feedbacks could be expected to have similarities with those in the results presented here. In the real atmosphere, angular momentum and energy obey local conservation laws, whereas in the experiments performed here these quantities have sources and sinks. However, it will be shown that the response to an applied torque is not sensitive to the structure of the torque, and it is shown in chapter 6 that the response is not strongly dependent on the initial conditions on time scales up to about two weeks.

Therefore, it is not expected that including a change in the wave structure in order to satisfy the conservation laws would qualitatively change the results, so that the results could help understanding of the HT relationship. The imposed torque is likely to be a simpler perturbation to the extratropics than the forcing by the QBO, which may have some dependence on the atmospheric state, may be zonally asymmetric and may be stochastic in nature, though the mean effect may be qualitatively similar to that of a fixed zonally symmetric torque.

The studies of Song and Robinson [2004] and Chen and Zurita-Gotor [2008] have previously shown the response to stratospheric zonal torques in models of the troposphere and stratosphere with zonally-symmetric boundary conditions, with their main objective being to examine the tropospheric response. However, the models that were used had weak planetary waves and are unlikely to have correctly simulated the changes in stratospheric waves. Cohen *et al.* [2013] reported that changing parameters in a momentum-conserving gravity wave parameterisation, in an atmospheric primitive equation model with zonal wavenumber-2 topography, results in the EP flux due to resolved planetary waves adjusting to cancel out most of the change in the gravity wave drag. They found that this also occurs in the steady-state response to a fixed zonally symmetric zonal torque, with the applied torque being similar to a change in gravity wave drag except it does not respond to changes in the circulation and is zonally symmetric. The experiments performed here extend upon this work by considering a wider range of torques under varied boundary conditions and examining more aspects of the circulation change.

The main simplifications of the SMM compared to the real atmosphere are that the GPH at 16 km is fixed as a lower boundary condition and gravity wave drag is parameterised by Rayleigh friction. The SMM is used here to understand *qualitatively* how the feedbacks from large-scale stratospheric dynamics, which are expected to be the most important feedbacks of those internal to the stratosphere, affect the response to a torque. The experiments are idealised, with wave generation by the troposphere unchanged by the torques. This allows the stratospheric dynamical response to a torque to be isolated, so the role of internal stratospheric dynamics is clear. In a

model that also simulates the troposphere, it is difficult to know what portion of the stratospheric response is due to internal stratospheric dynamics and what to forcing associated with tropospheric changes. In addition, the relative simplicity of the SMM compared to more complex GCMs allows the response to a wide range of torques to be studied, which would be too computationally expensive in the latter, and also allows the effects of modifying the boundary conditions to be investigated. The response between 16–50 km is focussed on as the mesospheric response is not likely to be simulated realistically due to the Rayleigh friction. The simulated response on long time scales may be different to the response of the real stratosphere to an externally-induced torque if feedbacks from tropospheric changes and gravity waves are important. This is discussed in more depth in section 5.5. (Note it is shown that greatly weakening the Rayleigh friction does not qualitatively change the results in appendix 5.A.)

5.1 Experimental method

Control runs of the SMM, for which the model equations were unmodified, were performed both under perpetual January conditions and with a seasonal cycle, defined by particular specifications of the solar radiation and 16 km GPH. Using both types of boundary conditions indicates the robustness of features of the stratospheric responses to applied torques, and shows which features depend on the climatology and on the nature of the boundary conditions. Section 5.1.1 describes the control runs used in most of the experiments, referred to as the "standard" control runs. Other control runs set up with different boundary forcings are used in some experiments and are described in the main text. To examine the responses to applied torques, runs were performed with the model equations modified to include a forcing term in the zonal momentum equation, as described in section 5.1.2. Data were sampled at daily intervals.

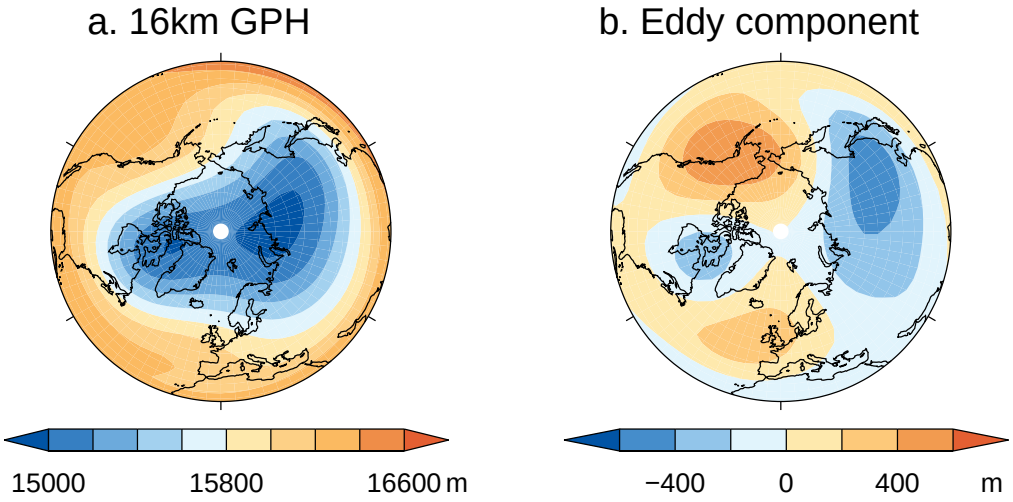

Figure 5.1: (a) shows the specified 16 km GPH used as the bottom-boundary forcing of the SMM in most of the perpetual January simulations, and (b) shows the zonally asymmetric component.

5.1.1 Control runs

5.1.1.1 Perpetual January control run

The model was run for 20 years with radiative conditions set to those for January 15 in the MIDRAD scheme. The first two years were not used in the following analysis in order to allow the model sufficient time to adjust to the forcing conditions. The 16 km GPH was set as the January climatology of the 100 hPa GPH in the HadGEM2-CCS 240-year pre-industrial control run (section 3.4), with the zonally asymmetric component multiplied by a factor of 2 in order for the model to exhibit irregular vacillations that have qualitative similarities with observed stratospheric variability (figure 5.1). The wavenumber-1 and 2 amplitudes at 62.5N are 250 m and 360 m respectively, which are similar to the amplitudes of specified bottom boundary single-wavenumber GPH eddies used in previous studies of stratospheric internal variability [e.g. Gray *et al.*, 2003; Scott and Polvani, 2006]. The need to scale up the wave amplitude at 16 km is not surprising, as in the real stratosphere the tropospheric wave forcing is time-varying and climatological averaging smooths bursts of intense wave activity that cause deceleration of the ZMZW as the waves propagate upward

[Matsuno, 1971], so that steady wave forcing is less effective at bringing about ZMZW reversals. As shown in section 3.4, the 100 hPa GPH in HadGEM2-CCS has a realistic wave structure. Data from HadGEM2-CCS were used in preference to observations because, for the runs with a seasonal cycle, the HadGEM2-CCS data provide over two hundred independent winters that are necessary to gain good statistical significance in those experiments. For consistency, HadGEM2-CCS is therefore also used to construct the bottom boundary in the perpetual January runs so they are more easily compared with the runs with a seasonal cycle.

Overall, it is shown below that this control run has a qualitatively reasonable extratropical stratosphere but quantitatively it is quite different from January observations. Since the seasonal cycle control run exhibits a wintertime climatology in much closer agreement with observations (section 5.1.1.2), this seems to be due to the simple perpetual January boundary conditions which are not realistic. These simulations are useful for understanding qualitatively the response to a torque and for indicating how the response depends on the model set up and climatology.

Figure 5.2 shows the climatology of the perpetual January control run. Figure 5.2(a) shows the climatological ZMZW and the difference from the ERA-40 January climatology (due to the simple boundary conditions, the model climatology would not be expected to resemble ERA-40 closely and the difference is just presented for comparison). The modelled polar vortex is considerably weaker than that in ERA-40 and is further equatorward, with the peak ZMZW around 30N in the middle and upper stratosphere compared to around 60N in ERA-40 (figure 2.1). Figure 5.2(b) shows the standard deviation of the ZMZW, which is up to $\sim$10 ms^{-1} less than that in ERA-40 in the NH extratropical mid-stratosphere, though it still peaks at high latitudes despite the vortex being further south. In the tropical stratosphere the modelled standard deviation is up to $\sim$20 ms^{-1} too small due to the absence of the QBO and SAO. Figure 5.2(c) shows a time-height section of the ZMZW at 57.5N over 365 days, illustrating the irregular vacillation cycles with ZMZW reversals and intensifications

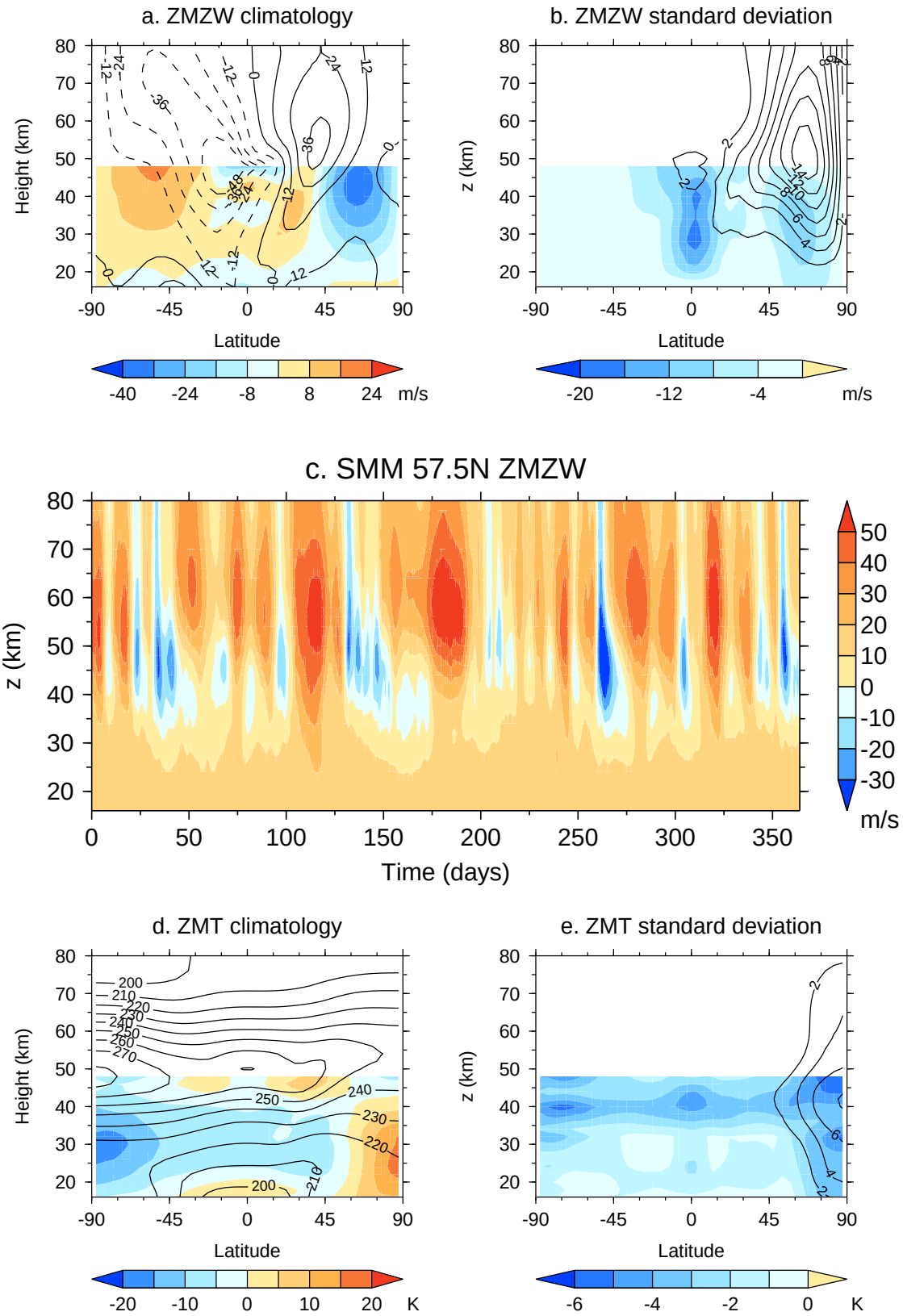

Figure 5.2: (Continued on the following page.)

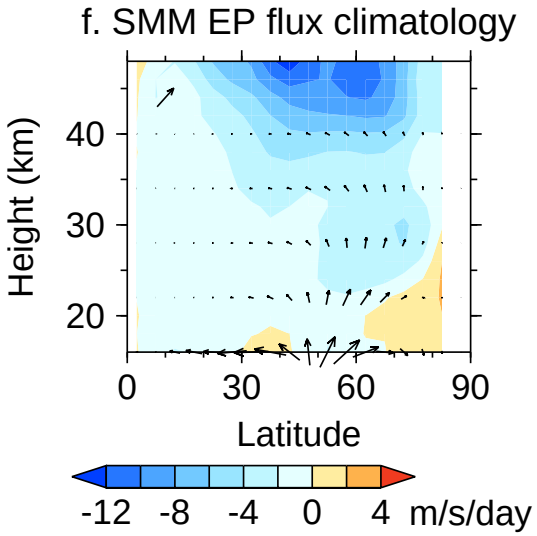

Figure 5.2: (Continued.) The perpetual January SMM standard control run. (a) shows the mean ZMZW (contours) and the difference from the ERA-40 January climatology (colours) and (b) shows the same for the ZMZW standard deviation. (c) shows a time-height section of the ZMZW at 57.5N for one year. (d) and (e) show the same as (a) and (b) but for the ZMT. (f) shows the mean EP flux (arrows, shown every 6 km in height) and D_F (colours) below 50 km. A reference arrow in the top left corner has a magnitude of $(10^7,\ 10^5)\,\mathrm{kg/s^2}$. D_F at 87.5N is not plotted as the differentiation error is large there.

that often propagate down to the lower stratosphere[1] in a qualitatively similar way to observed variability. Figure 5.2(d) shows the ZMT climatology and the differences from that in January in ERA-40. The NH polar lower and middle stratosphere is too warm by up to ∼15 K and the tropical and NH mid-latitude mid-stratosphere too cold by ∼5–10 K, consistent with the vortex being weaker. The SH polar mid-stratosphere is too cold by up to ∼20 K. Figure 5.2(e) shows the ZMT standard deviation which is less than that in ERA-40 by up to ∼5K in the NH polar mid-stratosphere, consistent with the ZMZW variability also being less. Figure 5.2(f) shows the climatological stratospheric EP flux and D_F. This agrees qualitatively with ERA-40 (figure 3.2(b)) with predominantly upward flux in the mid-latitude lower and middle stratosphere that turns more equatorward in the upper stratosphere, and which is equatorward

[1]When referring to the SMM experiments, "lower stratosphere" refers to a height range of approximately 16–25 km or pressure range 30–100 hPa, "middle stratosphere" to the approximate ranges 25–40 km and 3–30 hPa, "upper stratosphere" to the approximate ranges 40–48 km and 1–3 hPa and "mesosphere" to levels above about 48 km or 1 hPa.

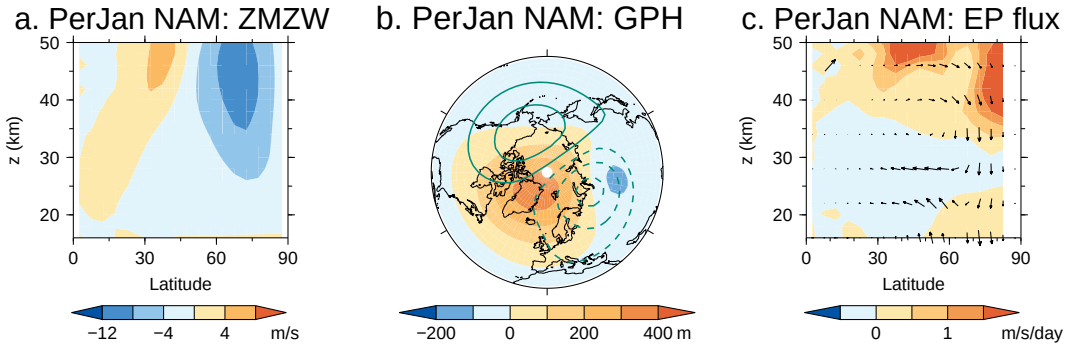

Figure 5.3: The NAM in the perpetual January SMM standard control run. (a) shows the regression of monthly mean ZMZW onto the NAM index. (b) shows the regression of GPH at 32 km north of 20N (colours), with the climatological GPH eddy component plotted as green contours, at 200 m and 600 m (solid) and at −200 m and −600 m (dashed). (c) shows the same for the EP flux (arrows) and D_F (colours). The reference arrow in the top left of the plot represents a flux $(F^\phi, F^z) = (5 \times 10^5, 5 \times 10^3)\,\mathrm{kg/s^2}$. D_F at 87.5N is not plotted as the differentiation error is large there.

at low latitudes. However, the flux is too poleward in the high-latitude lower stratosphere and it is too small, with the upward component being ∼25–40% weaker in the model than in ERA-40 in the lower stratosphere. D_F is represented reasonably in the model, however.

Figure 5.3 shows the regression of monthly mean ZMZW, GPH at 32 km (10 hPa) and EP flux and D_F onto the NAM index in this control run. The index is defined in a similar way to that in chapter 4 and is the leading principal component of monthly mean pressure- and area-weighted 3D GPH between 16–48 km north of 20N. The sign convention is opposite to that of the usual definition so that a more positive NAM index corresponds to weaker high-latitude ZMZW, for easier comparison with the ZMZW responses to torques. The ZMZW NAM signature is a dipole with negative ZMZW north of about 30N at 30 km and positive ZMZW further south, with the positive ZMZW tilting northwards with increasing height (figure 5.3(a)). This is qualitatively similar to that derived from observations [e.g. Kodera, 1995; Kushner, 2010] and to the leading EOF of ZMZW in this control run (not shown).

The NAM signature of 32 km GPH in figure 5.3(b) shows raised heights north of ∼55N in the zonal mean and lowered heights at lower latitudes, with the peak above Greenland, near a node of the climatological eddy pattern. This is fairly similar to the signature derived from ERA-40 in January (figure 4.4), although the GPH signature is less zonally symmetric in the model. The EP flux and D_F NAM signatures (figure 5.3(c)) exhibit poleward and divergent flux above ∼35 km that turns downward near the pole, with equatorward, convergent flux below. The EP flux signature is quite different to that in ERA-40 (figure 4.2).

The time scale of variability may influence the response to a forcing [Leith, 1975], so this was examined using a similar method to that of Baldwin *et al.* [2003b]. A new "single-level NAM" pattern in GPH was defined as the leading EOF of December–February mean GPH at 32 km and the daily 32 km GPH was regressed onto this pattern, with grid points weighted by the cosine of the latitude in the regression, to produce a daily NAM index. Baldwin *et al.* [2003b] defined a time scale of variability as the value of τ that minimises the sum of the squared deviations between the autocorrelation function of the daily NAM index and $\exp(-t/\tau)$, where t is the lag. However, unlike in observations, the autocorrelation function in this control run has a substantial oscillatory component, perhaps because in the normal seasonal cycle the winter is not long enough for oscillations to be discernible. It can be fitted quite well with a function of the form $\cos(\omega t)\exp(-t/\tau')$ (not shown). Varying ω and τ' to minimise the sum of the squared deviations between this and the autocorrelation function up to a lag of 60 days gives a value for τ' of 42 days. This is not easily compared to the data of Baldwin *et al.* [2003b], but it is similar to the January-mean value of 36 days obtained using the same method in the seasonal cycle control run (section 5.1.1.2) which shows reasonable agreement with the observational values when the method of Baldwin *et al.* [2003b] is applied.

5.1.1.2 Seasonal cycle control run

Five ensemble members each of length 45 years were run, each beginning with different initial conditions, and collectively these are called the seasonal cycle control run. Radiative conditions varied with a repeating constant seasonal cycle, and daily mean 100 hPa GPH values from the pre-industrial control run of HadGEM2-CCS were used as the bottom boundary condition, with linear interpolation between the middle of each day. 100 hPa GPH time series from different segments of the HadGEM2-CCS control run were used for each ensemble member so that their winters are independent. The first two years of each ensemble member were not used in the data analysis, so there are 215 years of data. (Several ensemble members were used rather than one long run as the different members could be run in parallel, reducing the computation time.)

Figure 5.4 shows the climatology of the control run with a seasonal cycle. Figure 5.4(a) shows the climatological November–February mean ZMZW and the difference from the ERA-40 climatology. The vortex has a realistic climatological structure with maximum ZMZW near 60–70N. It is slightly too weak in the mid-stratosphere by $\sim$5 ms^{-1}. Figure 5.4(b) shows the standard deviation of the November–February mean ZMZW, showing that the model displays realistic variability in the extratropics and again too little in the tropics due to the absence of the QBO and SAO. Figure 5.4(c) shows a time-height section of the 57.5N ZMZW for one winter in the model control run, showing the vortex strengthening into January followed by an SSW in early March. Figure 5.4(d) shows a similar sequence of events in the 1983–4 winter in ERA-40, indicating the model variability is physically realistic. It was found by inspection that the 57.5N ZMZW evolution in most winters of the first control ensemble member has a reasonable analogue in ERA-40. Overall, there are 7.5 major SSWs per decade according to the criterion of Charlton and Polvani [2007], which is similar to the frequency of about 6 per decade in ERA-40, and the fraction of SSWs falling in each month November–March is also similar to that in ERA-40 (not shown). Figure 5.4(e) shows the seasonal cycle of the ZMZW at (60N, 32 km)

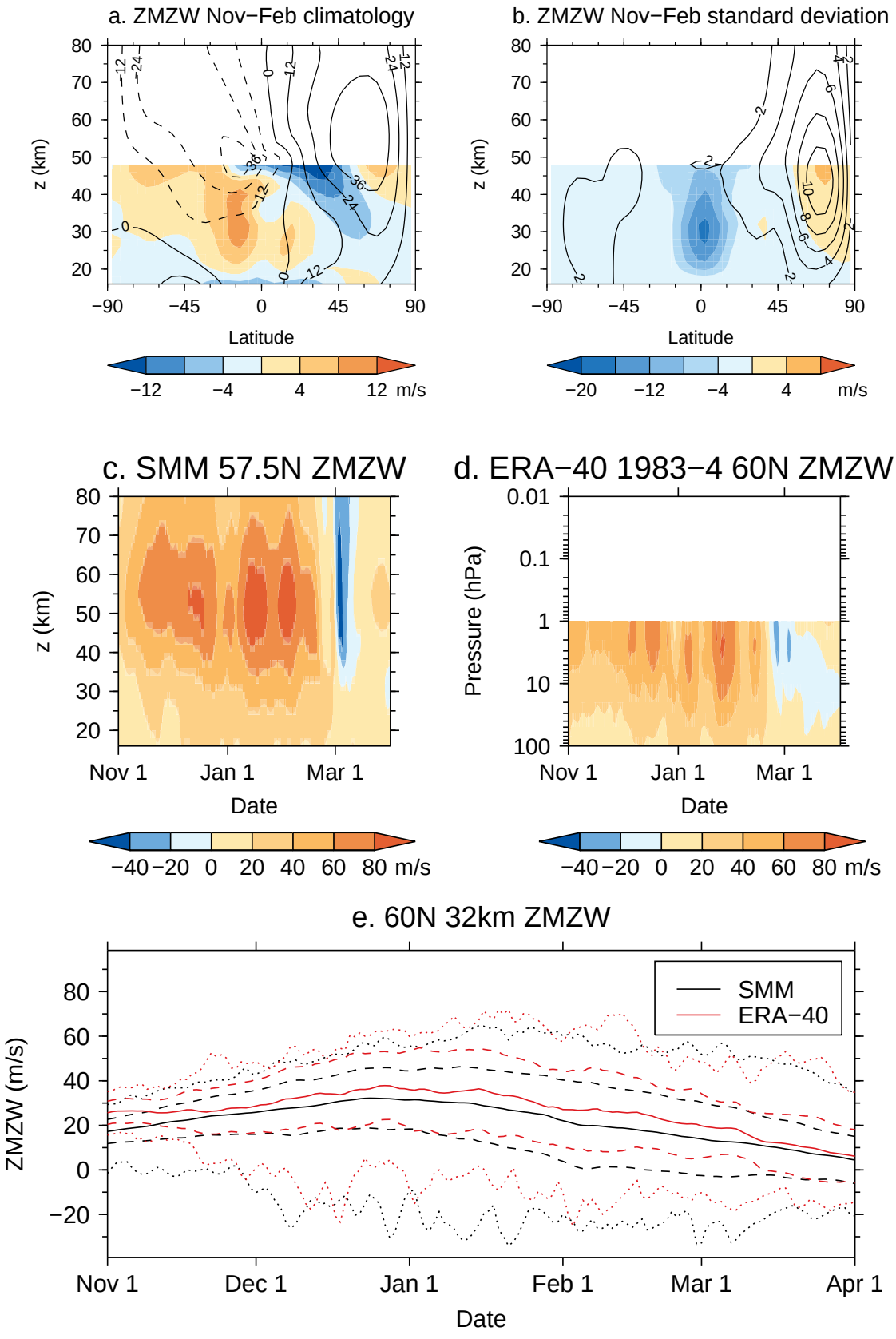

Figure 5.4: (Continued on the following page.)

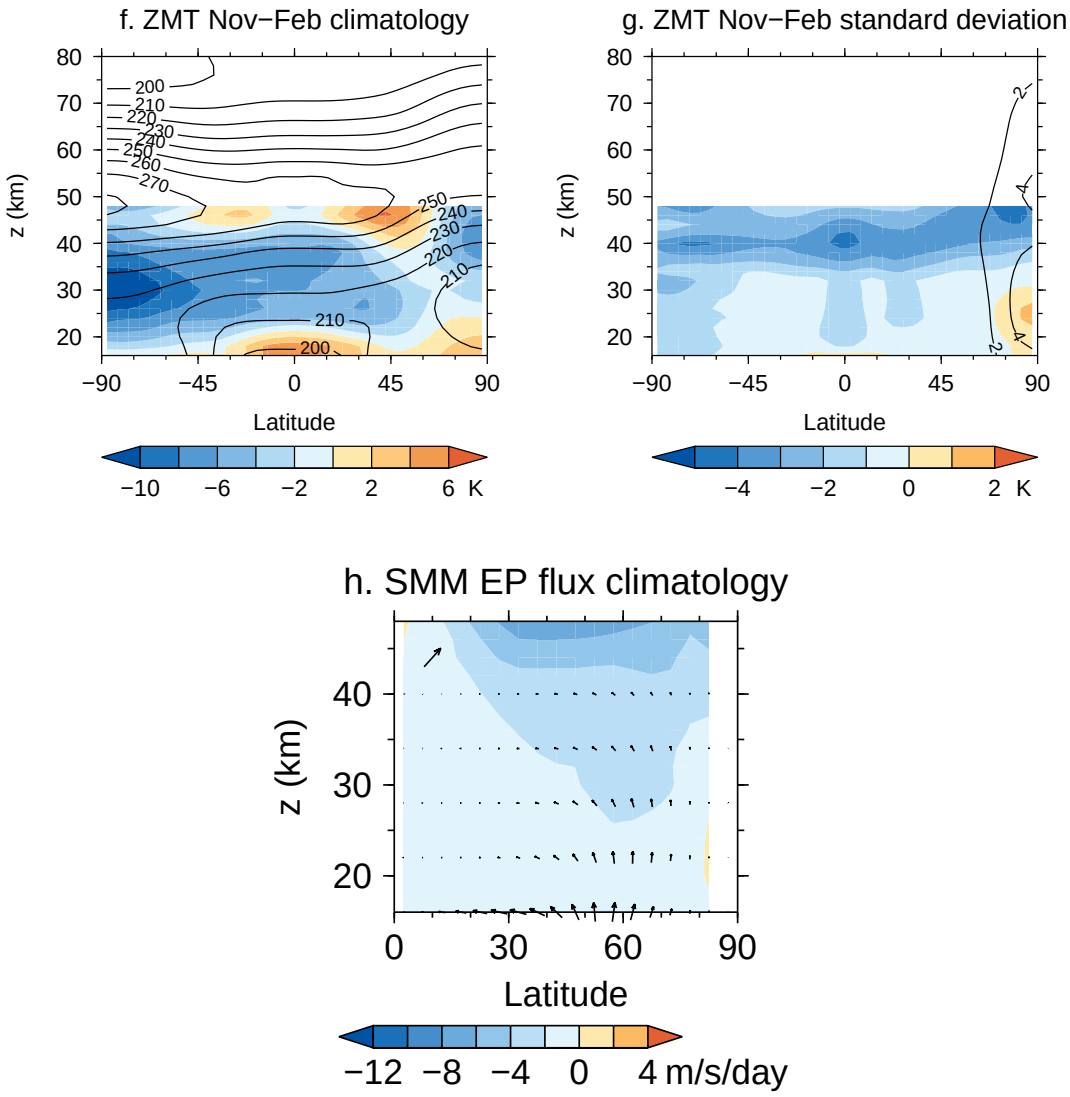

Figure 5.4: (Continued.) The SMM standard control run with a seasonal cycle. (a) shows the November–February mean ZMZW (contours) and the difference from the ERA-40 November–February climatology (colours) and (b) shows the same for the ZMZW standard deviation. (c) shows a time-height section of the ZMZW at 57.5N over November–March of one winter. (d) shows the same for the 60N ZMZW in ERA-40 in the 1983–4 winter for comparison. (e) shows the climatological seasonal cycle of ZMZW at (60N, 32 km) in the SMM (black solid line) and the same for the (60N, 10 hPa) ZMZW in ERA-40 (red solid line). Dashed lines show the mean ZMZW plus or minus one standard deviation in the SMM (black) and ERA-40 (red). Dotted lines show the extreme values in the first 45 years of SMM data and in ERA-40. (f) and (g) show the same as (a) and (b) but for November–February mean ZMT. (h) shows the November–February mean EP flux (arrows, shown every 6 km in height) and D_F (colours) below 50 km. The reference arrow in the top left corner has a magnitude of $(10^7, 10^5)$ kg/s^2. D_F at 87.5N is not plotted as the differentiation error is large there.

in the control run, along with its standard deviation and its range in the first 45 years of model data, alongside that at (60N, 10 hPa) in ERA-40. This indicates that the model seasonal cycle in the mid-stratosphere has a reasonable time evolution and variability but the ZMZW is too weak by $\sim 5\,\mathrm{ms}^{-1}$.

Figure 5.4(f) shows the ZMT climatology and (g) shows its standard deviation, indicating the model stratosphere is too cold on average by several Kelvin and its variability in the upper stratosphere is too small (though ERA-40 is also less trustworthy in this region).

Figure 5.4(h) shows the climatological November–February mean EP flux and D_F in the seasonal cycle control run. The EP flux is in better agreement with that in ERA-40 (figure 3.2(b)) than that in the perpetual January control run, with the flux being only weakly poleward at high latitudes. D_F is in reasonable agreement with that in ERA-40, but again the EP flux is too weak and the upward component is similar to that in the perpetual January control run.

The NAM index is defined in a similar way to that in the perpetual January control run, based on the time series of November–February mean GPH. Figure 5.5(a) shows the regression of November–February mean ZMZW onto the NAM index and 5.5(b) shows that of the EP flux and D_F. The ZMZW NAM signature is similar to that in the perpetual January control run, although the positive ZMZW south of ∼45N is not as prominent. This suggests that the NAM signature is not strongly dependent on the zonal mean climatology or the time-dependence of the stationary wave forcing. The EP flux signature is strongly upward at high latitudes and turns equatorward with increasing height, and is generally convergent – these features are quite consistent with those in the ERA-40 NAM signature (figure 4.2).

The time scale of variability τ was calculated for the 32 km daily NAM index as described in section 5.1.1.1. The monthly mean value of τ decreases from 20 days in November to 15 days in March and this is in good agreement with observations [Baldwin *et al.*, 2003b]. At 20 km the time scale is shorter than that in observations by about a week in November at 21 days, about right in December at 28 days and longer by 1–2 weeks in January–March at ∼30–40 days. Therefore the time scales tend to

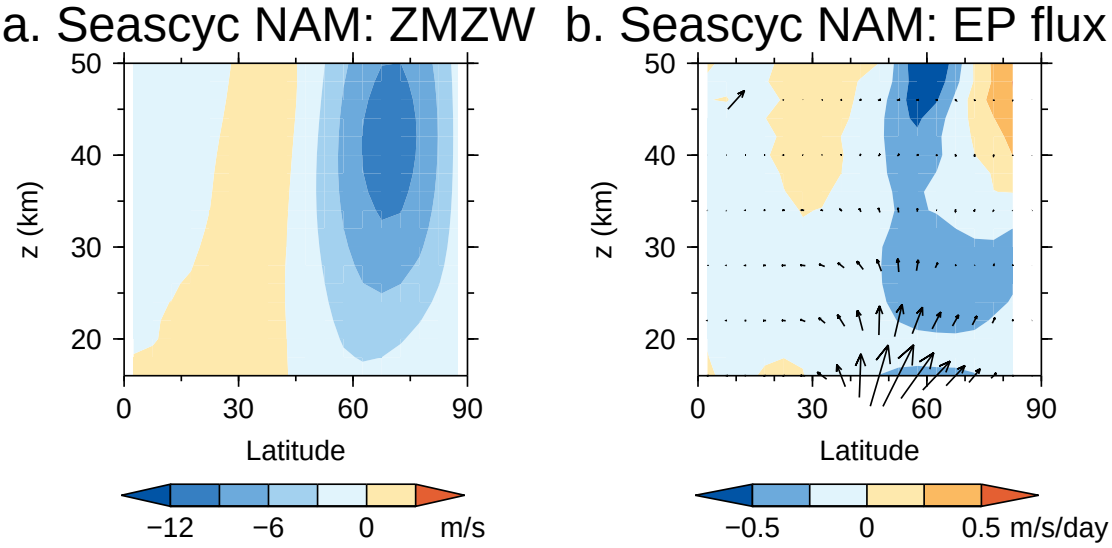

Figure 5.5: The NAM in the SMM control run with a seasonal cycle. (a) shows the regression of November–February mean ZMZW onto the NAM index. (b) shows the same for the EP flux (arrows) and D_F (colours). D_F at 87.5N is not plotted as the differentiation error is large there. The reference arrow in the top left represents a flux $(F^\phi, F^z) = (3 \times 10^5, 3 \times 10^3)\,\mathrm{kg/s^2}$.

be too long in the lower stratosphere in late winter but are not wholly unrealistic and the time scale is about right in the mid-stratosphere.

Overall the extratropical climatology of the SMM control ensemble members with a seasonal cycle seems qualitatively and quantitatively reasonable, with differences from ERA-40 that are comparable to those in modern sophisticated GCMs (e.g. HadGEM2-CCS [Osprey *et al.*, 2013]).

5.1.2 Runs with applied torques

A term of the form

$$f_0(\phi_0, \delta\phi, z_0, \delta z) \exp\left(-\left[\frac{\phi - \phi_0}{\delta\phi}\right]^2 - \left[\frac{z - z_0}{\delta z}\right]^2\right)$$

was added to the zonal momentum equation, where ϕ is the latitude, z is the height, $(\phi_0,\, z_0)$ is the position of the torque maximum and $\delta\phi$ and δz set the meridional and vertical scales of the applied torque respectively. f_0 is a constant chosen such that, except in experiments designed to test the effect of varying the torque magnitude,

the total zonal momentum added is the same for each choice of the other constants, chosen to be that of a torque with $\phi_0 = 60$N, $z_0 = 30$ km, $\delta\phi = 11°$, $\delta z = 10$ km and $f_0 = 2.5$ m/s/day, which will be referred to as the "standard torque". This torque strength was found to produce responses in quantities such as ZMZW of the order of one standard deviation in the control runs, which is comparable to the magnitude of the influence of natural forcings such as the QBO on the vortex in the real atmosphere. These $\delta\phi$ and δz values are the same as those used by Ring and Plumb [2007] and Song and Robinson [2004] respectively. A torque with oppositely signed f_0 is also applied in the opposite hemisphere (with the sign of ϕ_0 reversed), so no net zonal momentum is added globally. (Excluding this was not found to affect the results substantially. In appendix 5.B it is shown that moving the oppositely signed torque to the NH mesosphere does not affect the conclusions of this chapter.)

A variety of experiments were performed, where normally only one of the torque parameters is different from that of the standard torque, to test the dependence of the response to varying each individual parameter, though a few experiments have more than one parameter varied as explained in the text. In most experiments one of five other (ϕ_0, z_0) values in the extratropical stratosphere were used, to test the effect of changing the location of the torque maximum: (30N, 30 km), (45N, 30 km), (75N, 30 km), (60N, 20 km) and (60N, 40 km). Also one of two other values of $\delta\phi$ (5° and 20°), and of δz (2.5 km and 40 km), and four other peak torque magnitudes (-2.5, 1, 5 and 7.5 m/s/day), were usually used.

To examine the steady-state response to an applied torque under perpetual January conditions (section 5.2), model runs were performed that are similar to the standard perpetual January control run, but with a torque applied that does not vary in time.

To examine the wintertime response to an applied torque in runs with a seasonal cycle (section 5.3), model runs similar to the standard control run with a seasonal cycle were performed with a torque applied from October 22, from when it was linearly increased to reach full strength on Nov 1, and then was kept constant until the end of March, from when it was linearly decreased to zero on April 10. The runs with

applied torques with different ϕ_0 and z_0 values and other parameters kept the same as those for the standard torque had five ensemble members each. Torques with other parameter values changed were tested using only one ensemble member compared against the same ensemble member of the control run. Examining the response to torques with different ϕ_0 and z_0 values using only one ensemble member does not change the results substantially, though the statistical significance of the response is lower.

5.2 Perpetual January run responses

5.2.1 Zonal mean zonal wind response

Figure 5.6 shows the mean ZMZW response to each applied torque under perpetual January conditions, defined as the difference between the mean ZMZW in runs with applied torques and that in the standard control run. The anomaly correlation[2] between the response and the NAM signature (figure 5.3(a)) is also written on each panel. Apart from the responses to the (ϕ_0=30N, z_0=30 km) and (ϕ_0=45N, z_0=30 km) torques (figure 5.6(a)), for which the responses are small, the ZMZW responses all share many characteristics with the NAM signature, with the responses to positive torques showing decreased ZMZW at high latitudes and increases at lower latitudes, with the region of positive response sloping poleward with increasing height. The anomaly correlations of the ZMZW responses with the NAM signature north of 20N are 0.9 or greater in most cases. Many responses have the same sign as the torques just below where the torques peak, however, which is a departure from the NAM signature. This is particularly noticeable in the responses to torques with (ϕ_0=75N, z_0=30 km) and with (ϕ_0=60N, z_0=30 km) and peak torque 7.5 m/s/day, which consequently have smaller anomaly correlations with the NAM signature than

[2]The anomaly correlation is defined in the same way as in section 4.1, but with uniform weighting of the data.

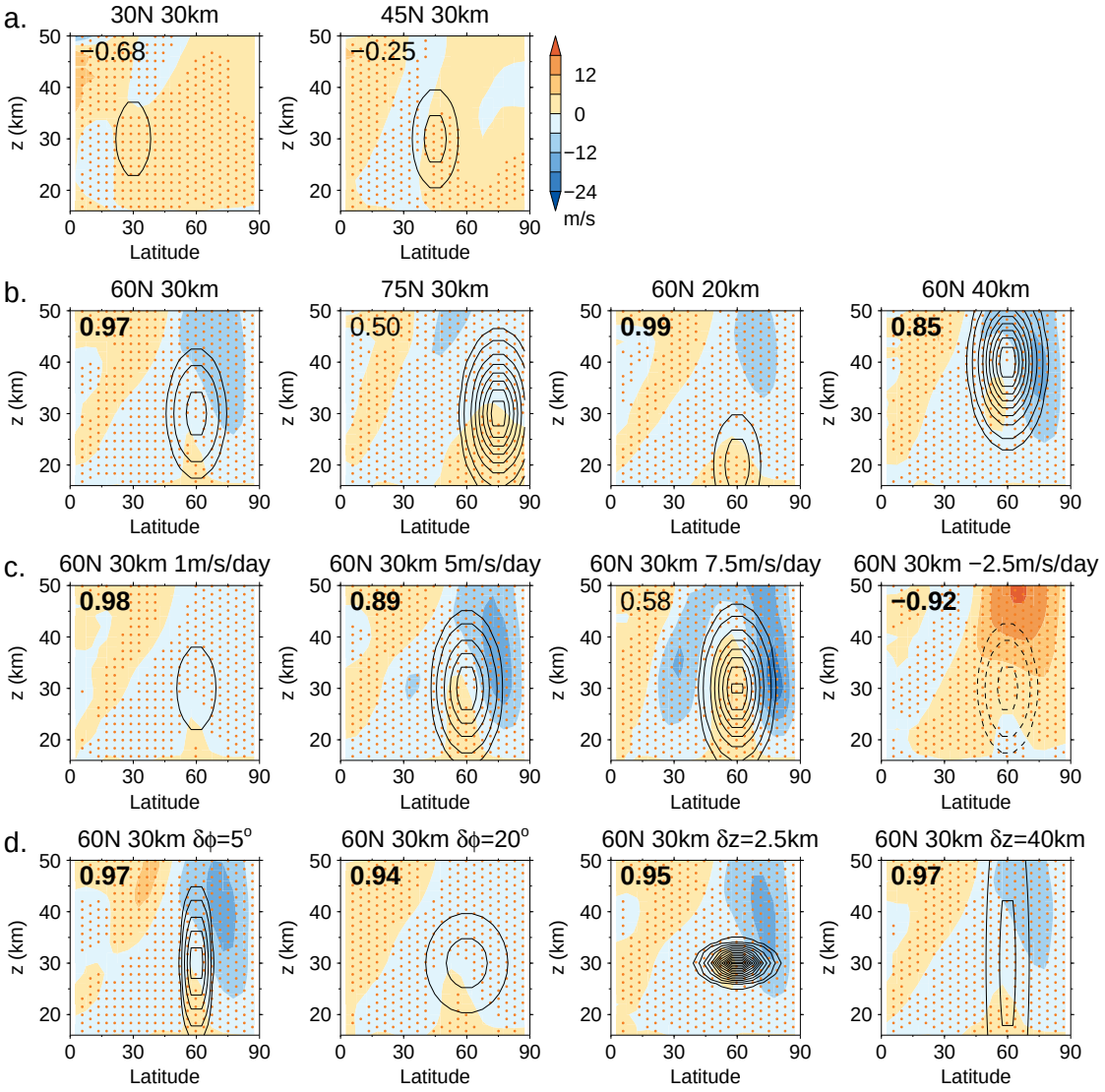

Figure 5.6: Steady state ZMZW responses to applied torques under perpetual January conditions (colours). (a) shows the response to torques centred in the mid-latitude mid-stratosphere, (b) the response to torques centred in the high-latitude stratosphere, (c) the response to torques centred at (60N, 30 km) with different peak strengths and (d) the response to torques with different meridional and vertical scales. A NAM-like response is displayed in many cases in (b)–(d). The number in the top left corner of each panel is the anomaly correlation north of 20N between the ZMZW response and its NAM signature in the standard control run (figure 5.3(a)). Correlations larger than 0.85 are written in bold font. The applied torque is shown by contours at ±0.5 m/s/day, then every integer value in units m/s/day, with negative contours dashed. Plot titles indicate the latitude ϕ_0 and height z_0 where the torque is strongest. Unless otherwise specified in the title, the torques have meridional scale $\delta\phi = 11°$, vertical scale $\delta z = 10$ km, and a strength such that the magnitude of the total zonal momentum added to the NH equals that of the standard (ϕ_0=60N, z_0=30 km) torque with peak strength 2.5 m/s/day. Stippling shows where the responses are statistically significant above the 95% level.

other high-latitude torques, although there is still a reasonable qualitative resemblance at higher altitudes. The response is stronger for torques with smaller meridional and vertical scales.

It is interesting to note that applying a *positive* torque at high latitudes results in a *decrease* in the ZMZW in most of the high-latitude stratosphere. This shows that non-linear feedbacks play an important role in bringing about this response.

Stippling in figure 5.6 shows where the data are statistically significant above the 95% level. This was calculated according to the null hypothesis that 6-month averages of the data have the same distribution in the runs with applied torques and the control run. The probability that the magnitude of the responses would exceed that in the data under this null hypothesis was calculated at each grid point using a Monte Carlo (MC) permutation test, by creating 1000 surrogate control and perturbed run time series of the same length as those in the actual data, by randomly combining 6-month averages from both the control and perturbed run data in each surrogate and finding the distribution of differences between these. Taking 6-month averages accounts for serial autocorrelation of the data – the envelope of the autocorrelation function of ZMZW at points in the extratropical stratosphere is about 0.2 or less at a lag of 6 months (not shown) – and the calculated statistical significance is not very different if 12-month averages are used instead. Statistical significances of the GPH, EP flux and residual circulation responses presented below were calculated in a similar way.

The response to the standard (ϕ_0=60N, z_0=30 km) torque is not sensitive to the precise winter conditions. The response is similar in runs with perpetual November, December and February radiative conditions and bottom boundary, and in perpetual January conditions with the eddy component of the prescribed 16 km GPH equal to 1.5 times that of the HadGEM2-CCS January climatology rather than two times (which causes the control run to exhibit periodic rather than irregular vacillations).

It is possible that the Rayleigh friction in the model affects its simulated response to a torque [Shepherd *et al.*, 1996; Shepherd and Shaw, 2004]. In appendix 5.A it is shown that removing Rayleigh friction in the NH extratropics does not change the qualitative nature of the response that is of interest here.

There is some evidence of non-linearity in the ZMZW response to the torques. The response to torques with (ϕ_0=60N, z_0=30 km) is quite linear with respect to varying the peak torque amplitude in the range 1–5 m/s/day and to changing its sign, but the response to the 7.5 m/s/day torque is more positive around (60N, 30 km) and more negative around (30N, 35 km). The size of the response to the −2.5 m/s/day torque is also about 1.5 times that to the 2.5 m/s/day torque in the high latitude upper stratosphere. The NAM signature in the run with the torque that peaks at 7.5 m/s/day is quite different from that in the standard control run and shows a better qualitative resemblance to the ZMZW response to this torque, with positive ZMZW between ∼40–60N and ∼16–50 km (not shown). This indicates that the NAM signature itself depends on the applied torque, so that agreement between the ZMZW response and the NAM signature may be better than is indicated by the anomaly correlations in figure 5.6. This makes agreement between the ZMZW response and the NAM signature more difficult to quantify. In runs with the other torques, the NAM signature is similar to that in the standard control run.

The ZMZW responses to the (ϕ_0=75N, z_0=30 km) and (ϕ_0=60N, z_0=40 km) torques may be less like the NAM signature than the response to the standard (ϕ_0=60N, z_0=30 km) torque because the peak torque values are quite large in these cases, peaking around 7.5–9 m/s/day – the responses to the (ϕ_0=75N, z_0=30 km) and (ϕ_0=60N, z_0=40 km) torques have a higher correlation with the NAM signature if the torque strength is reduced by 60% (not shown). The torque with (ϕ_0=60N, z_0=30 km) and vertical scale 2.5 km has a peak of over 13 m/s/day and the ZMZW response is very NAM-like, however, so it is not simply the case that responses to torques with greater peak accelerations are less NAM-like.

Ring and Plumb [2007] showed that the magnitude of the annular mode part of the ZMZW response to an applied torque in a tropospheric primitive equation model is linearly proportional to the projection of the torque onto the leading EOF of ZMZW (figure 5.7). This relationship was tested in the SMM experiments and figure 5.8 shows for each perpetual January SMM experiment, on the y-axis, the inner product of the zonal mean zonal momentum (ZMZM) response and the NAM ZMZM

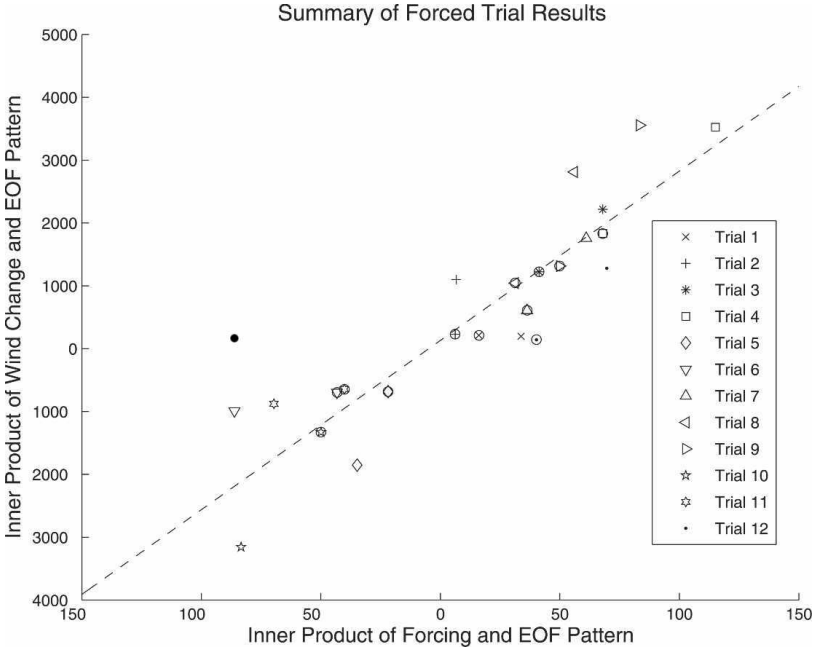

Figure 5.7: The inner product of the zonal wind change and EOF pattern plotted against the inner product of the forcing and EOF pattern in a tropospheric primitive equation model, from Ring and Plumb [2007]. The dashed line is the best least squares linear fit.

signature, plotted against, on the x-axis, the inner product of the ZMZM added by the torque at each time step with the NAM ZMZM signature. Since the response to an applied torque would be expected to scale more closely with the amount of ZMZM being added rather than the ZMZW perturbation it seems appropriate to do the calculation using ZMZM. (The results are qualitatively similar if the calculation is performed for ZMZW.) The inner product of a 2D field $x_{i,j}$ with the NAM signature $\mathrm{NAM}_{i,j}$ is defined as

$$\frac{\sum_{i,j} x_{i,j}\mathrm{NAM}_{i,j}}{\sum_{i,j} \mathrm{NAM}_{i,j}^2},$$

where the summations run over all grid-points north of 20N.

Unlike the findings of Ring and Plumb [2007], the ZMZM response projection onto the NAM signature has a negative correlation with the projection of the torque, equal to -0.67. Note, however, that the inner product of the ZMZM response and NAM signature does not always reflect the degree to which the response is NAM-like – for example, the response to the (ϕ_0=30N, z_0=30 km) torque has a large inner product

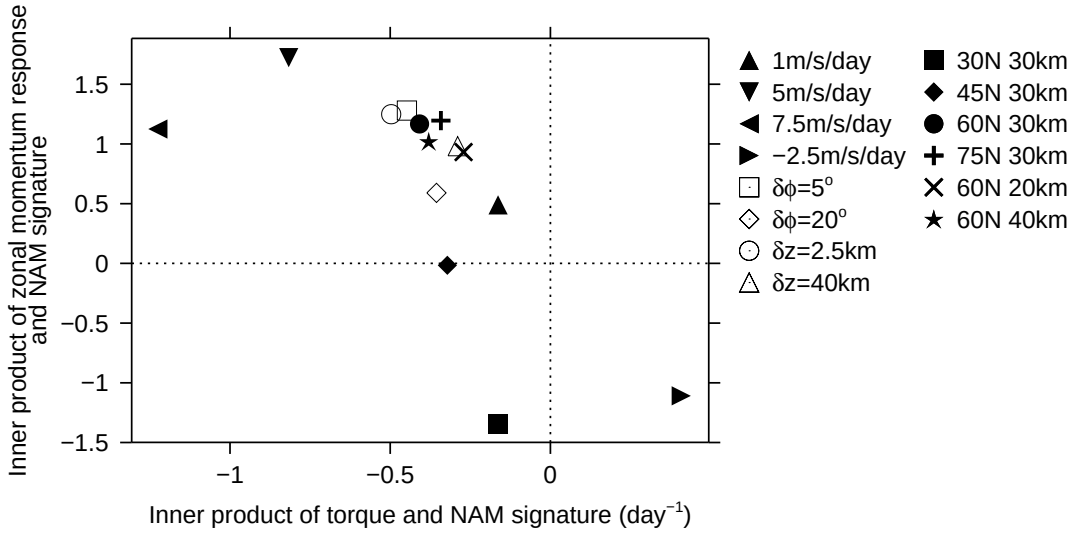

Figure 5.8: The inner product of the NAM zonal mean zonal momentum (ZMZM) signature and the steady state ZMZM responses under perpetual January conditions, plotted against the inner product of the NAM ZMZM signature and the torque. Unless otherwise specified in the legend, the torques peak at ($\phi_0 = 60$N, $z_0 = 30$km), have meridional scale $\delta\phi = 11°$ and vertical scale $\delta z = 10$ km and have a magnitude such that the total zonal momentum added to the NH equals that of the standard (ϕ_0=60N, z_0=30 km) torque with peak magnitude 2.5 m/s/day. A ZMZM response equal to the NAM signature would have a value of 1 on the vertical axis and a torque equal to the NAM ZMZM signature per day would have a value of 1/day on the horizontal axis.

with the NAM ZMZM signature even though the ZMZW response is not NAM-like (figure 5.6). The torques tested here sample from a wider range of positions, magnitudes and structures than those tested by Ring and Plumb [2007], which gives greater opportunity for non-linear behaviour to be exhibited.

5.2.2 Response of wave components

Figure 5.9(a) shows the mean response in GPH at 32 km to the standard (ϕ_0=60N, z_0=30 km) torque. The GPH response is again very like the NAM signature (figure 5.3(b)) with an anomaly correlation of 0.86, calculated in the same way as those in section 4.1. The anomaly correlation at 50 km is similar at 0.89 (not shown).

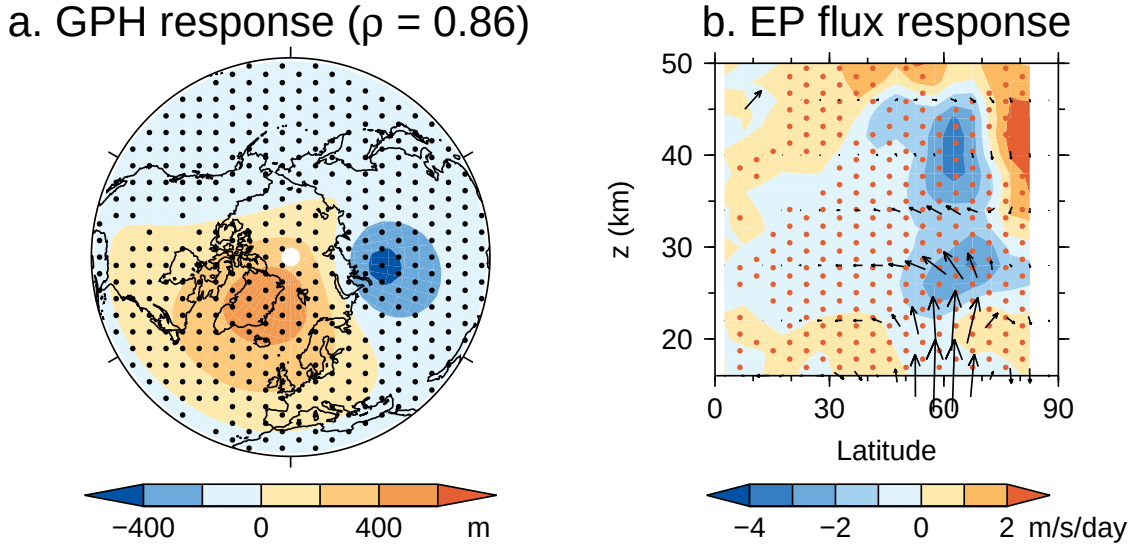

Figure 5.9: (a) The steady state 32 km GPH response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions, shown north of 20N. The ρ value in the title indicates the anomaly correlation with the NAM GPH signature north of 20N. (b) The steady state EP flux (arrows) and D_F (colours) responses to the same torque. Stippling shows where the GPH and D_F responses are statistically significant above the 95% level and the EP flux response in (b) is only plotted where either the F^ϕ or F^z response is statistically significant above the 95% level. The arrow in the top left corner of (b) has a length a flux $(10^6, 10^4)$ kg/s^2 would have. EP fluxes are plotted every 6 km in height and D_F data at 87.5N are not plotted, as differentiation error causes them to be spuriously large.

Figure 5.9(b) shows the mean response in the EP flux and D_F to the standard (ϕ_0=60N, z_0=30 km) torque. The D_F response is large and negative between about 50–70N. This is associated with an EP flux response that is predominantly more upward in the lower stratosphere in this latitude range, and which turns equatorward in the middle stratosphere. The negative D_F response in the region of (60N, 30 km) effectively cancels the direct effect of the torque on the zonal mean flow. (Note that in the case of a torque resulting from an EP flux anomaly directly caused by an external forcing, the measured EP flux and D_F response would be the sum of the direct forcing and the feedbacks, so this negative D_F in the mid-stratosphere would be less prominent.) These features are qualitatively similar for responses to the other torques centred at 60N or further polewards (not shown), with the EP flux being

more upward around 50–70N and 16 km and convergent where the torque strength peaks, so that the planetary wave response largely cancels the direct effect of the torque on the zonal mean circulation. The responses to torques centred at 60N all show increased convergence in the high-latitude upper stratosphere as well. This is not apparent in the response to the (ϕ_0=75N, z_0=30 km) torque, for which the EP flux response turns poleward in the lower stratosphere to become convergent near (75N, 30 km) (not shown).

The EP flux response resembles the NAM signature (figure 5.3(c)) only in some respects. Both exhibit a convergent equatorward EP flux in the mid-latitude stratosphere and a divergent poleward flux in the extratropics near 50 km. However, the NAM signature does not display the convergent upward EP flux in the high-latitude stratosphere shown in the response.

In appendix 5.B, it is shown that moving the oppositely-signed torque, that cancels the addition to the global momentum made by the torque in the NH stratosphere, from the SH to the NH mesosphere alters the EP flux response (but not the qualitative nature of the D_F response). Therefore this response depends on the way momentum is conserved. Responses of the other variables presented here are not strongly affected by this change, however.

Figure 5.10 shows the mean response to the standard (ϕ_0=60N, z_0=30 km) torque of the residual circulation $(\overline{v}^*, \overline{w}^*)$ and the associated acceleration term in equation 2.13, given by

$$-\overline{v}^* \left[(a\cos\phi)^{-1}(\overline{u}\cos\phi)_\phi - f\right] - \overline{w}^*\overline{u}_z.$$

The response of this acceleration term is generally equal and opposite to the sum of the torque and the D_F response, as expected since the mean $\partial\overline{u}/\partial t$ response is small (and the Rayleigh friction is small below 50 km). In the lower and middle stratosphere, the circulation response is equatorward between ∼16–22 km, mostly poleward just above and downward near the pole between ∼20–40 km. This is not associated with a large acceleration response as the applied torque and the D_F response largely cancel in the middle stratosphere as noted previously. There is a poleward circulation response in

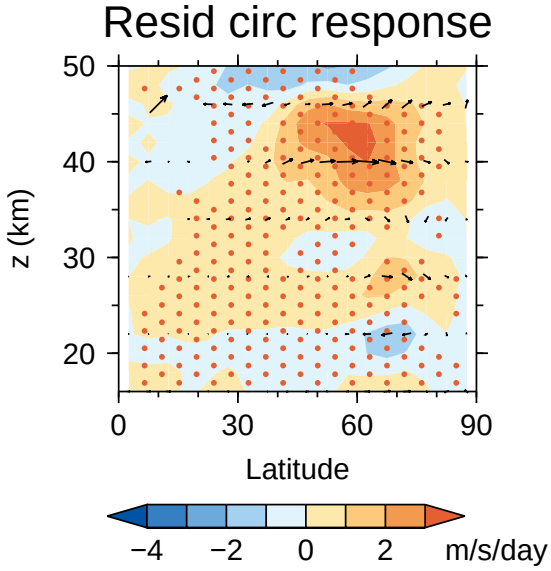

Figure 5.10: The steady state residual circulation (arrows) and associated zonal acceleration (colours) responses to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions. The arrow in the top left corner has a length (0.2, 0.002) ms^{-1}. The residual circulation response is only plotted where either the v^* or w^* responses are statistically significant above the 95% level and stippling shows where the associated zonal acceleration response is statistically significant above the 95% level. The residual circulation is plotted every 6 km in height.

the upper stratosphere and an upward response near the pole above ∼40 km. This residual circulation response is very different to what would be predicted if D_F did not respond to the applied torque (section 2.2.3).

5.3 Seasonal cycle run responses

5.3.1 Zonal mean zonal wind response

In this section the November–February mean response to an applied torque in runs with a seasonal cycle is investigated. Figure 5.11 shows the composite ZMZW responses to torques centred at different positions in the NH stratosphere between November and February, defined as composite differences between runs with the applied torques and the standard control run. The responses to the (ϕ_0=60N, z_0=30 km), (ϕ_0=75N, z_0=30 km) and (ϕ_0=60N, z_0=20 km) torques are qualitatively similar, be-

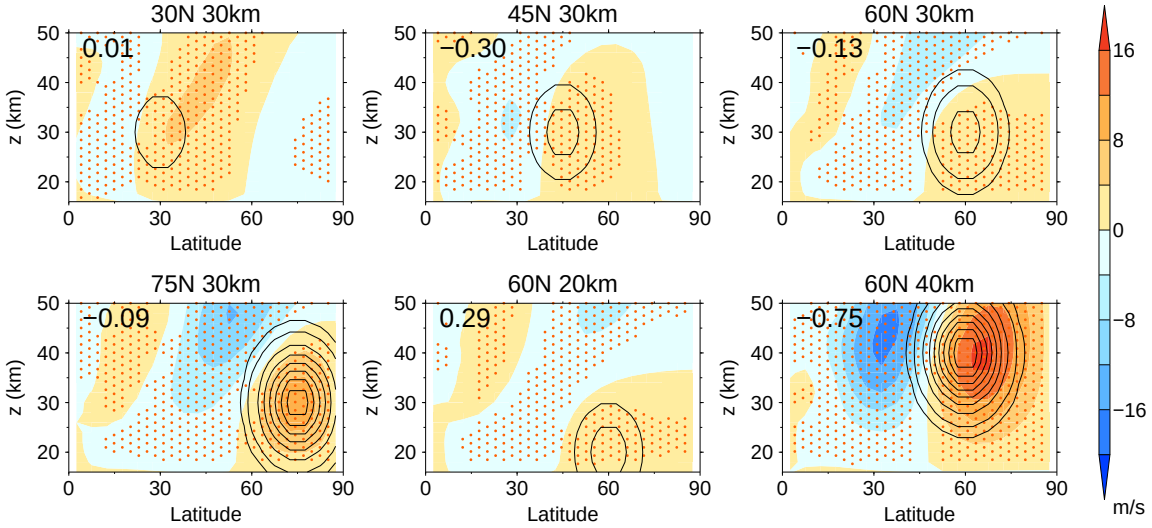

Figure 5.11: The November–February mean ZMZW responses to applied torques centred at latitude ϕ_0 and height z_0, as indicated in the plot titles, in runs with a seasonal cycle (colours). The numbers in the top left corners of each panel are the anomaly correlations north of 20N between the ZMZW responses and the ZMZW NAM signature in the standard control run (figure 5.3(a)), which are generally not large. The applied torque is shown by contours at ± 0.5 m/s/day, then every integer value in units m/s/day. The torques have meridional scale $\delta\phi = 11°$, vertical scale $\delta z = 10$ km, and a magnitude such that the total zonal momentum added to the NH equals that of the standard (ϕ_0=60N, z_0=30 km) torque with peak magnitude 2.5 m/s/day. Stippling shows where the responses are statistically significant above the 95% level.

ing positive in the high latitude lower and middle stratosphere, negative above and equatorward of this region and positive in the sub-tropical upper stratosphere.

The ZMZW response to the (ϕ_0=60N, z_0=40 km) torque remains positive in the upper stratosphere north of ∼60N and does not show much sign of a positive ZMZW response at low latitudes. The responses to the (ϕ_0=30N, z_0=30 km) and (ϕ_0=45N, z_0=30 km) torques are positive near where the torques peak and negative to the north and south.

Statistical significance of the responses is calculated in a similar way to that of the perpetual January runs, using an MC permutation test according to the null hypothesis that data in different years in the control runs and runs with torques are independent and drawn from the same distribution.

The responses to all torques with (ϕ_0=60N, z_0=30 km) and different torque magnitudes or meridional or vertical scales as described in section 5.1.2 (not shown) are very qualitatively similar to that of the standard (ϕ_0=60N, z_0=30 km) torque. The response is quite linear with respect to varying the torque magnitude between 1–7.5 m/s/day, and larger for torques with $\delta\phi = 5°$ and $\delta z = 2.5$ km than those with larger meridional and vertical scales, similar to the results for the perpetual January runs.

The ZMZW responses shown in figure 5.11 are not very like the NAM ZMZW signature (figure 5.5(a)). The response to the (ϕ_0=60N, z_0=40 km) torque appears somewhat like the negative NAM signature, though this may simply be due to the strong direct influence of the torque rather than feedbacks.

5.3.2 Response of wave components

Figure 5.12 shows the November–February mean response of the EP flux and D_F to the standard (ϕ_0=60N, z_0=30 km) torque. In the middle and lower stratosphere, the EP flux response to the torque is upward between ∼50–75N and turns equatorward with increasing height, with greater convergence where the torque is strongest, in common with the perpetual January experiments. This response is qualitatively similar to the NAM signature (figure 5.5(b)), except that the lower stratospheric EP flux response is not so poleward and there is greater convergence in the region where the torque is strongest.

The EP flux responses to the other torques centred at 45N or poleward are generally qualitatively similar, and that to the (ϕ_0=30N, z_0=30 km) torque differs in being divergent and more downward at high latitudes (not shown).

The GPH responses to the torques (not shown) are qualitatively consistent with the ZMZW changes and geostrophic balance, and are not NAM-like overall.

Again it is shown in appendix 5.A that greatly weakening Rayleigh friction in the NH extratropics does not qualitatively change these results.

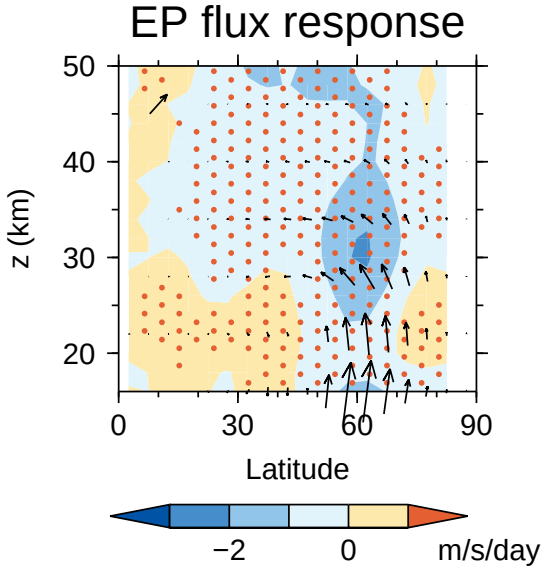

Figure 5.12: The November–February mean wintertime response of EP flux (arrows) and D_F (colours) to the standard (ϕ_0=60N, z_0=30 km) torque in runs with a seasonal cycle. The arrow in the top left corner represents a flux $(10^6, 10^4)$ kg/s^2. The EP flux response is only plotted where either the F^ϕ or F^z responses are statistically significant above the 95% level, and stippling shows where the D_F response is statistically significant above the 95% level. EP fluxes are plotted every 6 km in height and D_F data at 87.5N are not plotted, as differentiation error there causes them to be spuriously large.

5.4 Understanding the differences between the perpetual January and seasonal cycle simulations

To understand better the reasons why the circulation response is more NAM-like in the perpetual January runs than in those with a seasonal cycle, two experiments were done to test the sensitivity of the responses to torques to changing the boundary conditions. In the first experiment, one control ensemble member was produced with the same seasonal cycle in radiation as in the standard seasonal cycle control run but with the 100 hPa GPH seasonal cycle specified as the monthly mean climatology of HadGEM2-CCS, with twice the eddy component, with mid-month GPH having the value of the monthly mean and with the GPH at each time step given by linear interpolation between the mid-month values. This was to test the effect of removing the high-frequency wave forcing and increasing the wave amplitude. A similar

run with the standard (ϕ_0=60N, z_0=30 km) torque was also produced. The ZMZW climatology and response to the torque were quite similar to those under perpetual January conditions (not shown), so the presence of seasonally-varying radiation does not prevent the response from being NAM-like.

In the second experiment, the daily-mean 16 km GPH was specified as in the standard control run with a seasonal cycle, but with eddy component multiplied by a factor of 1.5 to test the effect of increasing the wave amplitude alone. The ZMZW response to the standard (ϕ_0=60N, z_0=30 km) torque in this case (not shown) is similar to that in the seasonal cycle experiments shown in section 5.3. (Recall from section 5.2.1 that the response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions is NAM-like when using an eddy multiplication factor of 1.5. Using an eddy amplification factor greater than 1.5 with a seasonal cycle makes the SMM unstable.) Therefore the difference between the responses to torques in seasonal cycle and perpetual January simulations is not simply due to the mean wave amplitude being larger in the latter. These results indicate that the differences between the responses to torques in simulations with perpetual January conditions and with a seasonal cycle arise due to the high-frequency variability in the specified 16 km GPH in the seasonal cycle runs, either directly or indirectly due to its effects on the control run climatology, rather than being due to the presence of the seasonal cycle or to the smaller mean wave amplitudes at 16 km.

In order to see if the responses would become more NAM-like if the torque were switched off in mid-winter in runs with a seasonal cycle, three ensemble members were run with a seasonal cycle with the standard (ϕ_0=60N, z_0=30 km) torque set at full strength up to the end of December, then linearly decreased to zero in the first 10 days of January, but this also did not lead to a NAM-like response (not shown).

The transient responses to applied torques (examined in chapter 6) show the response in the perpetual January runs becomes NAM-like after about one month, so it is not the case that the responses are similar in the runs with a seasonal cycle, with it taking longer than one winter for the response to become NAM-like.

Son and Lee [2006] examined the response of a tropospheric primitive equation model to heating perturbations and found the response to be more AM-like in situations when the leading EOF dominates the variability more, so that the ratio of the variance of circulation statistics accounted for by the second EOF to that of the first EOF is smaller. In the standard perpetual January control run, this ratio for the variance in the mass-weighted GPH north of 20N is 0.39, and in individual months in the standard seasonal cycle control run it is 0.7 in November and ∼0.4 in December–February. Therefore, as the ratio only differs substantially in November of the seasonal cycle control run, the difference in the response to a torque in the two cases does not seem to be due to the NAM being a more dominant mode of variability in the perpetual January runs.

Therefore the only factor causing the differences between the perpetual January runs and those with a seasonal cycle seems to be the high-frequency variability in the 16 km GPH in the latter, which may affect the response directly or indirectly through changing the climatology.

5.5 Summary and discussion

The results presented here show that the long-term mean response of the zonal mean flow in the stratosphere to a fixed torque in the SMM is strongly influenced by feedbacks from the response of the large-scale wave part of the flow. This indicates that understanding such feedbacks is very important for understanding the extratropical stratospheric response to external forcings, on top of understanding the processes involved in the direct influence of these forcings. In the perpetual January runs, the feedback is strong enough to produce a steady-state ZMZW response to torques placed at high latitudes that is opposite in sign to the torque, and which overall appears very NAM-like (figure 5.6). This behaviour is qualitatively similar to that of tropospheric models that often show an AM-like response to applied torques or heating perturbations [Son and Lee, 2006; Ring and Plumb, 2007, 2008], except for the sign of the response. My results show that the response is quite linear with respect to

changing the torque magnitude in the range 1–5 m/s/day or reversing the sign of the torque, and the response is qualitatively similar if the meridional and vertical scales of the torque are varied substantially.

However, the ZMZW is not NAM-like in simulations with a seasonal cycle and specified 100 hPa GPH that varies day-to-day. There is also a substantial zonally asymmetric response that is unlike the NAM for both types of boundary conditions. The resolved planetary waves respond to produce an EP flux convergence roughly equal and opposite to the applied torque where the torque is strongest, with greater convergence in the upper stratosphere as well. The response to torques placed in the mid-stratosphere in middle latitudes is not NAM-like with either type of boundary conditions.

Overall, these results are generally consistent with those of other studies. The cancellation of the direct effect of a torque placed in the stratosphere by the planetary wave response in the steady state was also seen by Cohen *et al.* [2013] in both primitive equation and more comprehensive models with an interactive troposphere. They also report that a similar effect was seen by McLandress and McFarlane [1993] in experiments on parameterised gravity wave drag, but that they suggested it may be fortuitous. However, the negative response of the high latitude stratospheric ZMZW to a positive torque seen in the perpetual January experiments has not been previously reported – indeed, the studies of Song and Robinson [2004] and Chen and Zurita-Gotor [2008] indicate that the stratospheric ZMZW change in models with weak stationary wave forcing has the same sign as the torque. However, the strong EP flux response in the results presented here indicates that the stationary wave forcing is very important for bringing about the steady-state responses in the SMM simulations.

To further illustrate the importance of the stationary waves, figure 5.13 shows the ZMZW response to the standard ($\phi_0 = 60$N, $z_0 = 30$ km) torque in perpetual January runs performed in the same way as the runs described in sections 5.1.1.1 and 5.1.2, but with the zonally asymmetric part of the specified 16 km GPH set to half rather than twice that of the HadGEM2-CCS January climatology. In this case the torque causes a large positive ZMZW response at high latitudes, in contrast to the result shown

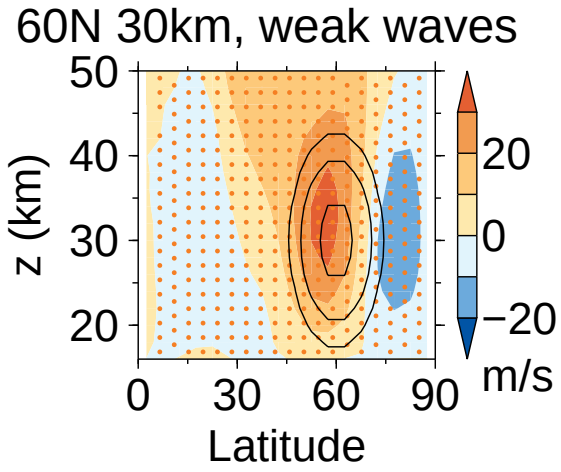

Figure 5.13: The steady state ZMZW response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions, with the eddy component of the prescribed 16 km GPH a quarter that in the standard control run, plotted as for the responses in figure 5.6. With weak stationary wave forcing, the ZMZW response is positive where the torque has its greatest strength.

in figure 5.6(b). This suggests that a certain strength of planetary wave forcing is required in order to bring about the responses presented in this work. (In this experiment, the climatological polar night jet in the control run is also much stronger than in the standard control run and peaks further north near 60N, so this may also have influenced the response to the applied torque.)

Overall, the results presented here indicate that there are things in common between AM dynamics and the stratospheric response to an applied torque, just as in tropospheric models. However, they do not strongly support the suggestion made in section 4.1 that the QBO-E minus QBO-W composite differences presented in chapter 4 are very NAM-like because this is the case for an arbitrary forcing perturbing the vortex, since the response to a torque in runs with a seasonal cycle was not found to be NAM-like. The climatology of these runs is more like that of the real atmosphere than that of the perpetual January runs, so the response in the runs with a seasonal cycle might be expected to be closer to the response of the real atmosphere to an externally-forced torque. This could indicate that the differences between the way the QBO influences the vortex and the fixed torque are important for the QBO

bringing about a NAM-like response in the real atmosphere, or that the SMM may not simulate the feedbacks that shape the response correctly.

It is difficult to understand the feedbacks using only the long-term mean response because this is a sum of the direct response and the feedbacks. The responses to torques in the seasonal cycle runs may not be AM-like because the feedbacks do not tend to make the response AM-like, or because they are not large enough to dominate the direct response to the torque, or both. More probing analysis of the feedbacks is presented in chapter 6 using the transient response to applied torques, and this indicates that the feedbacks are actually qualitatively similar in both perpetual January runs and those with a seasonal cycle, and do indeed tend to make the responses to torques placed at high latitudes more AM-like, but that they are weaker in the latter so they do not dominate the long-term mean response as they do in the perpetual January runs. However, the SMM does not realistically simulate all the processes that determine the strength of the feedbacks, as discussed below. Therefore it is not clear if the response to an externally-forced torque in the real atmosphere would be more like that in the perpetual January experiments or in those with a seasonal cycle.

It is interesting to verify in the perpetual January experiments that in principle the steady-state stratospheric response to a forcing may have the opposite sign to the forcing[3]. This has an important implication for studies of the HT relationship, namely that, based on the long-term vortex response to QBO forcing that has been the focus of most studies, it could be the case that the direct effect of the QBO-E phase is to *strengthen* the vortex, and that it is the resulting feedbacks that lead to the vortex being weaker on average in QBO-E. This possibility has not been considered by previous studies of the HT relationship, nor apparently with regard to the effect of other forcings on the vortex.

It is emphasised that the SMM is a simplified model of the stratosphere that is only expected to show behaviour that is qualitatively similar to the true dynamics.

[3]The response can also project negatively onto the forcing in simpler systems, for example the system of Lorenz [1963] [Palmer, 1999].

In particular, imposing fixed GPH at 16 km and parameterising gravity wave drag using Rayleigh friction may introduce unrealistic effects which should be considered.

Regarding the effect of fixing GPH at 16 km, comparison of the responses to the torques centred at (60N, 20 km), (60N, 30 km) and (60N, 40 km) with δz=10 km and those centred at (60N, 30 km) with δz=2.5 km and δz=40 km indicates that the qualitative nature of the response is not very sensitive to changing the torque strength near the lower boundary. Since boundary effects would be expected to be stronger the closer the torque is to the boundary, the bottom boundary therefore does not seem to have a strong effect on the response. The responses may be artificially small near the bottom boundary, however.

In the real atmosphere, the troposphere will respond to changes in stratospheric circulation (section 2.8), and Kushner and Polvani [2004] argued that the tropospheric response to stratospheric heating perturbations in a primitive equation model substantially altered the stratospheric response. However, the tropospheric response did not become strong for several hundred days after the stratosphere was perturbed in their experiments (though the tropospheric annular mode time scale in their model is too large by an order of magnitude [Chan and Plumb, 2009], suggesting that in the real atmosphere the response could become strong after about a month). Conversely, Garfinkel *et al.* [2012b] simulated the response of the vortex to forcing by QBO-E and argued that the resulting tropospheric response was not large enough to substantially affect the vortex. Therefore, on the basis of those experiments, the tropospheric response may become important after at least about month, though it is unclear how big a difference it would make to the results.

In the case of the steady-state response in the perpetual January runs, the weakening of high-latitude stratospheric ZMZW would be expected to cause an equatorward movement of the tropospheric jet (section 2.8) which is associated with greater upward EP flux in the stratosphere [Kushner, 2010]. This would be expected to reinforce the ZMZW decrease. Therefore, even if the tropospheric response feeds back to alter the stratospheric response, it is still plausible that a positive zonal torque could cause weakening of the high-latitude stratospheric ZMZW in the real atmosphere.

The absence of a tropospheric response may be part of the reason why in runs with a seasonal cycle the feedbacks do not bring about a NAM-like response to the torque, yet natural forcings acting on the real vortex do produce a NAM-like response in observations (section 4.1). It would be valuable to study the effect of a fixed torque in a model that includes the troposphere to see if the response is different. Comparing results from that experiment with those presented here would indicate how important feedbacks onto the stratosphere due to tropospheric changes are in determining the extratropical stratospheric response to external forcings such as the QBO and solar cycle, which has not been considered much in previous studies.

The SMM also parameterises the effects of gravity waves quite crudely using a Rayleigh friction term, so feedbacks from changes in these waves are not properly represented here. It is expected that the qualitative nature of the results presented here for the stratosphere would not change if a more realistic gravity wave parameterisation were included, as the contribution from planetary waves usually dominates the stratospheric wave drag, and models using Rayleigh friction to parameterise gravity wave drag have been found to simulate stratospheric variability well, for example in the studies discussed in section 3.3. The strength of the feedbacks relative to the direct effect of the torque may change, though. The mesospheric response may depend more strongly on the gravity wave parameterisation. However, it is shown in appendix 5.A that the qualitative nature of the stratospheric response to a torque is similar if Rayleigh friction is greatly weakened, indicating that the results are not strongly affected by including this in the SMM. It would be interesting to repeat some of these experiments in a model with a more sophisticated gravity wave parameterisation to see if the behaviour is similar to that of the SMM.

5.A Appendix: sensitivity to varying Rayleigh friction

Shepherd *et al.* [1996] and Shepherd and Shaw [2004] showed that Rayleigh friction can introduce error into the simulated residual circulation response to an applied torque. In this appendix it is shown that removing or greatly weakening the Rayleigh friction in the NH does not affect the qualitative nature of the simulated long-term mean stratospheric responses to torques that is of interest, and therefore the conclusions deduced from the results presented in this chapter do not depend on the use of Rayleigh friction to parameterise gravity wave drag.

5.A.1 Perpetual January run responses

A control perpetual January simulation and simulations with (ϕ_0=60N, z_0=30 km) and (ϕ_0=60N, z_0=40 km) torques (with other torque parameters set to their "standard" values) were performed as described in sections 5.1.1.1 and 5.1.2. In these runs the Rayleigh friction coefficient $1/\tau$, where τ is the relaxational time scale, was set to zero north of 20N, set to the value given in section 3.3 south of 10N, and linearly decreased to zero between 10–20N. Rayleigh friction is imposed south of 20N because without this the simulations become unstable, and the arguments of Shepherd *et al.* [1996] and Shepherd and Shaw [2004] do not suggest that Rayleigh friction applied outside the NH extratropics would strongly affect the response to a torque applied in this region. Sets of simulations with other perturbations to the Rayleigh friction, such as reducing it by half everywhere or raising the height at which τ starts increasing above 10^7 s, give similar results to those shown here.

Figure 5.14(a) shows the ZMZW response to the (ϕ_0=60N, z_0=30 km) torque in these runs. Comparing the ZMZW response with NH extratropical Rayleigh friction removed with the corresponding response shown in figure 5.6(b), it is seen that it is qualitatively similar but generally larger than if Rayleigh friction is applied north of 20N, and is positive rather than negative between ∼50–60N and ∼25–35 km. However, the response still has an anomaly correlation of 0.93 with the NAM signature in the

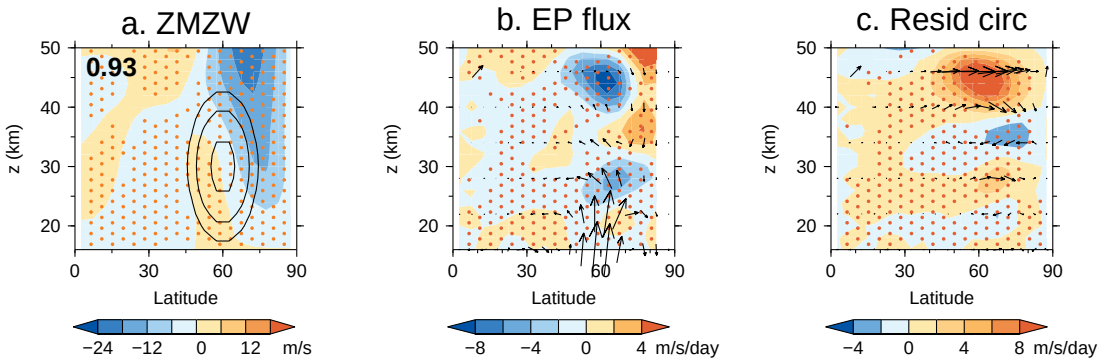

Figure 5.14: Effect of removing NH extratropical Rayleigh friction on the response. (a) shows the steady-state ZMZW response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions, plotted as in figure 5.6. (b) shows the EP flux and D_F responses to the same torque, plotted as in figure 5.9(b). (c) shows the response of the residual circulation and associated acceleration to the same torque, plotted as in figure 5.10. The responses are qualitatively similar to those in simulations with Rayleigh friction applied north of 20N.

corresponding control run (not shown). Therefore removing Rayleigh friction does not affect the tendency of feedbacks to make the response NAM-like. Note that the quantitative changes in the response may not just be directly due to removing Rayleigh friction but may also be due to the indirect effect this has though changing the climatology.

Figures 5.14(b) and 5.14(c) show the EP flux and D_F responses and the responses of the residual circulation and the associated acceleration respectively to the same torque, with zero NH extratropical Rayleigh friction. Again the responses are qualitatively similar to those with Rayleigh friction applied north of 20N (figures 5.9(b) and 5.10 respectively). The EP flux convergence peak in the upper stratosphere is larger and moved upwards slightly when Rayleigh friction is removed. The residual circulation response in the upper stratosphere is also correspondingly stronger and moved upward, and with a more equatorward extratropical response at ∼35 km it has a more wave-like vertical structure.

Torques placed closer to the region where Rayleigh friction is strong (above 50 km in the SMM experiments in previous sections) may be affected more strongly by the Rayleigh friction [Shepherd *et al.*, 1996]. Figure 5.15 shows the ZMZW response to

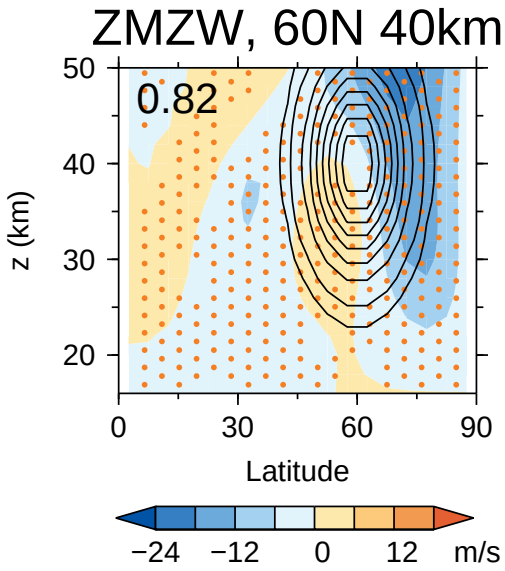

Figure 5.15: The steady-state ZMZW response to the (ϕ_0=60N, z_0=40 km) torque under perpetual January conditions with no NH extratropical Rayleigh friction, plotted as in figure 5.6. The response is larger but qualitatively similar to that shown for the same torque in figure 5.6(b).

the (ϕ_0=60N, z_0=40 km) torque with no NH extratropical Rayleigh friction. Again the response is larger but qualitatively similar to that with Rayleigh friction applied (figure 5.6(b)) and has a strong correlation with the NAM signature. The EP flux, D_F and residual circulation responses to this torque with and without NH extratropical Rayleigh friction are also qualitatively similar (not shown).

5.A.2 Seasonal cycle run responses

A control simulation with a seasonal cycle and simulations with (ϕ_0=60N, z_0=30 km) and (ϕ_0=60N, z_0=40 km) torques (with other torque parameters set to their "standard" values) were also performed as described in section 5.1. However, in these runs the Rayleigh friction coefficient was not set to zero between 20–90N above 50 km as this makes the simulations too unstable. Rather in this region it was set to 0.3 times that given in section 3.3. Below 50 km north of 20N the coefficient was set to zero. It was set to the value given in section 5.1 south of 10N with a linear taper between 10–20N.

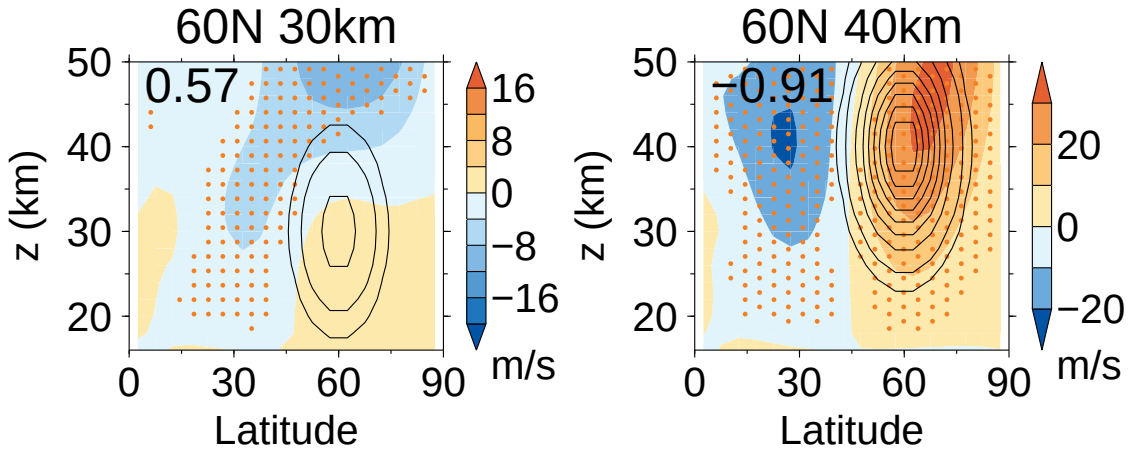

Figure 5.16: November–February mean ZMZW responses to the standard (ϕ_0=60N, z_0=30 km) and (ϕ_0=60N, z_0=40 km) torques in runs with a seasonal cycle with weakened NH Rayleigh friction, plotted as in figure 5.11. The responses are larger but qualitatively similar to those shown for the same torques in figure 5.11.

The control simulation and that with the (ϕ_0=60N, z_0=30 km) torque consisted of one 45-year ensemble member each. However, the simulation with the (ϕ_0=60N, z_0=40 km) torque became unstable after 25 years, but this provided enough data to see the differences in the response compared to the simulation with full Rayleigh friction.

Figure 5.16 shows the wintertime composite ZMZW responses to the (ϕ_0=60N, z_0=30 km) and (ϕ_0=60N, z_0=40 km) torques. Again these responses are stronger but qualitatively similar to those with full Rayleigh friction (figure 5.11). Figure 5.17 shows the EP flux and D_F responses to the (ϕ_0=60N, z_0=30 km) torque, which are just slightly stronger than those with full Rayleigh friction (figure 5.12). The residual circulation response to this torque is also largely similar with weak and full strength Rayleigh friction, as are the EP flux, D_F and residual circulation responses to the (ϕ_0=60N, z_0=40 km) torque (not shown).

Therefore the simulated long-term mean stratospheric responses to applied torques are qualitatively similar if NH extratropical Rayleigh friction is substantially weakened.

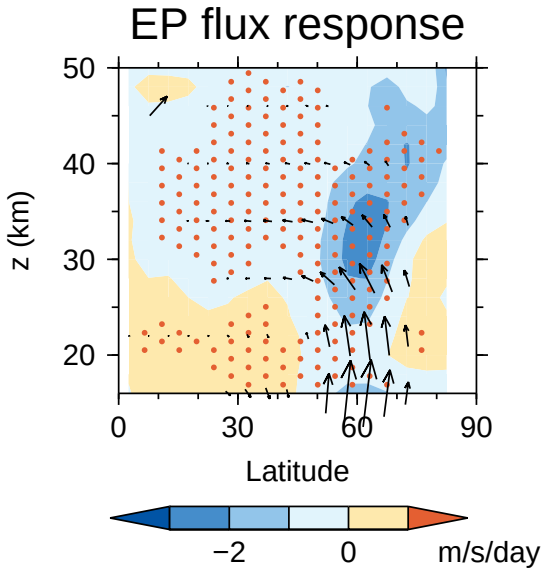

Figure 5.17: November–February mean EP flux and D_F responses to the standard (ϕ_0=60N, z_0=30 km) torque in runs with a seasonal cycle with weakened NH Rayleigh friction, plotted as in figure 5.12. The responses are similar to those shown in figure 5.12.

5.B Appendix: sensitivity to moving the momentum sink from the SH to the NH mesosphere

In the experiments described previously, a torque was applied in the NH stratosphere and an opposing torque was applied in the SH so that no net momentum was added globally. However, the location where the opposing torque is applied may affect the circulation response, and this may be relevant to understanding the effects of different forcings on the stratosphere. For example, a decrease in stratospheric easterly forcing by parameterised gravity wave drag in a model may be associated with an easterly forcing at higher altitudes if the parameterisation is momentum-conserving, so it is useful to test the effect of applying the opposing torque in the NH mesosphere rather than in the SH. In this appendix it is shown that doing this under perpetual January conditions does not strongly affect the qualitative nature of the long-term mean NH ZMZW, D_F and residual circulation responses, although the EP flux response changes character.

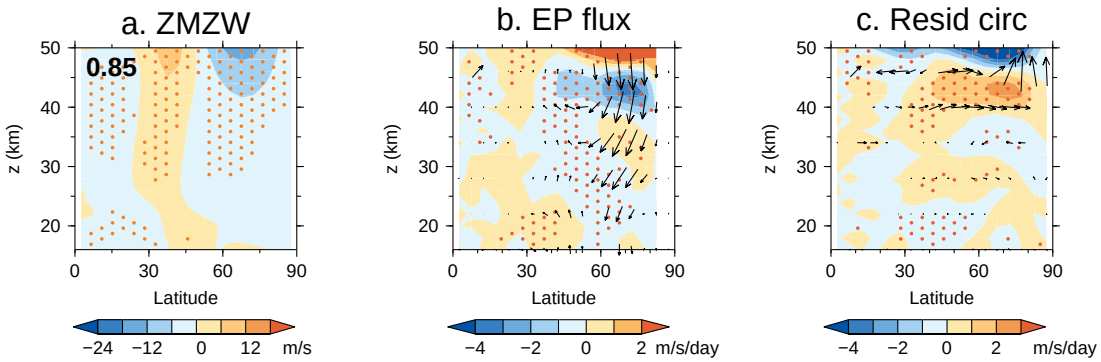

Figure 5.18: Effect of moving the opposing torque from the SH into the NH mesosphere on the response. (a) shows the steady-state ZMZW response to a (ϕ_0=60N, z_0=30 km) torque with peak magnitude 0.5 m/s/day under perpetual January conditions, plotted as in figure 5.6. (b) shows the EP flux and D_F responses to the same torque, plotted as in figure 5.9(b) with the arrow in the top left corner representing a flux $(3 \times 10^5, 3 \times 10^3)$ kg/s^2. (c) shows the response of the residual circulation and associated acceleration to the same torque, plotted as in figure 5.10 with the arrow in the top left corner representing a velocity $(5 \times 10^{-2}, 5 \times 10^{-4})$ ms^{-1}. The responses are qualitatively similar to those with the opposing torque in the SH (figure 5.14) except for the EP flux respnse.

Figure 5.18 shows the response under perpetual January conditions to a westerly torque centred at (60N, 30 km) with "standard" torque parameters and peak magnitude 0.5 m/s/day. The opposing easterly torque is constant at 60 km and above and zero below and has the same meridional profile as the westerly torque, so zero net momentum is added in any vertical column. The Rayleigh friction is zero north of 20N, as in the experiments in appendix 5.A. (If the NH extratropical Rayleigh friction were not zero then momentum would not be conserved and it would not be clear how to interpret the results, and for this reason a similar experiment with a seasonal cycle could not be performed, as the SMM becomes unstable with a seasonal cycle if NH extratropical Rayleigh friction is not present.) The SMM becomes unstable if the opposing torque is confined to be closer to the model top for this torque strength, as a very large easterly acceleration is then required, and similarly if the torque strength is increased the opposing torque must be applied at lower levels to maintain stability. The response to the torque was not found to be very sensitive to the torque strength

and minimum level at which the easterly torque is applied and that shown in figure 5.18 is a representative example. The effect of moving the opposing torque from the SH to the NH mesosphere can be seen by comparing figure 5.18 to figure 5.14.

Figure 5.18(a) shows the ZMZW response is again qualitatively like the NAM signature in the control run with no NH extratropical Rayleigh friction (not shown), with an anomaly correlation of 0.85. However, the ZMZW response in the upper stratosphere is larger relative to that at lower altitudes compared to the response shown in figure 5.14(a), and the resemblance to the NAM signature is better in the upper stratosphere.

Figure 5.18(b) shows the EP flux and D_F responses, showing again that there is increased EP flux convergence where the westerly torque peaks. Overall the D_F response is largely qualitatively similar to that in figure 5.14(b), though the upper stratospheric response is stronger relative to that at lower altitudes in figure 5.18(b). Correspondingly the residual circulation response shown in figure 5.18(c) is also qualitatively similar to that in figure 5.14(c), with a relatively larger response in the upper stratosphere.

The EP flux response (figure 5.18(b)) is quite different, however, being downward at high latitudes rather than upward. In chapter 6 it is shown that the upward EP flux response to the standard (ϕ_0=60N, z_0=30 km) torque, with the opposing torque placed in the SH, is consistent with the physical explanation that the direct effect of the torque increases the stratospheric refractive index (section 2.2.1) near 60N so that Rossby wave propagation from the 16 km boundary into the stratosphere increases. This reasoning predicts that when the opposing torque is placed in the mesosphere it decreases the refractive index there, so there is less wave propagation into this region, leading to a downward EP flux response below. This seems qualitatively consistent with the response in figure 5.18(b) being downward at high latitudes.

These results indicate that the qualitative natures of the ZMZW and D_F responses are not very sensitive to moving the opposing torque into the NH mesosphere, but the EP flux response itself depends on the location of the opposing torque. This does not affect the conclusions arrived at in section 5.5.

Chapter 6

Stratospheric response to an extratropical zonal torque: transient response

In chapter 5 it was found that the long-term mean response to an applied torque in the UKMO Stratosphere-Mesosphere Model (SMM; section 3.3) differs between simulations with perpetual January boundary conditions and those with a seasonal cycle. In the former, the response to torques centred in the high-latitude stratosphere is NAM-like and of opposite sign to the torque at high latitudes, whereas this is not the case in the latter. Since the question of whether the stratospheric response to arbitrary forcings is AM-like may be important for understanding the observed response of the vortex to natural forcings and how observational data can be used to deduce information about forcing mechanisms (section 4.2), it is of interest to understand the physical reasons for the differences between the simulations in order to understand better how the real extratropical stratosphere would respond to an externally-induced torque or other forcing. It was also discussed in section 4.2 how the long-term mean tropospheric responses to applied forcings are not always AM-like, and understanding why the stratospheric response to a torque is AM-like under perpetual January conditions in the SMM, but not in runs with a seasonal cycle, may help explain what is required to bring about an AM-like response more generally. In this chapter the transient responses of the SMM to applied torques are presented in

order to give insight into these problems. These experiments additionally provide a simple case in which the processes of wave-mean flow interaction can be studied.

6.1 Experimental method

Both "standard control runs" described in section 5.1.1 are used here. The perpetual January control run was extended to a total length of 40 years, and a wide range of torques are applied in simulations as described in section 5.1.2. The response to the standard (ϕ_0=60N, z_0=30 km) torque is focussed on as the transient responses to other torques centred at 60N or poleward are qualitatively similar. To examine the transient response to an applied torque, 90 day branch runs were performed under both perpetual January and seasonally varying boundary conditions with initial conditions taken from the appropriate control run, as described below. In these the torque strength was linearly increased from zero to full strength in the first ten days. Raising the period of this increase to twenty days was found not to qualitatively change the results, nor did decreasing it to zero apart from in the first few days when spurious oscillations in the uppermost model levels were apparent, so the results are not sensitive to varying this time.

Branch runs under perpetual January conditions were started with initial conditions taken from the 40-year standard perpetual January control run every 6-months, not including the initial 2-year spin-up time. Thus 76 branch runs were performed for each applied torque.

Branch runs with a seasonal cycle were initiated at January 1 of each year of the first ensemble member of the seasonal cycle control run, again excluding the two-year spin-up time. There were 43 branch runs for each torque.

Data are sampled daily in each experiment.

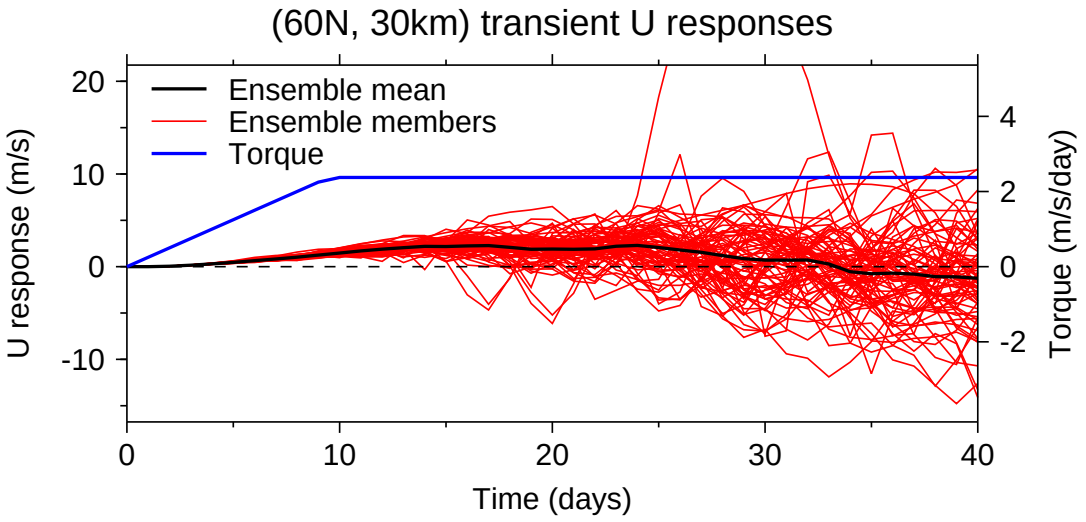

Figure 6.1: ZMZW responses to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions interpolated to (60N, 30 km) as a function of time of each branch run (red lines), the ensemble-mean response (black line) and the torque (blue line).

6.2 Perpetual January run responses

Figure 6.1 shows time series of the ZMZW responses at (60N, 30 km) over the first 40 days in perpetual January branch runs with the standard (ϕ_0=60N, z_0=30 km) torque. The responses are positive in all the ensemble members up to day 14, consistent with the expected direct response to the applied torque. Then shortly after this time the sign of the response begins to differ amongst the members, likely due to chaotic variability. The ensemble mean changes sign just after day 30.

Figure 6.2 shows the ensemble-mean transient ZMZW response as a function of latitude and height. On day 1, the structure of the ZMZW response resembles that of the torque very well near (60N, 30 km), consistent with the expected similarity between the response and the forcing on short time scales (section 4.3). The ZMZW response is already negative in the lower and middle stratosphere to the north and south of the torque position. The response resembles the steady state response to this torque when stationary wave forcing is weak (figure 5.13). Up to day 10, the positive ZMZW response spreads northwards, but between days 10 and 15 there is a substantial zonal deceleration in the polar upper stratosphere, giving a negative

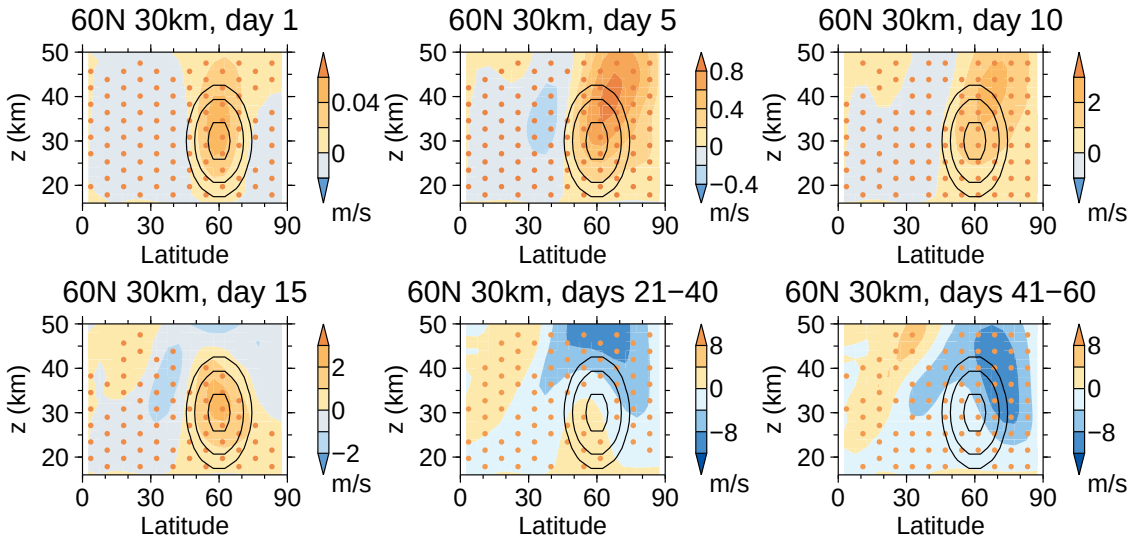

Figure 6.2: The ensemble-mean ZMZW response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions, averaged over the time periods of the branch runs indicated in the panel titles. Contours show the torque with levels at 0.5, 1 and 2 m/s/day from day 10, but these levels are scaled by factors 0.1 on day 1 and 0.5 on day 5. Stippling shows where the response is statistically significant above the 95% level.

ZMZW response. This descends with time and a positive ZMZW response develops in the subtropics to give an overall response that resembles the steady state response shown in figure 5.6 after ∼day 20, although the response does not become entirely steady within the 90-day length of the runs.

Statistical significances of the responses were calculated using a Monte Carlo (MC) bootstrap test. At each grid point, a surrogate data sample was generated according to the null hypothesis that the mean response is zero but other moments of the true distribution of responses equal those in the data. The mean of the responses for all pairs of branch and control runs was subtracted from the response for each pair and the results were resampled with replacement to produce surrogate ensembles of the same size as the originals. The probability of the mean of this resampled data being larger than that for the real data was estimated using 1,000 data resamplings.

The transient ZMZW responses to other torques placed at 60N or further poleward (not shown) are qualitatively similar. The responses to the (ϕ_0=30N, z_0=30 km) and (ϕ_0=45N, z_0=30 km) torques have substantial magnitudes up to ∼day 50, comparable

to the response to the standard (ϕ_0=60N, z_0=30 km) torque, but these subsequently decrease, consistent with the steady state responses being small (figure 5.6). So even though these torques produce substantial responses on short time scales, these are not maintained in the steady state.

In order to further investigate the dynamics of the stratospheric response, the transient response of each term in the TEM zonal momentum equation (equation 2.13) to the standard (ϕ_0=60N, z_0=30 km) torque is shown in figure 6.3. On day 1, $\partial\overline{u}/\partial t$ is dominated by the torque, and the EP flux response is small, with its convergence increasing where the torque is being applied, associated with increased upward flux from below and equatorward flux from higher latitudes. The residual circulation response is equatorward in most of the stratosphere with a negative associated acceleration term in equation 2.13

$$-\overline{v}^* \left[(a\cos\phi)^{-1}(\overline{u}\cos\phi)_\phi - f\right] - \overline{w}^*\overline{u}_z.$$

This is expected from the discussion in section 2.2.3, given that the sum of the torque and D_F is positive in the extratropical stratosphere, and this makes $\partial\overline{u}/\partial t$ less than half the peak torque strength of 0.25 m/s/day near (60N, 30 km). This response also causes $\partial\overline{u}/\partial t$ to be negative in the low latitude lower and middle stratosphere.

On days 6–10, $\partial\overline{u}/\partial t$ is larger due to the torque growing linearly in the first 10 days (figure 6.3). The EP flux response is also much larger, with the convergence response near (60N, 30 km) being more than half the strength of the torque. The residual circulation response is qualitatively similar to that on day 1 and stronger.

Over days 11–15, $\partial\overline{u}/\partial t$ becomes negative in the extratropical upper stratosphere due to the EP flux convergence response strengthening. The EP flux convergence response in the middle stratosphere now also cancels most of the direct effect of the torque. The negative $\partial\overline{u}/\partial t$ response is largest over days 16–20 when it also extends to the high latitude lower stratosphere. As a consequence of the EP flux convergence response almost completely cancelling the torque's direct effect, the residual circulation response has changed qualitatively so that the largest velocity response and associated acceleration is in the upper stratosphere, with an equatorward return

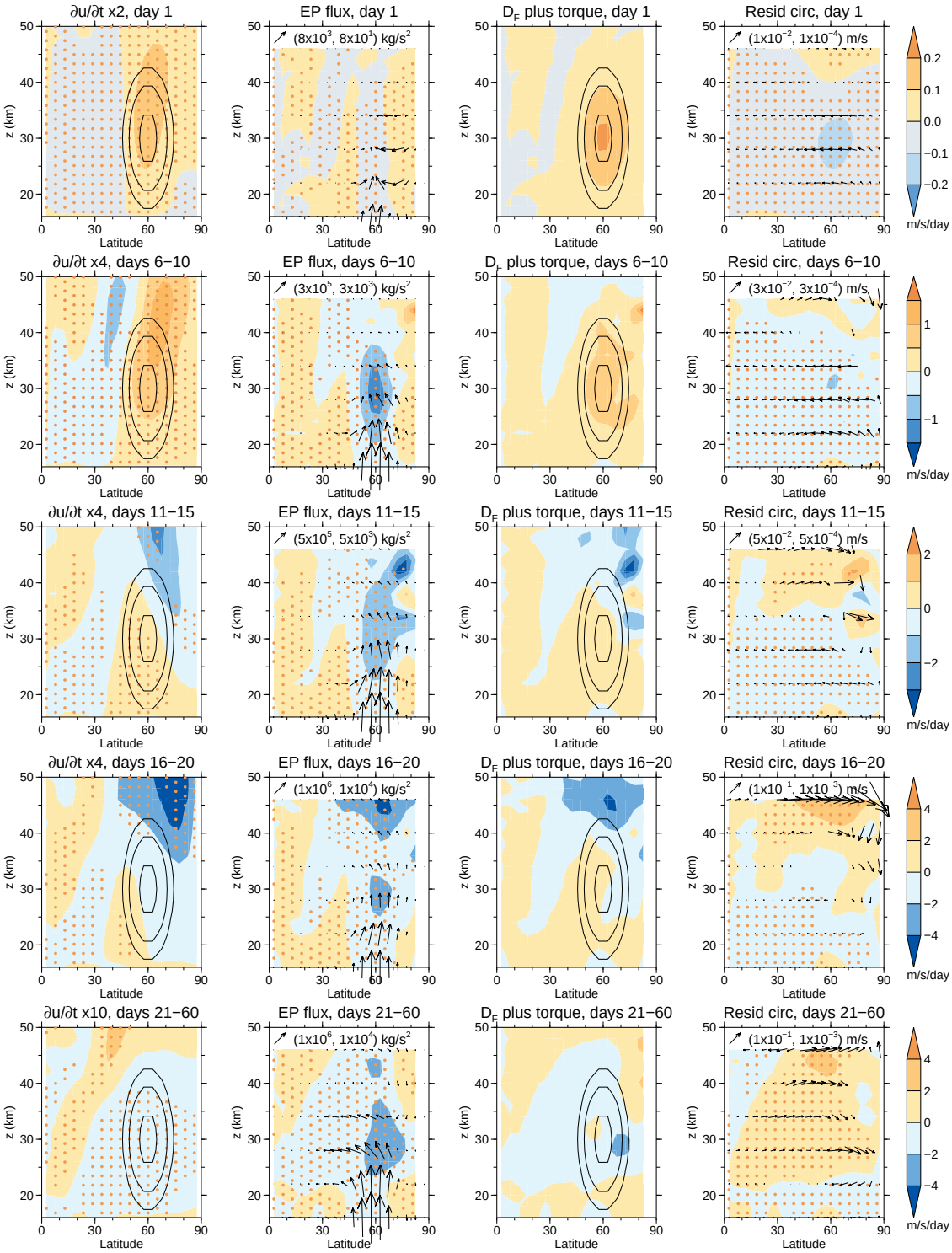

Figure 6.3: (Continued on the following page.)

Figure 6.3: (Continued.) In columns from left to right are the ensemble-mean responses to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions of: $\partial\overline{u}/\partial t$, the EP flux (arrows) and D_F (colours), the sum of D_F and the applied torque, and the residual circulation (arrows) and associated acceleration (colours). Data are averaged over a different time period of the branch runs on each row. $\partial\overline{u}/\partial t$ data are multiplied by a scaling factor as indicated in the plot titles. Contours in the first and third columns show the torque, with contours at 0.5, 1 and 2 m/s/day from the third row onwards, but these levels are scaled by factors 0.1 and 0.8 in the first and second rows respectively. In the second and fourth columns, arrows in the top left corner of each panel indicate the size of an arrow that represents the EP flux or residual circulation value written in brackets alongside. The EP flux and residual circulation response is only plotted where either the ϕ- or z-component is statistically significant above the 95% level, and are plotted every 6 km in height. Stippling in the first, second and fourth columns shows where responses in $\partial\overline{u}/\partial t$, D_F and the acceleration associated with the residual circulation are statistically significant above the 95% level respectively.

flow in the lower stratosphere, so the response begins to resemble the steady-state situation in figure 5.10. After day 20, the extratropical $\partial\overline{u}/\partial t$ is smaller and the EP flux and residual circulation responses evolve slowly to become more like the steady state responses. The ZMZW response in the high-latitude lower stratosphere becomes negative in this period (figure 6.2). In the subtropics, the ZMZW response becomes positive and similar to the NAM signature in this region (figure 5.3(a)) and this is associated with the Coriolis force acting on the poleward residual circulation response (figure 6.3).

Figure 6.4 shows time series of the ensemble-mean responses of $\partial\overline{u}/\partial t$ and terms in the TEM zonal momentum equation at (60N, 30 km) to the standard (ϕ_0=60N, z_0=30 km) torque. This shows that the ZMZW acceleration here is driven by the torque and resisted by responses in the EP flux and residual circulation up to about day 20. After this time the EP flux convergence response tends to be larger than the torque, so the ZMZW acceleration response tends to be negative, with the residual circulation response contributing positively. So it can be seen that the EP flux convergence response drives the easterly ZMZW acceleration in the high-latitude stratosphere after day 20. In the steady state (chapter 5), the sum of D_F and the torque is about zero. The local Rayleigh friction contribution is very small (not shown).

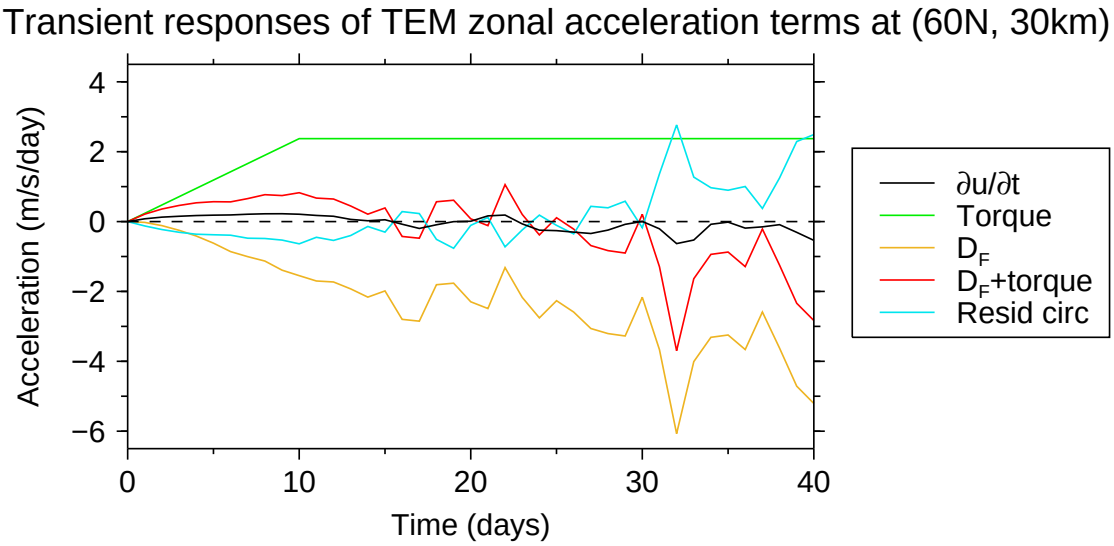

Figure 6.4: Ensemble-mean responses of $\partial u/\partial t$ (black), the torque (green), D_F (orange), the sum of D_F and the torque (red) and the acceleration associated with the residual circulation (blue) to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions interpolated to (60N, 30 km) as a function of time since initiation of each branch run.

The response of the terms in the TEM zonal momentum equation to other torques centred at 60N and polewards (not shown) is qualitatively similar, with the EP flux responding to cancel the applied torque and converge more in the upper stratosphere from about day 10, bringing about deceleration of the ZMZW.

The response to the (ϕ_0=30N, z_0=30 km) torque does not show increased EP flux convergence in the region of the applied torque, rather greater convergence develops in the lower and middle stratosphere between ∼0–30N and ∼60–80N with greater divergence in between (not shown). The EP flux response to the (ϕ_0=45N, z_0=30 km) torque is convergent near where the torque is centred but divergent north of 60N until after day 20, so it takes longer for the ZMZW response at high latitudes to turn negative than in response to the standard (ϕ_0=60N, z_0=30 km) torque (not shown).

It is shown in appendix 6.A that these results are not sensitive to varying Rayleigh friction in the SMM.

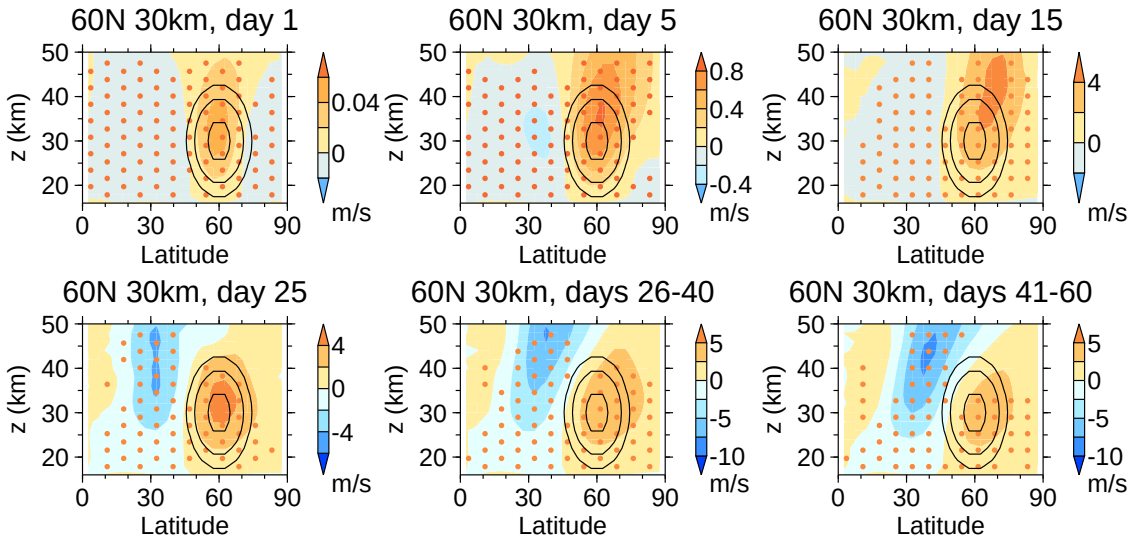

Figure 6.5: As in figure 6.2 but for the responses to the standard (ϕ_0=60N, z_0=30 km) torque in runs with a seasonal cycle.

6.3 Seasonal cycle run responses

Figure 6.5 shows the ensemble-mean transient ZMZW response to the standard (ϕ_0=60N, z_0=30 km) torque in branch runs with a seasonal cycle. The response in the first 25 days goes through approximately the same stages as that displayed by the perpetual January simulations in the first 15 days (figure 6.2), with acceleration initially at high latitudes and a negative ZMZW response developing further south. However, the response remains positive in most of the polar stratosphere, consistent with the wintertime composite responses (figure 5.11).

Figure 6.6 shows the ensemble-mean transient response of the terms in the TEM zonal momentum equation to the same torque. Again this shows the same sequence of stages as for the perpetual January response (figure 6.3), with the torque initially causing $\partial\overline{u}/\partial t$ to be greater, and causing a more equatorward middle and lower stratospheric residual circulation. The EP flux convergence increases in the region where the torque is strongest, and later it increases in the upper stratosphere, causing the ZMZW to decelerate there. However, the EP flux response is weaker and this sequence again unfolds more slowly than in the perpetual January case. It takes until days 16–20 for the EP flux response to largely cancel out the direct effect of the torque

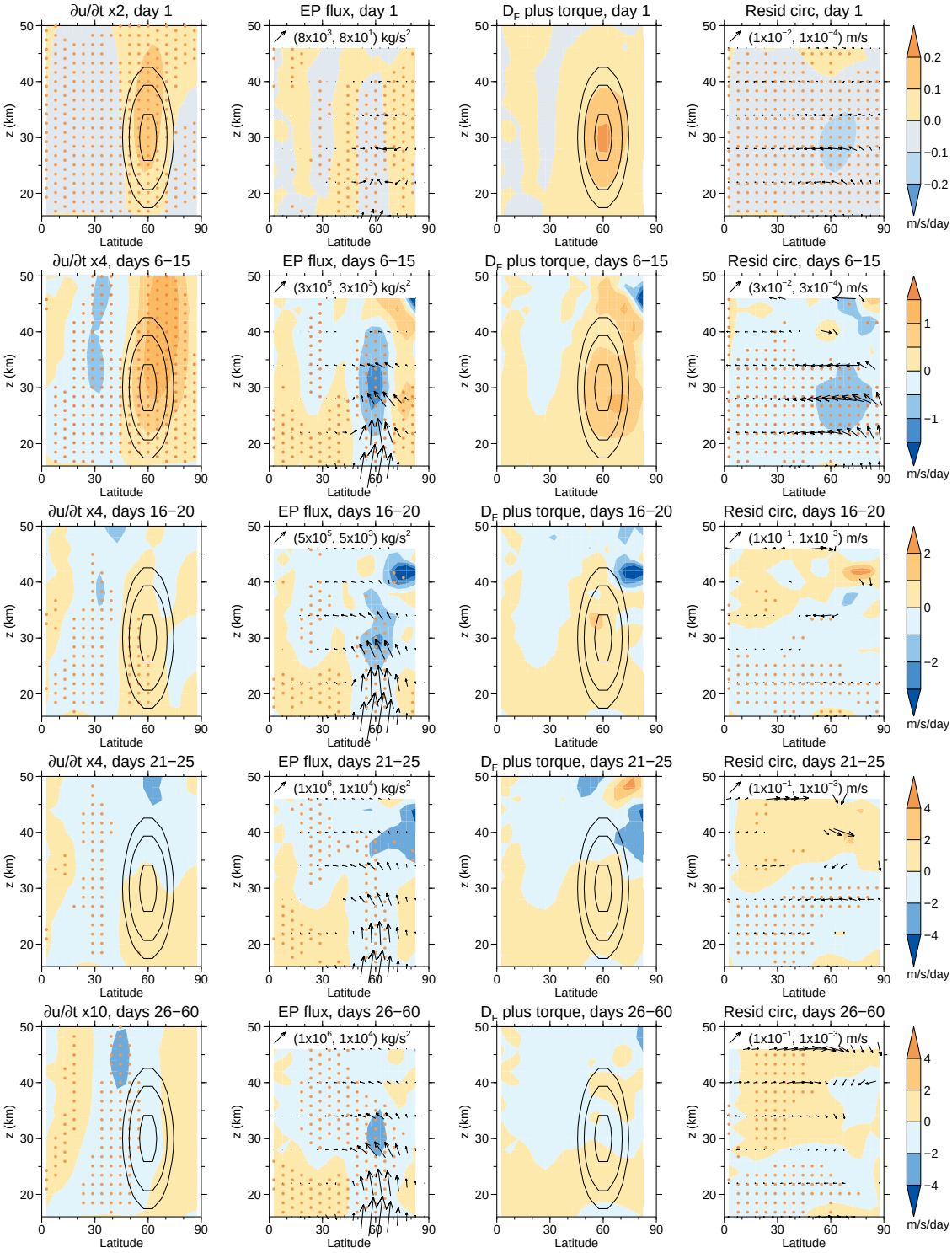

Figure 6.6: As in figure 6.3 but for the responses to the standard (ϕ_0=60N, z_0=30 km) torque in runs with a seasonal cycle. Contours in the first and third columns showing the torque are plotted slightly differently to those in figure 6.3, with contours again at 0.5, 1 and 2 m/s/day from the third row onwards with these levels scaled by a factor of 0.1 in the first row, but the scaling factor in the second row is 0.9.

in the high latitude middle stratosphere and to cause $\partial\overline{u}/\partial t$ to become negative in some places north of 60N. Under perpetual January conditions these events happen during days 11–15. This difference may be because the zonal mean flow and EP flux co-evolve and respond to changes in each other – the feedbacks are weaker in runs with a seasonal cycle which may make the time-derivatives of each smaller, so the sequence of changes proceeds more slowly. The EP flux convergence response in the high latitude upper stratosphere also never becomes as large as it does under perpetual January conditions, and in the lower stratosphere north of ∼70N it never becomes large enough to make $\partial\overline{u}/\partial t$ negative, though at (60N, 30 km) it does temporarily become larger than the torque and make $\partial\overline{u}/\partial t$ negative from ∼day 25. This weaker EP flux response appears to be the reason why the long-term mean ZMZW response does not become negative in the high latitude lower and middle stratosphere in runs with a seasonal cycle, and hence why it does not appear NAM-like (chapter 5).

The responses to other torques centred at 60N or poleward are qualitatively similar (not shown). The responses are also qualitatively similar for branch runs begun on October 22 rather than January 1, except that the EP flux response is weaker and the sequence unfolds more slowly (not shown), likely because wave forcing at the lower boundary is weaker at this time.

Therefore the transient responses to torques centred in the high-latitude stratosphere under perpetual January conditions and with a seasonal cycle each proceed through three stages:

1. The torque causes acceleration of the ZMZW and induces EP flux convergence where the torque is strongest, and it drives a residual circulation that is equatorward in the stratosphere and poleward above.

2. The EP flux becomes more upward and convergent in the extratropical upper stratosphere, causing the ZMZW here to decelerate and the residual circulation response to become poleward in the upper stratosphere.

3. The ZMZW acceleration reduces to approximately zero in the high latitude stratosphere, and in the case of the perpetual January runs becomes temporarily strongly negative, due to EP flux convergence increasing. The EP flux response is weaker in the runs with a seasonal cycle, and this produces quite a different long-term mean response to that in perpetual January runs (chapter 5).

6.4 Understanding the wave response

It can be understood from the discussion in sections 2.2.2 and 2.2.3 why the $\partial\overline{u}/\partial t$ response tends to have the same sign as the sum of the torque and D_F response, and why the residual circulation responds such that the associated acceleration contribution tends to have the opposite sign. A complete explanation of the response must account for the change in the EP flux, however.

Figure 6.7(a) shows the transient response of the refractive index for stationary waves squared (n^2, given by equation 2.8, dropping the s subscript and using reference temperature T_s=240 K and scale height H=6.95 km) to the standard (ϕ_0=60N, z_0=30 km) torque after 1 day under perpetual January conditions, along with the zonal wavenumber-1 component of the EP flux response. n^2 was introduced in section 2.2.1 as a diagnostic which may enable prediction of EP flux changes. The presented n^2 response is a "trimmed mean" [Wilks, 2006] of the responses over the ensemble members, where the top and bottom 10 values were excluded at each grid point to remove large outliers that arise after a few days at a few grid points because some ensemble members have small $\overline{u}$ values, making n^2 very sensitive to changes in $\overline{u}$ and $\overline{q}_\phi$. The important features identified in the following discussion are similar if the mean of the responses is used, and those in the first few days can also be seen in the responses for individual ensemble members.

On day 1, there is a very good agreement between the structure of the n^2 change in the extratropical lower and middle stratosphere and the wavenumber-1 EP flux response expected according to ray theory between about 50–70N and 16–40 km, which is also seen in the full EP flux response (figure 6.3). Here n^2 is increased and

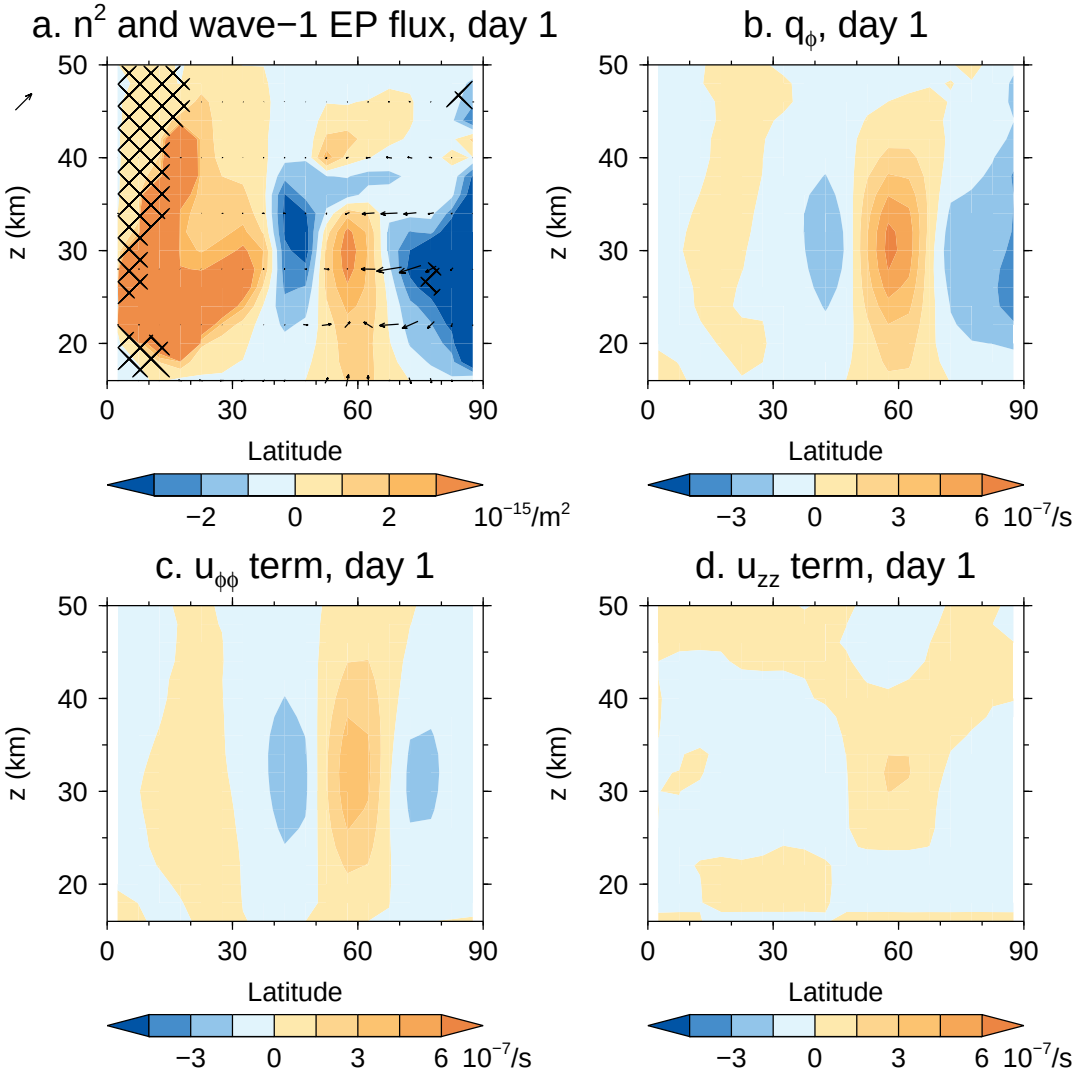

Figure 6.7: (a) shows the trimmed-mean response of the squared refractive index for stationary waves (n^2, colours) and zonal wavenumber-1 EP flux (arrows) to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions after 1 day. Hatching shows where trimmed-mean n^2 for zonal wavenumber-1 is negative in the control run. The reference arrow to the left of the plot represents EP flux (5000, 50) kg/s^2. (b) shows the response of the meridional PV gradient $\overline{q}_\phi$. (c) shows the contribution to the response in $\overline{q}_\phi$ associated with the "$u_{\phi\phi}$ term" $-\overline{u}_{\phi\phi}/a$ and (d) shows the same for the "u_{zz} term" $-af^2\overline{u}_{zz}/N^2$, which dominate the $\overline{q}_\phi$ response.

the EP flux below and to the north of (60N, 30 km) is directed more into this region. n^2 decreases to the north and south of this region and the EP flux response is directed away from these regions. There is also a large increase in n^2 south of ∼30N in the stratosphere which is not associated with a substantial change in the EP flux.

The D_F response is not directly predicted by ray theory, but it may be anticipated that EP flux convergence will tend to increase where wave propagation increases, for example due to dissipative processes acting on the waves.

The wavenumber-2 EP flux response (not shown) is similar to the wavenumber-1 response in figure 6.7(a) and corresponds well with what is expected given the n^2 response, but the vertical flux response below the region of convergence is larger and the equatorward response to the north is smaller.

The refractive index response δn^2 is approximately given by

$$\delta n^2 \approx (\delta \overline{q}_\phi - \delta \overline{u}\, \overline{q}_\phi / \overline{u})/a\overline{u}, \tag{6.1}$$

where $\delta\overline{q}_\phi$ and $\delta\overline{u}$ are the $\overline{q}_\phi$ and $\overline{u}$ responses respectively, derived from equation 2.8 taking $\delta\overline{q}_\phi/\overline{q}_\phi$ and $\delta\overline{u}/\overline{u}$ to be small, and noting that N^2 is assumed to be constant. On day 1 in the extratropical lower and middle stratosphere, the largest contribution to δn^2 is associated with $\delta\overline{q}_\phi$ (figure 6.7(b)) with the $\delta\overline{u}$ term providing a smaller, oppositely signed contribution (not shown), as can be inferred from figure 6.2. $\delta\overline{q}_\phi$ is contributed to mostly by a change in $\overline{u}_{\phi\phi}$ and also partly by a change in $\overline{u}_{zz}$. These decrease where $\overline{u}$ is being accelerated most rapidly by the direct effect of the torque around (60N, 30 km), and the terms $-\overline{u}_{\phi\phi}/a$ and $-af^2\overline{u}_{zz}/N^2$ in the expansion of equation 2.7 are associated with most of the increase in $\overline{q}_\phi$ here (figures 6.7(c) and 6.7(d)). $\overline{u}_{\phi\phi}$ increases on the meridional flanks of this region and $\overline{u}_{zz}$ increases above and below, giving changes in $\overline{q}_\phi$ of opposite sign.

Thus the convergence of EP flux into the region of the applied torque on day 1 is qualitatively consistent with the physical description of Rossby wave propagation increasing where the meridional PV gradient is increased due to curvature of $\overline{u}$ becoming more negative. The large increase in n^2 near the tropics is due mostly to $\overline{u}$

decreasing there (figure 6.2), so that $1/\overline{u}$ in equation 2.8 increases a lot, since $\overline{u}$ is already typically small in the control run (figure 5.2(a)).

There is a good correspondence between the extratropical lower and middle stratospheric transient EP flux and refractive index responses in the first few days for all the torques in the perpetual January simulations.

In simulations with a seasonal cycle, the EP flux response to the standard (ϕ_0=60N, z_0=30 km) torque in the first few days is also consistent with changes in the refractive index (not shown) caused by changes in the QG PV gradient, with the peak refractive index change after one day being about half that in the perpetual January runs shown in figure 6.7. This is because the climatological $\overline{u}$ (figures 5.2(a) and 5.4(a)) and $\overline{q}_\phi$ (not shown) are both about twice as large in the seasonal cycle control run as in the perpetual January control run, and $\delta\overline{u}$ and $\delta\overline{q}_\phi$ are similar, as expected since the direct $\overline{u}$ response to the torques ought to be similar. So by equation 6.1, δn^2 is half as large in the runs with a seasonal cycle. Therefore it seems that the initial EP flux response is less strong in the runs with a seasonal cycle due to the vortex being stronger and more poleward on average.

On days 6–10, the perpetual January n^2 response to the standard (ϕ_0=60N, z_0=30 km) torque is largely similar to that on day 1, though it has become positive in the stratosphere near the pole (figure 6.8(a)), due to the $\overline{q}_\phi$ response becoming positive there (not shown). The upward EP flux response near 60N in the lower stratosphere is still qualitatively consistent with the n^2 response and ray theory, if propagating waves can be said to be generated at the lower boundary.

The n^2 response is generally positive at all times in the lower stratosphere between ∼50–70N (not shown), with the decrease in $\overline{u}$ contributing positively after about day 20 (figure 6.2) as well as changes in $\overline{q}_\phi$ – note that the $\overline{q}_\phi$ response stays positive where the torque is centred even after the $\overline{u}$ response turns negative, since $\overline{u}$ decreases more to the north and south of where the torque is applied than where it is centred (figure 6.2). This is qualitatively consistent with the EP flux response being more upward here (figure 6.3) and ray theory.

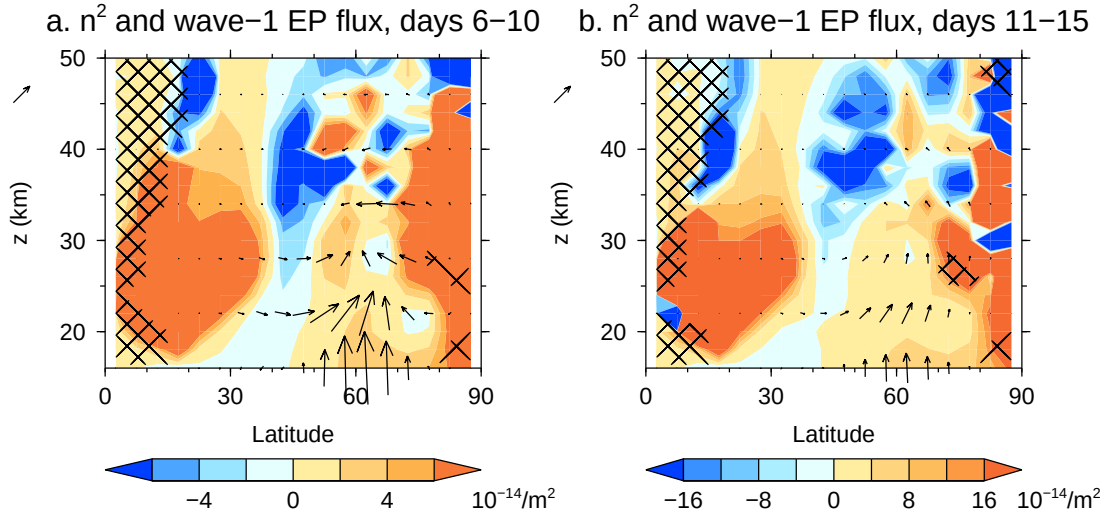

Figure 6.8: The trimmed-mean response of the squared refractive index for stationary waves (n^2, colours) and zonal wavenumber-1 EP flux (arrows) to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions averaged over (a) days 6–10 and (b) days 11–15 of the branch runs. Hatching shows where trimmed-mean n^2 for zonal wavenumber-1 is negative in the control run. The arrow to the left of each plot indicates the size of an arrow corresponding to EP flux of $(10^5, 10^3)$ kg/s^2 in (a) and $(5 \times 10^5, 5 \times 10^3)$ kg/s^2 in (b).

The EP flux response is upward and equatorward above ∼35 km during days 11–15 (figure 6.8(b)). This is not clearly consistent with what would be expected according to ray theory and the negative n^2 response in this region. The n^2 response here remains negative in the next few weeks and so it is not clear that the refractive index diagnostic is useful for explaining the second stage of the transient EP flux response, when the EP flux response becomes upward and equatorward and the D_F response becomes strongly negative in the upper stratosphere. The response is qualitatively similar for zonal wavenumber-2 (not shown).

Another way of physically understanding the second stage of the EP flux response is suggested by figure 6.9. This shows the ensemble and day 11–25 mean response of $\widetilde{D_F} = 2\pi a^2 \cos(\phi) \nabla \cdot \mathbf{F}$ integrated over 30–80N and 35–50 km, plotted against the day 6–10 mean F^z response (δF^z) at 30 km integrated between 45–80N for all the torques in the perpetual January and seasonal cycle experiments. The integral of δF^z

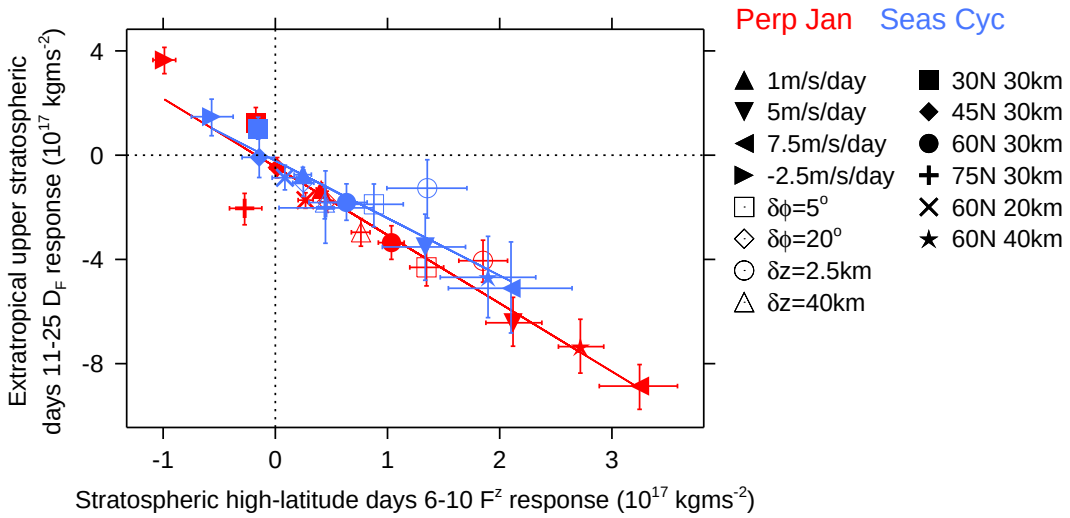

Figure 6.9: The ensemble and day 11–25 mean responses in $\widetilde{D_F}$ integrated over 30–80N and 35–50 km plotted against the ensemble and day 6–10 mean F^z responses at 30 km integrated between 45–80N, for all the torques in the perpetual January (red) and seasonal cycle (blue) experiments. The solid lines are the least-squares linear fits to each set of experiments. Error bars show the 2.5th–97.5th percentiles of the distribution of the mean according to an MC bootstrap estimate. Unless otherwise specified in the legend, the torques peak at (ϕ_0=60N, z_0=30 km), have meridional scale $\delta\phi = 11°$ and vertical scale $\delta z = 10$ km, and have a magnitude such that the total zonal momentum added to the NH equals that of the standard (ϕ_0=60N, z_0=30 km) torque with peak magnitude 2.5 m/s/day.

is defined as

$$\int_{45\mathrm{N}}^{80\mathrm{N}} 2\pi a^2 \cos(\phi)\delta F^z|_{z=30\,\mathrm{km}}\,\mathrm{d}\phi.$$

This expression follows from applying Stokes' theorem to equations 2.5 and 2.6 of Dunkerton *et al.* [1981] to infer the negative contribution of δF^z to the response of $\widetilde{D_F}$ integrated within a closed surface of which the surface at 30 km between 45–80N forms a part. The day 11–25 period includes that of greatest ZMZW deceleration and EP flux convergence responses in the upper stratosphere in both perpetual January runs and those with a seasonal cycle (figures 6.3 and 6.6).

Also shown in figure 6.9 are the 95% confidence intervals of the mean response in each quantity, calculated using an MC bootstrap method in which the responses for the individual ensemble members were resampled with replacement to produce

surrogate ensembles of the same size as the originals. The confidence intervals are the range between the 2.5th–97.5th percentiles of the distribution of the means of 1000 such surrogate data samples.

The high-altitude $\widetilde{D_F}$ response during days 11–25 is more negative for torques that cause a greater F^z response in the middle stratosphere between days 6–10, which as argued previously seems to be due to torques increasing $\overline{q}_\phi$ here, with correlations of -0.96 and -0.93 across the perpetual January experiments and those with a seasonal cycle respectively. The regression coefficient of the mean $\widetilde{D_F}$ response against the mean F^z response is slightly smaller in the experiments with a seasonal cycle than in the perpetual January experiments.

This is consistent with the physical explanation that the mid-stratospheric upward EP flux response in days 6–10 is associated with propagating waves with an upward group velocity, which carry easterly momentum into the upper stratosphere in the next couple of weeks, with deceleration of the ZMZW enhanced by the lower air density here. The data are also consistent with the D_F responses here in the seasonal cycle experiments being weaker than those in the perpetual January experiments being due mainly to the mid-stratospheric F^z response before day 10 also being weaker[1].

The response to the (ϕ_0=75N, z_0=30 km) torque under perpetual January conditions is an outlier in figure 6.9, showing mid-stratospheric F^z and high-altitude $\widetilde{D_F}$ responses that are both negative. Inspection of the **F** response (not shown) shows that between days 6–10 it is mainly poleward at 30 km and convergent near (75N, 30 km), and does not become substantially upward in the stratosphere until after day 10. It does not become substantially convergent at high altitudes until after day 15, later than in the response to the standard (ϕ_0=60N, z_0=30 km) torque (figure 6.3), so this still seems consistent with the above physical explanation.

[1]It is possible that the response could also be shown to be consistent with ray theory if a detailed ray tracing calculation were performed [Matsuno, 1970; Karoly and Hoskins, 1982], if greater upward wave propagation due the increase in n^2 in the lower stratosphere were predicted to lead to greater wave propagation to high altitudes despite the decrease in n^2. However, this would not yield any greater physical insight than the analysis performed above and this is not pursued here.

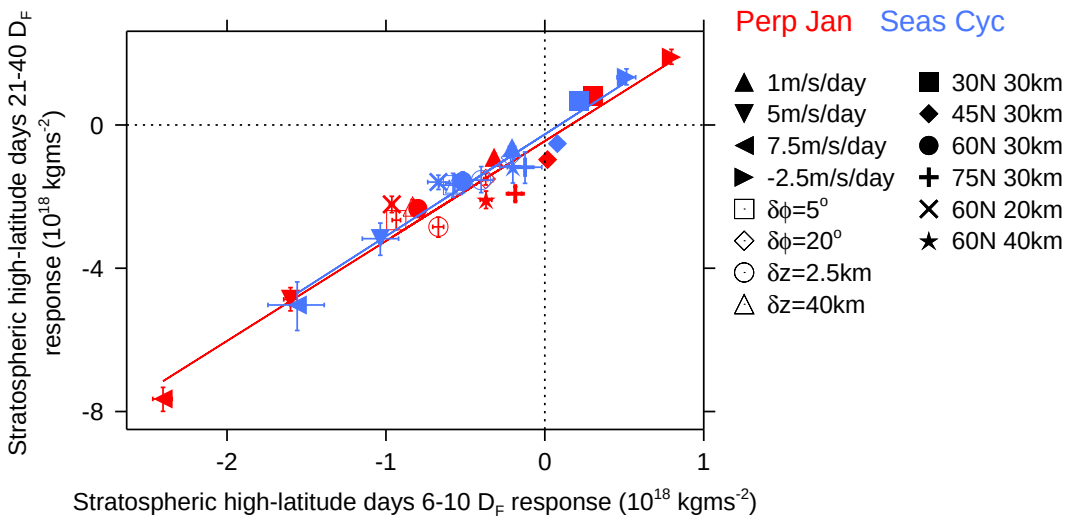

Figure 6.10: The ensemble and day 21–40 mean responses in $\widetilde{D_F}$ integrated over 45–80N and 16–40 km plotted against the day 6–10 mean $\widetilde{D_F}$ responses in the same region for all the torques in the perpetual January (red) and seasonal cycle experiments (blue), plotted as in figure 6.9.

The reason why, in the third phase of the response, the D_F response in the high-latitude stratosphere becomes large enough to cause the ZMZW response to turn negative in the perpetual January runs and not in those with a seasonal cycle seems harder to pin down in physical terms. The $\widetilde{D_F}$ response here between days 21–40 is well correlated with the $\widetilde{D_F}$ response here in days 6–10 (figure 6.10), so that the weaker response over days 21–40 in the runs with a seasonal cycle seems related to the weaker initial response, but this does not favour any one physical explanation in particular. For example, it could be related to δn^2 continuing to be larger in the perpetual January runs than in the seasonal cycle runs (not shown) and encouraging greater EP flux convergence. It could also be related to changes in the circulation that have happened in other locations by this time and which may also affect the high-latitude stratospheric EP flux. Given the complexity of the dynamics there is no particular reason to expect that a physically straightforward explanation can be found.

Cohen *et al.* [2013] argued that increased convergence of the EP flux associated with resolved waves in the region of an applied torque is necessary, because otherwise

the flow would eventually become unstable due to $\overline{q}_\phi$ changing sign. The fact that in our experiments the EP flux becomes more convergent in the region where the torques peak within a day, for torques centred at 45N or poleward (shown for the (ϕ_0=60N, z_0=30 km) torque in figures 6.3 and 6.6), implies though that wave feedbacks begin counteracting the torques before instability develops. $\overline{q}_\phi$ remains positive in our runs with applied torques, except very near 90N (not shown). This is consistent with the analysis of Cohen *et al.* [2014].

6.5 Discussion and conclusions

Examination of the transient response to an applied stratospheric torque in the SMM shows that the feedback contributions to the transient responses are qualitatively similar in perpetual January runs and in those with a seasonal cycle, but they are weaker in the latter, and this causes the long-term responses to be quite different (chapter 5). The perpetual January experiments here and in chapter 5 indicate that the feedback part of the responses is composed of a component that has a lot of similarity with the AM on time scales longer than a couple of weeks, and an EP flux convergence component that is approximately opposite to the torque. The combination of the feedback and the torque is therefore an AM-like total response on long time scales. In the runs with a seasonal cycle, the feedback does not seem strong enough to dominate over the direct effect of the applied torque, so the long-term response is not AM-like. This supports the conceptual view that the feedback parts of the responses to an arbitrary forcing tend to make the total response AM-like on long time scales in the stratosphere, so long as the feedback part dominates over the direct response, which builds upon the discussion in sections 4.2 and 4.3. This is difficult to understand from the long-term mean responses alone (chapter 5), which have been the focus of studies of the AM-like responses to forcings in the troposphere [e.g. Son and Lee, 2006; Ring and Plumb, 2007, 2008; Branstator and Selten, 2009], and this conceptual view may also apply in that situation. Simulating the relative strengths

of this feedback and the direct response accurately is therefore important in models for producing even qualitatively realistic responses to applied forcings.

The results presented here have also given insight into the physical nature of the feedbacks, which are associated with wave-mean flow interaction. They indicate that parts of the response are consistent with a physical description in terms of changing wave propagation and that therefore such a description can be appropriate even when changes in the EP flux are an order of magnitude smaller than the climatology. In the lower and middle extratropical stratosphere, the results are consistent with the physical explanation that, at first, the torque increases the meridional QG PV gradient where the torque is strongest, and hence increases n^2, so propagation of Rossby waves and their transport of easterly momentum into this region increases. At high latitudes this leads to an increase in EP flux convergence where the torque is strongest, that largely cancels out the direct effect of the torque when it is fully developed. This feedback is then weaker in the experiments with a seasonal cycle because the vortex is stronger and more poleward on average, so that the torque increases n^2 less. This stage is followed by EP flux convergence increasing in the upper stratosphere and causing deceleration of the ZMZW. The results are consistent with the explanation that the additional upward-propagating Rossby waves in the middle stratosphere produced in the first stage propagate into the extratropical upper stratosphere and transfer easterly momentum there. In the third stage, the EP flux convergence in the extratropical stratosphere increases and causes the ZMZW acceleration to become temporarily negative, and there could be multiple physical explanations for this, but it was shown that this is stronger for torques that produce a stronger EP flux response in the first stage. In the perpetual January experiments this causes the ZMZW response to become negative. This indicates that the initial response of n^2 to the torque is important for simulating the feedbacks correctly. The n^2 response depends on the ZMZW, and therefore this should be realistic in a model if it is to simulate a correct response.

It is possible that there are also other physical explanations that are consistent with the results of these experiments. It is not altogether clear why the result of

these physical processes is a circulation response that is similar to the NAM. Deeper analysis, beyond the scope of this work, may indicate if the NAM-like response occurs, for example, due to regime behaviour [Palmer, 1999], or to NAM-like anomalies having a long decay time scale [Branstator and Selten, 2009].

Much of the discussion in section 5.5 about differences between the behaviour of the SMM and the real atmosphere also applies here. The lack of an interactive troposphere in the SMM seems unlikely to affect the transient response to a torque for the first few weeks. The way in which gravity waves are parameterised may affect the strength of the feedbacks, though it seems unlikely to qualitatively change the feedback part of the response, and it is shown in appendix 6.A that the results are not sensitive to varying the Rayleigh friction. The SMM is not expected to predict the strength of the feedbacks with high quantitative accuracy, so it is not clear how strong they would be in the case of an externally-induced extratropical stratospheric torque in the real atmosphere. The NAM-like nature of the observed influence of natural forcings on the extratropical stratosphere (section 4.2) indicates that the feedbacks induced in response to those forcings are stronger than in the response to a torque simulated by the SMM in runs with a seasonal cycle. This is despite the mean ZMZW being slightly weaker in the control run than in ERA-40 (figure 5.4), so that its n^2 response to a torque would be expected to be slightly too large on average. This is perhaps related to the climatological F^z being too weak in these runs (section 5.1.1.2), since F^z has an important role in the feedback part of the response.

The degree to which the atmospheric response to a given forcing is AM-like may be important for predicting future trends in response to, for example, rising greenhouse gas concentrations or changes in stratospheric ozone, because if the response is largely AM-like then the work of making the prediction simplifies largely to predicting just the change in the AM index [Son and Lee, 2006; Kushner, 2010]. The work presented here suggests that examining the transient response to applied forcings, on top of investigating the long-term responses, may be helpful for understanding this, and this may help understanding of the tropospheric response to applied forcings as well.

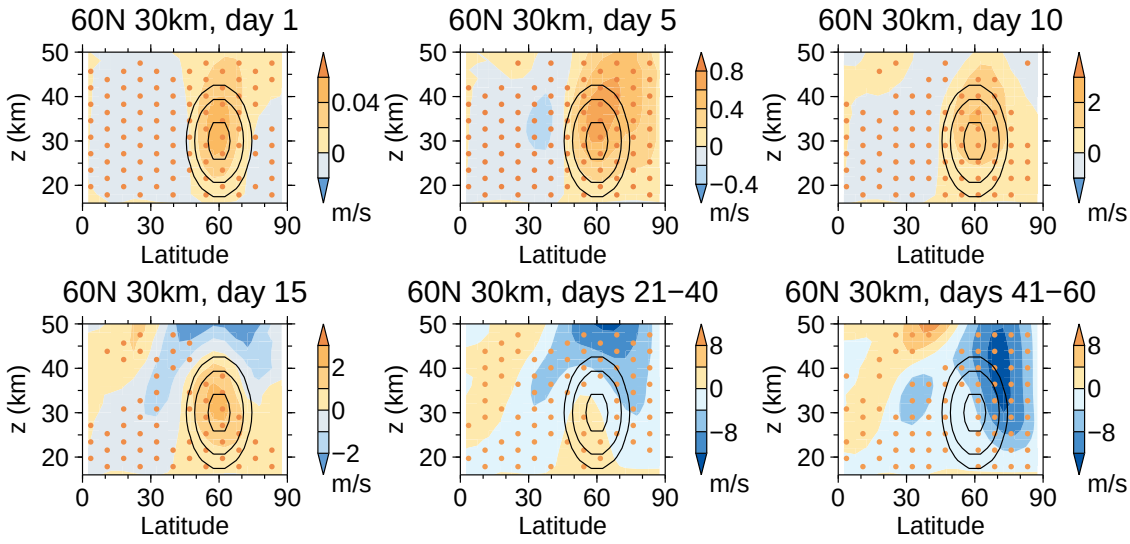

Figure 6.11: The ZMZW response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions with no NH extratropical Rayleigh friction, plotted in the same way as in figure 6.2, and showing similar results.

6.A Appendix: sensitivity to varying Rayleigh friction

In this appendix it is shown that removing Rayleigh friction in the NH does not qualitatively affect the simulated transient stratospheric response to an applied torque in the SMM under perpetual January conditions. Similar experiments were performed in runs with a seasonal cycle, and it was also found that weakening Rayleigh friction in the same way as described in section 5.A does not qualitatively alter the response in this case. Therefore the conclusions deduced from the results presented in this chapter are not affected by the Rayleigh friction in the SMM.

The perpetual January control run with no NH extratropical Rayleigh friction described in section 5.A was used, and 36 branch runs with the standard (ϕ_0=60N, z_0=30 km) torque were performed in the same way as described in section 6.1. Figure 6.11 shows the ensemble-mean transient ZMZW response as a function of latitude and height. This is very similar up to day 20 to that with full Rayleigh friction (figure 6.2) but a bit stronger after day 20.

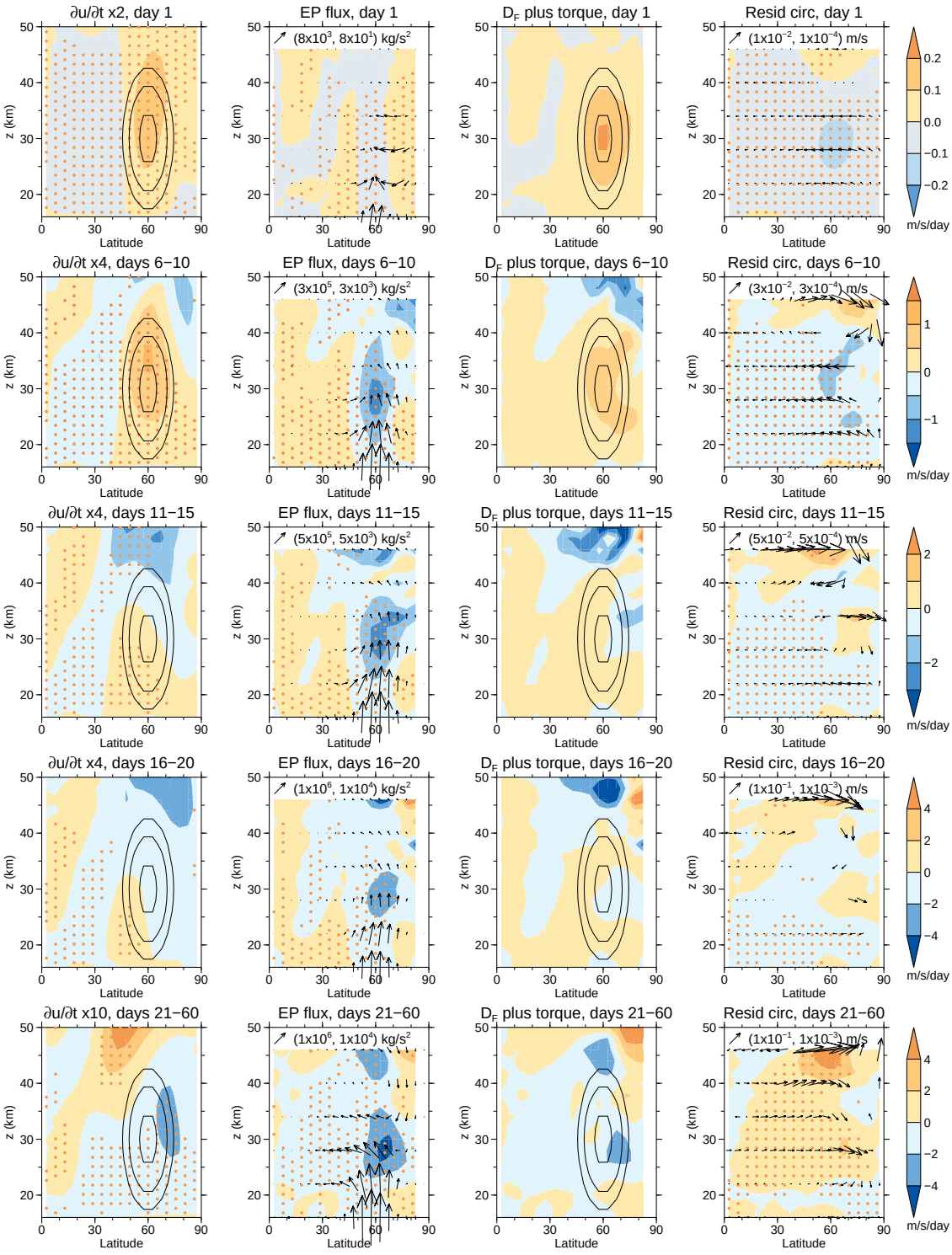

Figure 6.12: As in figure 6.3 but for the responses to the standard (ϕ_0=60N, z_0=30 km) torque with no NH extratropical Rayleigh friction, showing a similar response.

Figure 6.12 shows the ensemble-mean transient responses of the terms in the TEM zonal momentum equation in the same runs. Again the responses are very similar to those in runs with full Rayleigh friction (figure 6.3), though the D_F response is a bit stronger.

The responses to the (ϕ_0=60N, z_0=40 km) torque (not shown) are similarly not strongly affected by removing NH extratropical Rayleigh friction.

Therefore the simulated transient stratospheric responses to applied torques are not strongly affected by weakening NH extratropical Rayleigh friction.

Chapter 7

Evaluation of methods for deducing forcing mechanisms

In this chapter the stratospheric responses to imposed extratropical zonal torques in the UKMO Stratosphere-Mesosphere Model (SMM; section 3.3) that are presented in chapters 5 and 6 are examined further, in order to evaluate methods that could be applied to observational and model data to gain more information about forcing mechanisms[1]. Such methods can be thought of as operations on such data which return an answer about what the forcing mechanism is, with some degree of confidence. Since the forcing in the SMM experiments is known to be a torque with particular properties, it can be shown whether a particular method is successful at deducing information about the applied forcing from the response simulated by the SMM. A diagnostic that is useful for this purpose will at least return different values for different applied torques, so that in principle information is contained that allows for differentiation between torques. The diagnostic will be more useful if the results are related to the torque in a fairly straightforward way as well, else it may be necessary to calculate the values of the diagnostic for different possible forcings to use it practically. Methods which pass this test may then be applied with greater confidence in order to understand the HT relationship and other similar problems.

[1]The model data in mind are from runs with realistic forcings, including free-running control and historical simulations, as opposed to experiments in which various idealised forcings are imposed in order to test hypotheses (such as applying idealised equatorial wind profiles to test proposed mechanisms of the HT relationship), which form a complementary experimental design with its own strengths and weaknesses.

A considerable number of studies have used methods like composite differencing to infer information about forcing mechanisms [e.g. Hu and Tung, 2002; Sassi *et al.*, 2004; Naoe and Shibata, 2010; Calvo and Marsh, 2011; Yamashita *et al.*, 2011]. However, it was argued in chapter 4 that the tendency of the extratropical NH stratospheric response to external forcings to resemble the NAM on monthly time scales makes it difficult to use observational data and the quasi-steady state response to applied forcings in model runs to test hypotheses for forcing mechanisms. Some authors have previously recognised this problem and proposed different methods to use this data to understand forcing mechanisms. In this work the following three methods, in addition to using composite differences, are focussed on.

Firstly, Kodera [1995] argued that the observed NH extratropical stratospheric response to an applied forcing in the first month that the response becomes substantial is different for volcanic, QBO and solar forcings, and is indicative of the forcing mechanism in each case. His analysis was based on only thirteen winters' data, however, and since the response to the QBO phase in November appears very NAM-like in longer data records [Dunkerton and Baldwin, 1991; Ruzmaikin *et al.*, 2005, and chapter 4 of this thesis], and the solar cycle response does not seem statistically significant in early winter (section 2.7.3), it is not clear if this method is reliable.

Secondly, Kushner *et al.* [2001] proposed that the "residual" response following subtraction of the best-fit AM component of the circulation response to anthropogenic forcing can be thought of as the "direct" response to the forcing, and that this is informative about the nature of the forcing. Polvani and Kushner [2002] applied a similar method to try to understand how stratospheric circulation changes influence the troposphere, and Deser *et al.* [2004] showed the residual GPH response to sea ice and North Atlantic sea surface temperature changes is a change in GPH above the region of the applied forcing that is consistent with the expected direct temperature response to each applied forcing. However, this method could only be reliable as long as the feedbacks tend to act to make the "indirect" response very AM-like and the applied forcing does not have a pattern that is very similar to the AMs. Chen and Zurita-Gotor [2008] showed the residual response to a stratospheric torque in

a primitive equation model peaks near the location where the torque is strongest. However, stationary waves in their model were weak, and so feedback from changes in waves were also probably too weak – it was shown in chapter 5 that varying the strength of the tropospheric wave forcing can have a large effect on the response.

Thirdly, in chapter 4 it was proposed that the short-term transient response ought to be very informative about the applied forcing, but whilst this seems theoretically well-justified, the time scale in which the "feedback" part of the response becomes important is not known. It also has the disadvantages that it cannot be examined in the real atmosphere, it requires dedicated modelling experiments to be performed, and only information on the early stages of the forcing mechanism can probably be clearly obtained (though this can still be useful for testing hypotheses). An estimate of the time scale within which feedbacks become important may affect the interpretation of experiments examining the transient response of the atmosphere to applied forcings [e.g. Simpson *et al.*, 2009; Garfinkel *et al.*, 2012b, and chapter 8 of this thesis].

To my knowledge, Deser *et al.* [2004] and Chen and Zurita-Gotor [2008] presented the only results that suggest any of these methods can be used to work backwards from the measured response to correctly identify a known applied forcing, so it is uncertain whether any of these methods are generally reliable. The experiments presented in chapters 5 and 6 provide an opportunity to test whether these methods applied to the SMM data will give the correct answer when used to work backwards to infer the known applied torque, and indicate whether they could be reliable for learning about forcing mechanisms of interest in the atmosphere. The experiments examining the transient vortex response to the torques also indicate the time scale in which the "feedback" part of the response becomes important in the extratropical stratosphere.

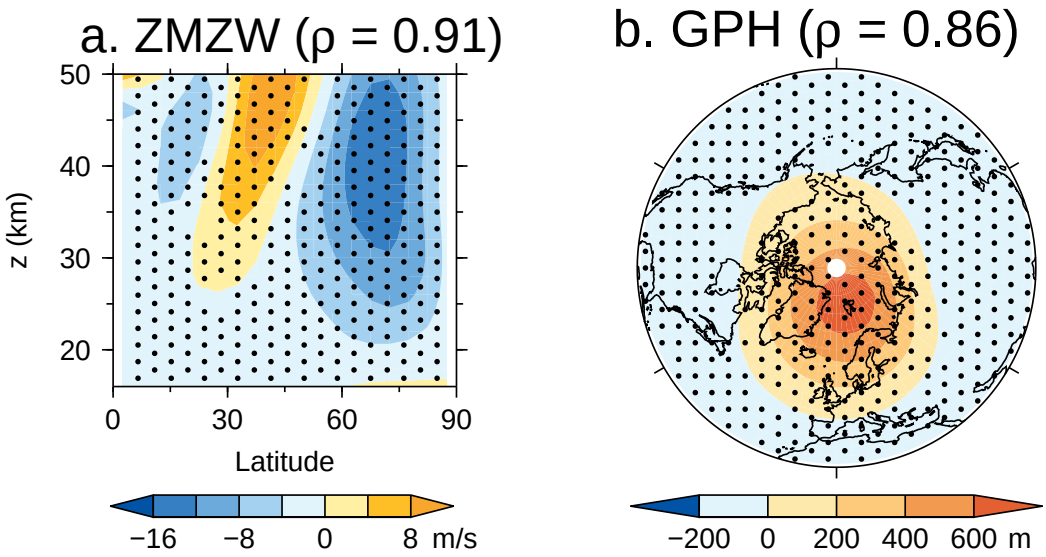

Figure 7.1: (a) shows the steady state ZMZW differences under perpetual January conditions between an SMM run with prescribed 16 km GPH equal to the HadGEM2-CCS 100 hPa January climatological GPH with eddy component multiplied by a factor of 2 and a similar run with the eddy component multiplied by a factor of 1.5. (b) shows the differences in 32 km GPH between the same pair of runs. Stippling shows where the differences are statistically significant above the 95% level. ρ values in the plot titles are the anomaly correlations between the differences and the NAM signatures (figures 5.3(a) and 5.3(b)).

7.1 Composite differences averaged over long time periods

The steady-state ZMZW response to a torque in the SMM in perpetual January conditions (figure 5.6) and the composite wintertime response in runs with a seasonal cycle (figure 5.11) have many similarities for torques with different positions and structures, meaning these data do not contain much information about the applied torque. The same is true of the GPH and EP flux responses as described in sections 5.2 and 5.3. So inferring the forcing from these data would be very difficult even if these responses to different forcings were well known.

This point is reinforced by figure 7.1, which shows the steady-state ZMZW and 32 km GPH differences between two different SMM runs without torques. The first is the "standard" perpetual January control run with 16 km GPH specified as the

January climatological 100 hPa GPH in HadGEM2-CCS, with eddy component multiplied by 2 (section 5.1.1). The second is a similar run with the eddy component of HadGEM2-CCS 16 km GPH multiplied by 1.5. In other words, these differences are due to a steady forcing by a 16 km GPH perturbation equal to a quarter that shown in figure 5.1(b), which may represent a forcing on the vortex due to a change in the tropospheric state. The ZMZW differences appear very like the NAM ZMZW signature and the steady-state ZMZW responses to torques applied at 60N or poleward (figures 5.3(a) and 5.6), with an anomaly correlation of 0.91 with the NAM signature (calculated in the same way as in section 5.2.1). The GPH differences resemble the NAM 32 km GPH signature (figure 5.3(b)) much more than they do the GPH forcing pattern in figure 5.1(b), and have an anomaly correlation of 0.86 with the former. This further illustrates that the atmospheric response to a forcing on long time scales can be very different to the direct effect of the forcing, and can be strongly influenced by AM dynamics. It would be very difficult to confidently conclude whether a given long-term mean response like that in figure 7.1 is due to a direct torque or tropospheric forcing, even though the forcings have very different characters.

7.2 Composite differences in first month that response is substantial

Figure 7.2 shows the November mean ZMZW responses to applied torques that were centred on different positions in the NH stratosphere in SMM runs with a seasonal cycle (described in sections 5.1.1.2 and 5.1.2). The responses are defined as composite differences between runs with the torques applied between October 22 and April 10 and a control run with no torque, as described fully in chapter 5. To evaluate the method of Kodera [1995], the responses in November are examined here. Statistical significance of the differences is calculated in the same way as in section 5.3.

The position of the maximum November ZMZW response shown in figure 5.11 is in the vicinity of the torque location in all cases, and appears different for torques at different positions, so the necessary condition that these responses should be different

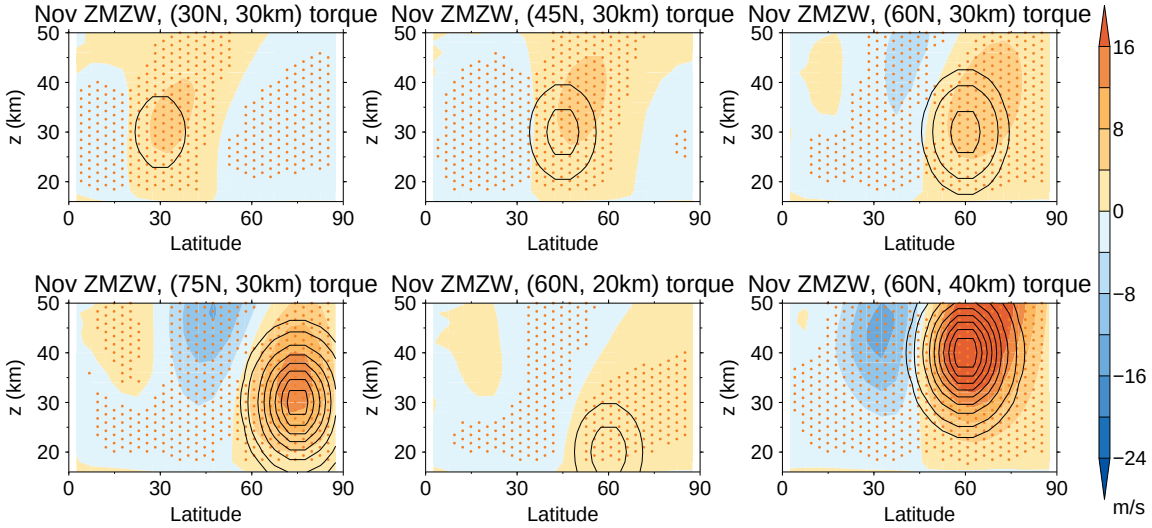

Figure 7.2: The November mean ZMZW responses to applied torques centred at latitude ϕ_0 and height z_0, indicated in the plot titles, in runs with a seasonal cycle (colours). The applied torque is shown by contours at $\pm 0.5\,\mathrm{m/s/day}$, then every integer value in units m/s/day, with negative contours dashed. The torques have meridional scale $\delta\phi = 11°$, vertical scale $\delta z = 10\,\mathrm{km}$, and a strength such that the magnitude of the total zonal momentum added to the NH equals that of the standard (ϕ_0=60N, z_0=30 km) torque with peak strength 2.5 m/s/day. Stippling shows where the responses are statistically significant above the 95% level.

if information about the torque location is to be deduced seems to be satisfied, in support of Kodera [1995]. However, the response is not simply a ZMZW increase where the torque is applied, so prior knowledge of the response to each torque is required to make an accurate deduction from any individual response.

On the other hand, in experiments examining the transient response to an applied torque with perpetual January boundary conditions, it was found that the stratospheric response becomes NAM-like in a time scale of ∼2–3 weeks (e.g. figure 6.2), due to the feedback of waves responding to changes in the zonal mean circulation being stronger in these simulations than in those with a seasonal cycle. The fact the QBO-E minus QBO-W composite differences in the real atmosphere are very NAM-like in November (chapter 4) suggests these feedbacks could also be substantially stronger in the real atmosphere than in the SMM simulations with a seasonal cycle. The date at which external forcings start to have a substantial influence on the vortex

is difficult to know to a precision of 2–3 weeks, so it is not clear that this method could be relied on to give information about forcing mechanisms.

7.3 Residual following subtraction of AM component

Figure 7.3 shows the "residual" of the ZMZW responses to applied torques in perpetual January simulations, calculated by subtracting the "best-fit" NAM signature from the full responses shown in figure 5.6. Using a similar method to Deser *et al.* [2004] and Chen and Zurita-Gotor [2008], the best-fit NAM signature was defined as the NAM signature scaled by a constant whose value was chosen to minimise the sum over the gridpoints north of 20N of the squared residual ZMZW values. The residuals are small for responses to torques with a high anomaly correlation with the NAM signature (figure 5.6). The residual has the same sign as the torque where the torque peaks in every case. It is also negative in the regions to the south and north of where the torque is applied in most cases. The residuals for torques with different meridional and vertical scales are similar, however, indicating that the residual does not contain information about these parameters.

These results support the idea of Kushner *et al.* [2001] in so far that the residual displays information that relates more clearly to an applied forcing than does the total response, since the residual has a relatively large value with the same sign as the torque near where the torque peaks. The parts of the residuals to the north and south with opposite sign to the torque may be understood as being due to the Coriolis force acting on an induced residual circulation (section 2.2.3).

The residuals are similar if the best-fit NAM signature is chosen to minimise the sum of the squared ZMZW weighted by the horizontal areas of the grid boxes, or to minimise the sum of residual zonal mean zonal momentum values.

On the other hand, the GPH residual contains features that are not directly related to the applied torque, and which may give a misleading impression of the character of the applied forcing. Figure 7.4(a) shows the residual of the response in GPH at 32 km

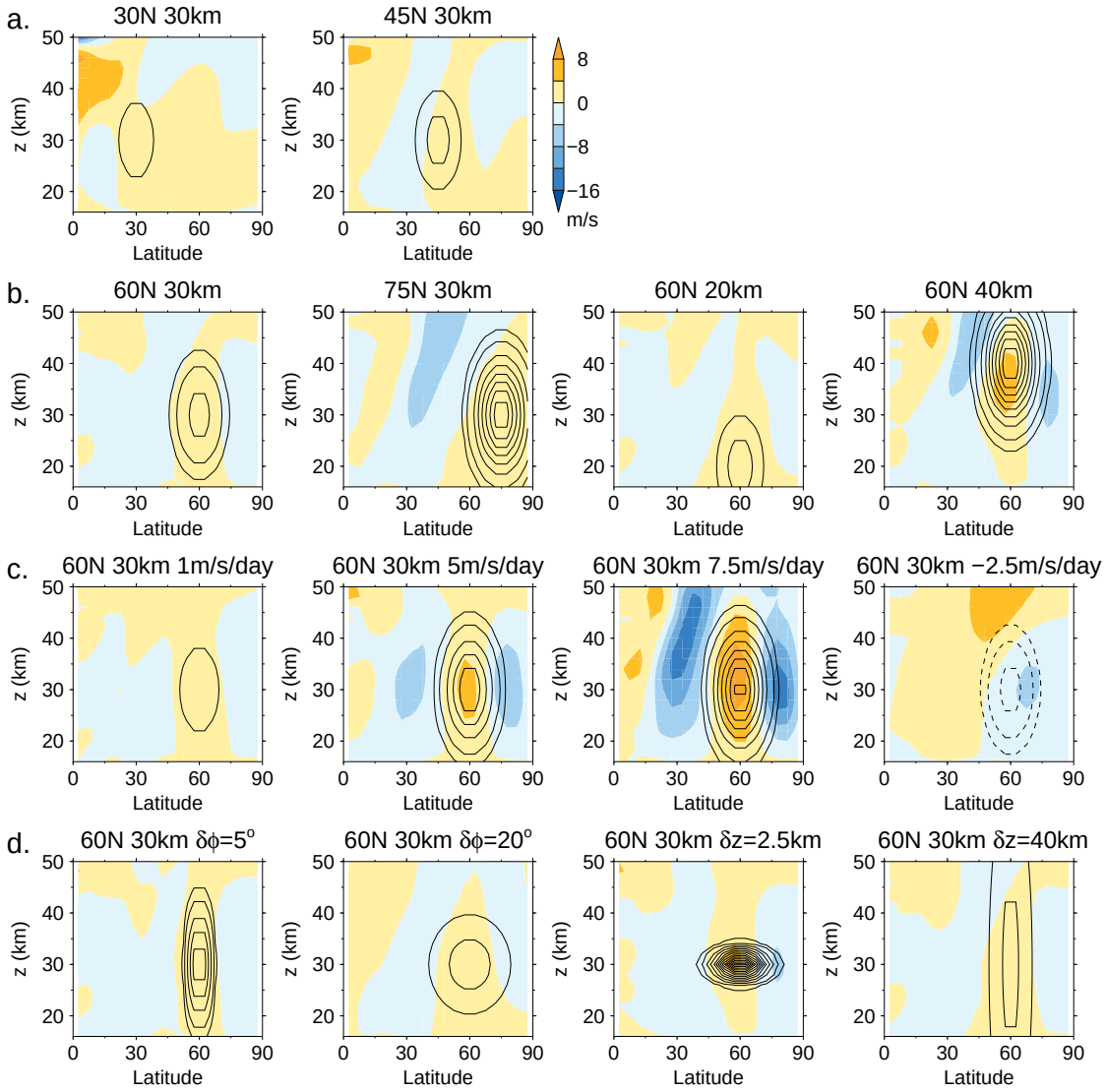

Figure 7.3: The "residual" of the steady state ZMZW responses to applied torques under perpetual January conditions (colours), calculated by subtracting from the ZMZW responses the NAM signature multiplied by a constant chosen so that the sum of the squared residual values north of 20N is minimised. Data are shown for the same torques as in figure 5.6. (a) shows the residual for torques centred in the mid-latitude mid-stratosphere, (b) the residual for torques centred in the high-latitude stratosphere, (c) the residual for torques centred on (60N, 30 km) with different peak strengths and (d) the residual for torques with different meridional and vertical scales. Contours show the applied torque with levels at ± 0.5 m/s/day, then every integer value in units m/s/day, with negative contours dashed.

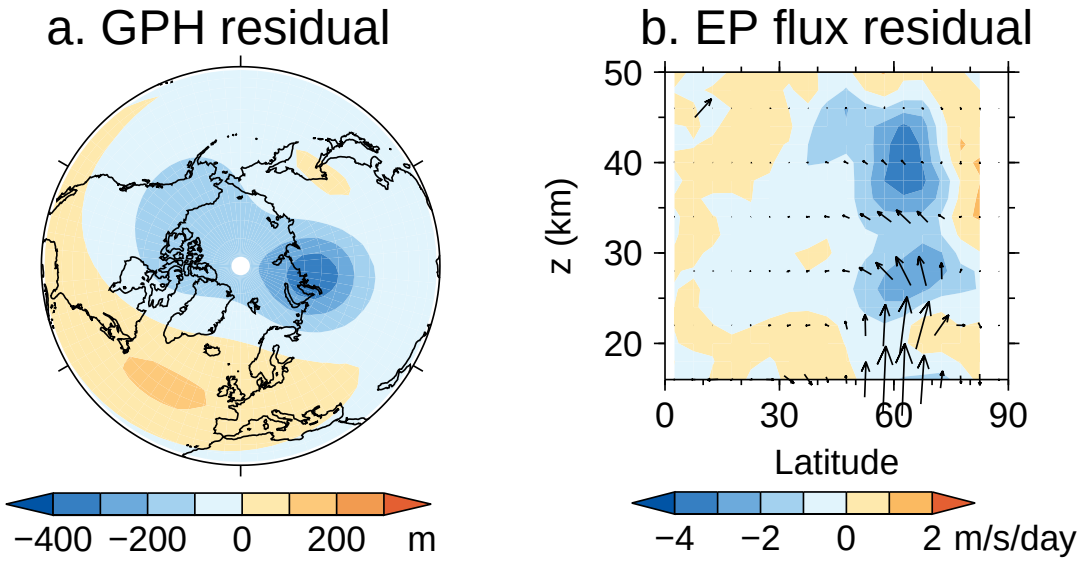

Figure 7.4: (a) Residual of the steady state 32 km GPH response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions, shown north of 20N. This is calculated by subtracting from the total GPH response the NAM signature multiplied by a constant chosen so that the sum of the residual values multiplied by cos(ϕ) and then squared north of 20N is minimised. (b) The residuals of the responses of the EP flux (arrows) and D_F (colours) to the same torque. These are calculated by subtracting from the total responses the NAM signature multiplied by a constant chosen so that the sum of the squared D_F residual values between 22.5–82.5N is minimised. The arrow in the top left corner of (b) has a length a flux $(10^6, 10^4)$ kg/s^2 would have. EP fluxes are plotted every 6 km in height, and D_F data at 87.5N are not plotted as differentiation error causes them to be spuriously large.

to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions, calculated by subtracting the best-fit NAM signature from the total response (shown in figure 5.9(a)). The best-fit NAM signature is defined in a similar way as for the ZMZW residuals, but with the sum of squared products of the GPH residuals with the cosine of the latitude minimised (the results are similar if the residuals at each grid point have equal weight). (The total response shown in figure 5.9(a) is very NAM-like.) The residual is dominated by a trough centred near 80E and 65–70N. The residual of the 50 km GPH response is smaller and has a zonal wavenumber 2 pattern (not shown). These results are similar for the other torques centred at 60N or polewards, and are not obviously to be expected to be part of the direct response to an applied torque.

Figure 7.4(b) shows the residual EP flux and D_F response to the standard (ϕ_0=60N, z_0=30 km) torque under perpetual January conditions following subtraction of the best-fit NAM signature, which is defined as the NAM signature scaled by a constant whose value was chosen to minimise the sum over the gridpoints between 22.5–82.5N of the squared residual D_F values (the data at 87.5N were not used as differentiation error causes them to be spuriously large). The total responses are shown in figure 5.9(b). The residual includes the upward EP flux response in the lower stratosphere between ∼50–70N, and negative D_F in the high latitude middle and upper stratosphere. This appears to capture the direct effect of the torque on wave propagation, as it increases the high-latitude lower stratospheric meridional PV gradient, which enhances vertical propagation of planetary waves into the stratosphere (section 6.4). The residuals for other torques centred at 60N or poleward are similar.

It was shown in chapters 5 and 6 that a major component of the feedback part of the response to a high-latitude torque is increased convergence of EP flux where the torque is applied that largely cancels the direct effect of the torque. This may not be visible in the observed response to a torque which is itself caused by an externally forced perturbation to the EP flux. Therefore it was checked that the EP flux and D_F residuals for each torque are not very different if the torque is firstly added to D_F in order to remove this effect before the calculation of the residual is done. The results are also not sensitive to defining the best-fit NAM component by minimising the sum of squared residuals of F^ϕ and F^z (either with F^z multiplied by a scaling factor of 111.1, which is the number of kilometers per degree of latitude, or with the fluxes scaled according to acceleration scaling as defined by Gray *et al.* [2003]).

In the SMM simulations with a seasonal cycle, the ZMZW responses (figures 5.11 and 7.2) do not tend to closely resemble the NAM signature, so the residuals following subtraction of the best-fit NAM ZMZW signature calculated in the same way as in section 5.2 (not shown) appear largely similar to the overall ZMZW responses. As in the case of the perpetual January experiments, the residual is always positive where the torque peaks. However, there is also a lot of structure which is not directly

identifiable with the torque. In these experiments the residual does not seem to contain more useful information about the applied forcing than the November composite responses discussed previously. Similarly, the residuals of the GPH and EP flux responses in the simulations with a seasonal cycle are not any more clearly related to the applied torque than the total responses.

Overall, the results from the perpetual January simulations suggest that using the residual part of the response to infer a forcing mechanism may be useful when feedbacks are strong enough to make the response NAM-like, because this is more clearly related to the direct effect of an applied torque than is the total response. In this case it also seems that the total response can be decomposed reasonably well into a "direct response" and a NAM-like feedback part. However, in such situations the residuals will also be small and perhaps statistically uncertain if short time series are used. The GPH residual in these experiments also has features that are hard to relate to the applied torque, and suggests the residual could be misleading regarding the nature of the applied forcing, if the responses to given forcings are not known beforehand. The residual was not found to be useful in experiments with a seasonal cycle, for which the response to a torque is not NAM-like.

7.4 The short-term transient response

The simulations with applied torques branched from control runs as explained in section 6.1 are analysed here to examine the information about an applied torque contained in the short-term transient response. Figure 7.5 shows the responses to torques with different positions and different meridional and vertical scales in runs with a seasonal cycle after one day. The responses in the perpetual January simulations (not shown) are very similar. There is clearly increased ZMZW where the torque is strongest in each case. The decreased ZMZW to the north and south and increased ZMZW above can be understood as due to the Coriolis force acting on the induced residual circulation (section 6.3). Therefore the position of the peak torque can be inferred quite easily from these data. The ZMZW response is also more meridionally

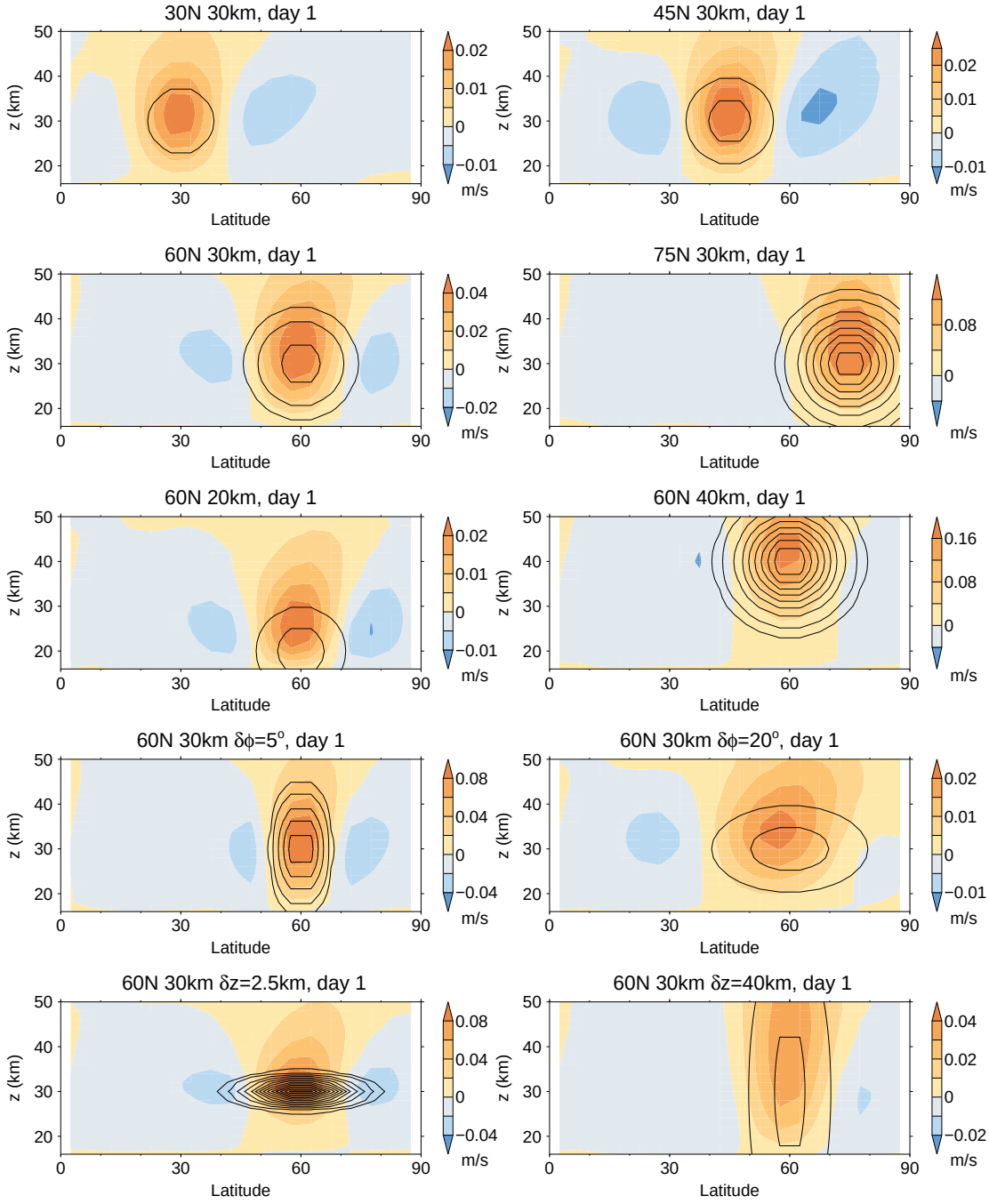

Figure 7.5: The transient ZMZW response after one day to applied torques in runs with a seasonal cycle (colours). The applied torque on day 1 is shown by contours at $\pm 0.05\,\mathrm{m/s/day}$, then every integer multiple of $0.1\,\mathrm{m/s/day}$, with negative contours dashed. Plot titles indicate the latitude ϕ_0 and height z_0 where the torque is strongest. Unless otherwise specified in the title, the torques have meridional scale $\delta\phi = 11°$, vertical scale $\delta z = 10\,\mathrm{km}$, and a magnitude such that the total zonal momentum added to the NH equals that of the standard (ϕ_0=60N, z_0=30 km) torque with peak magnitude $0.25\,\mathrm{m/s/day}$ after one day. The ZMZW differences are related quite simply to the spatial structure of each torque.

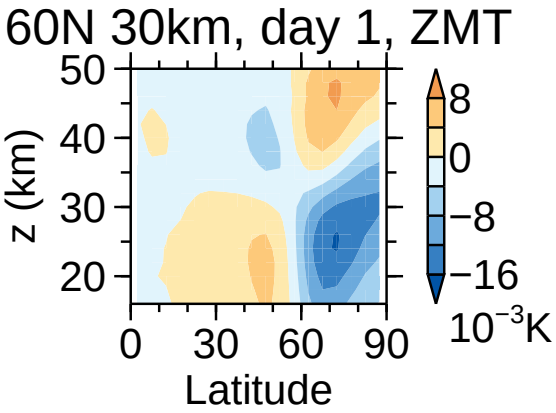

Figure 7.6: The transient ZMT response after one day to the standard (ϕ_0=60N, z_0=30 km) torque in runs with a seasonal cycle.

extended for torques with greater meridional scales and more vertically extended for torques with greater vertical scales, so there is information about these torque parameters. After a few days the responses for different vertical and meridional scales are quite similar, however (not shown). The responses to torques centred at (60N, 30 km) with magnitudes in the range 0.1–0.75 m/s/day after one day scale very linearly with the magnitude (not shown), so the magnitude could also be deduced.

The short-term transient response also shows that the applied forcing is a torque and not something else such as heating or a change in bottom boundary GPH. The ZMT response to the standard (ϕ_0=60N, z_0=30 km) torque after one day in the runs with a seasonal cycle (figure 7.6) shows cooling of the polar lower and middle stratosphere and heating above, with an oppositely signed response in mid-latitudes. If this were the direct response to an applied heating, it would be associated with a residual circulation response that would partially cancel the heating, mostly by adiabatic vertical motion [Plumb, 1982]. Thus the residual circulation response would be downward in the high latitude lower stratosphere and upward above, and vice versa in mid-latitudes. However, this is opposite to the response shown in the top row of figure 6.6, so the applied forcing cannot be a heating. The GPH response (not shown) is very zonally symmetric on day 1, so there is no indication that the response could be due to a change in tropospheric wave forcing. A direct forcing of the residual circulation to match that shown in figure 7.6 would cause deceleration of the lower

and middle stratospheric ZMZW rather than the acceleration where the torque peaks shown in figure 7.5, leaving a direct zonal forcing in the TEM zonal momentum equation (equation 2.13) as the only possibility.

The fact that the response may become NAM-like in 2–3 weeks, as found in the perpetual January simulations (section 6.2), implies that in order to infer the direct effect of a forcing on the stratosphere using the transient response, time scales considerably shorter than this ought to be examined, and it may be necessary to examine even shorter time scales in the troposphere as dynamical time scales there are shorter. Simpson *et al.* [2009] and Garfinkel *et al.* [2012b] presented the transient response to a forcing averaged over the first 10 and 15 days respectively, and it seems possible given the results presented here that this averaging period may be too long to clearly see the direct influence of each forcing (though personal communication from C. Garfinkel indicates the response on shorter time scales is not qualitatively different from that averaged over the first 15 days in the experiments of Garfinkel *et al.* [2012b]).

7.5 Summary and conclusions

Results have been presented from experiments where a known zonal torque has been applied in the SMM in order to understand how the simulation output can be processed to correctly infer that the forcing is a torque, and to infer properties such as its spatial structure and magnitude. This improves understanding of what methods can be used to learn about unknown physical processes behind the statistical relationships between measurements in different parts of the atmosphere. These experiments relate directly to studies of the response of the extratropical stratosphere to external forcings, but may also be informative about what methods can help understanding of mechanisms linking other parts of the climate system.

These results support the argument given in chapter 4 that composite differences, or other diagnostics of the total long-term mean response to a forcing, do not contain

much information about the applied forcing, since these are very similar for different applied forcings (section 7.1). As well as suggesting that such measurements do not provide much information about the mechanism behind the HT relationship (chapter 4), this also suggests that it may be worth revisiting the problem of how other external forcings affect the extratropical stratosphere. For example, Sassi *et al.* [2004], Garcia-Herrera *et al.* [2006] and Manzini *et al.* [2006] argued, based on composite differences between different ENSO phases, that El Niño events cause weakening of the stratospheric polar vortex by affecting the planetary wave structure in middle latitudes, thereby increasing upward EP flux into the polar stratosphere. However, it is not clear if this increase causes the circulation changes or is a response to them. Whilst there is no reason to think that these authors' explanation of the mechanism is wrong, there could be other effects of El Niño that are also important, such as on GPH in the high-latitude Pacific as identified by Garfinkel *et al.* [2012a], and which may be revealed using a different experimental method.

The other three methods tested above showed at least some improvement over using long-term mean responses at being useful for inferring known applied torques, given the responses to those torques in the SMM. Given that the mechanisms by which natural forcings affect the vortex are likely to be more complicated than the fixed torque tested here, and the SMM simulations do not include feedbacks from changes in gravity waves and the troposphere, each of these methods may be expected to perform less well if applied to gain information about a real forcing mechanism of interest. It is not clear if feedbacks in the real atmosphere or in more realistic atmospheric models are weak enough for the method, suggested by Kodera [1995], of using composite differences in the first month in which the response becomes substantial to work (section 7.2). The method of using the residual part of the response proposed by Kushner *et al.* [2001] may be misleading about the nature of the forcing in some respects, and was not found to be useful in the experiments with a seasonal cycle (section 7.3), so it is difficult to have confidence in this method as well. Examination of the short-term transient response to a forcing, on the other hand, works robustly in these experiments in both the perpetual January and seasonal cycle simulations,

allows forcings of different types to be distinguished and contains information about all the torque parameters (section 7.4). (However, as noted in the introduction to this chapter and at the end of chapter 4, when considering realistic forcings that act on the extratropical stratosphere such as the QBO, the most important effects in the short-term transient response are not necessarily the most important for determining the long-term response[2]. The short-term transient response also cannot be determined from observations and requires performing experiments in models, which will not simulate the response perfectly.)

Altogether, this indicates that none of the methods tested here can unambiguously show the mechanisms by which different parts of the atmosphere interact, and that combining experimental approaches is likely to be necessary to come to firm conclusions. The methods of Kodera [1995] and Kushner *et al.* [2001] do not clearly give reliable information about forcing mechanisms. Examination of the short-term transient response to a forcing in a model may give a useful signal, and therefore this is focussed on in chapter 8 to better understand the mechanism behind the HT relationship.

To my knowledge, using the transient response of the atmosphere to a forcing on time scales of a few days to understand a forcing mechanism is novel and may be of use in solving various other outstanding problems, such as understanding the downward influence of the stratosphere on the troposphere [e.g. Baldwin and Dunkerton, 2001], the dynamical influence of the solar cycle [e.g. Labitzke, 2005] or of ENSO, as discussed above.

[2]For example, diabatic effects that take time to become apparent were found by Polvani *et al.* [1995] to have an important influence on the formation of the stratospheric surf zone.

Chapter 8

The transient response of the vortex to QBO forcing

This chapter presents the results of experiments examining the short-term transient response of the vortex to nudging tropical stratospheric winds towards a QBO-E state in the HadGEM2-CCS GCM (section 3.4). As argued in chapters 4 and 7, this is expected to be much more informative about the forcing mechanism behind the HT relationship than QBO-E minus QBO-W composite differences or the transient response on time scales of two weeks or more that have been presented in previous studies. This chapter includes and expands upon part of the analysis of Watson and Gray [2014].

8.1 The Holton-Tan relationship in HadGEM2-CCS

Figure 8.1 shows the QBO-E minus QBO-W composite difference in the January–February mean ZMZW in ERA-40 (section 3.1) and in the 240-year pre-industrial control run of HadGEM2-CCS (section 3.4) – the monthly-mean difference is not statistically significant earlier in winter in HadGEM2-CCS. For analysis of ERA-40, the QBO is defined as being in its easterly or westerly phase when the 5S–5N November–February mean ZMZW is easterly or westerly respectively at 50 hPa, as in section 4.1. In the HadGEM2-CCS data, the correlation between the November–February mean ZMZW averaged over 5S–5N and that at (60N, 10 hPa) is greatest for equatorial winds at 30 hPa, so the sign of the 5S–5N November–February mean ZMZW on this level

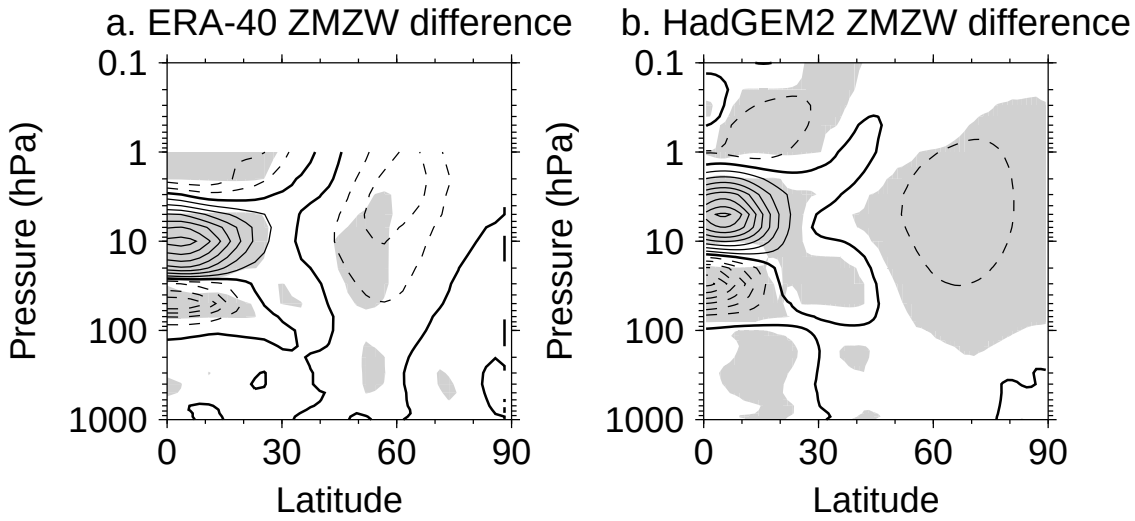

Figure 8.1: QBO-E minus QBO-W composite differences of January–February mean ZMZW in ERA-40 (a) and in HadGEM2- CCS (b). HadGEM2-CCS reproduces the differences in ERA-40 reasonably well, with smaller high-latitude differences. Note the equatorial ZMZW on different pressure levels is used to define the QBO phase in ERA-40 and HadGEM2-CCS. Contour levels are every 4 m/s with negative contours dashed and the zero contour thickened. Shading shows where differences are statistically significant above the 95% level.

is used to define the QBO phase for analysis of model data, which seems the fairest way to compare the modelled and observed HT relationships. The model reproduces the weakening of the vortex seen in observations, with ZMZW differences that are somewhat smaller at high latitudes. The QBO-E minus QBO-W composite differences in stratospheric GPH in January and February (not shown) are very NAM-like, in agreement with the differences in ERA-40 (section 4.1). The pre-industrial control run excludes forcing by anthropogenic emissions, volcanoes and the solar cycle so that the influence of the QBO is more easily distinguished than in runs with historical forcings over 1860–2005, which show a similar but weaker HT relationship (not shown).

Shading in figure 8.1 shows where the QBO-E minus QBO-W differences are statistically significant above the 95% level, calculated using an MC permutation test in the same way as in section 4.1. The same test is used for the EP flux differences below.

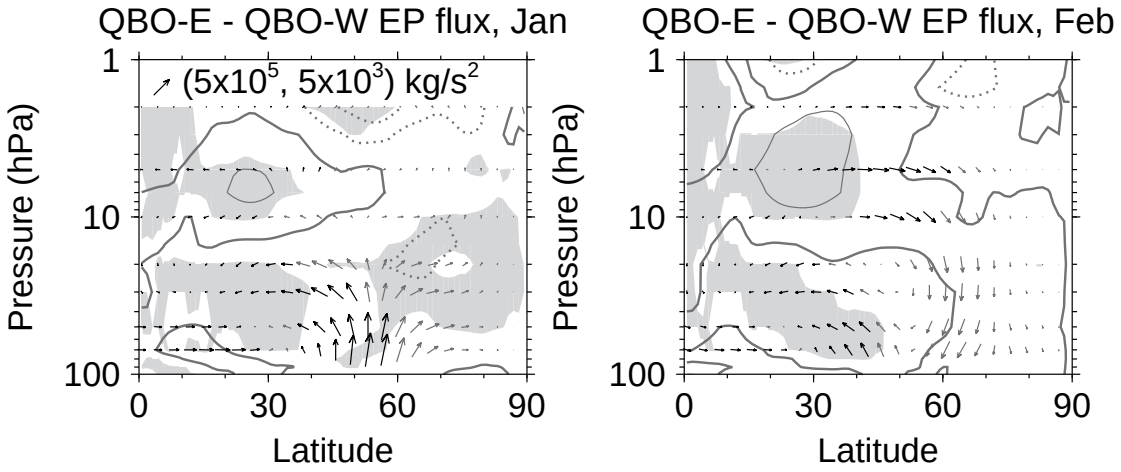

Figure 8.2: As in figure 4.1 but for HadGEM2-CCS, showing only January and February when the ZMZW differences (figure 8.1) are statistically significant. In these months HadGEM2-CCS reproduces the pattern of QBO-E minus QBO-W composite EP flux differences in ERA-40 within observational error.

The lack of an HT relationship in November and December in the model may be due to the equatorial winds simply having too weak an influence on the vortex, and it taking time for their impact to accumulate and give rise to an appreciable vortex response, following the suggestion of O'Sullivan and Young [1992]. It could also be related to the model exhibiting slightly less total variability than in observations in early winter [Osprey *et al.*, 2013], which is a common problem in stratosphere-resolving GCMs. As noted in section 2.7.1, few modelling studies have reported an HT relationship in early winter. Nevertheless, the mechanism by which the equatorial winds influence the vortex in the model is likely to be qualitatively similar to the mechanism in the real atmosphere.

Figure 8.2 shows the QBO-E minus QBO-W composite difference in the monthly mean NH EP flux in January and February in the model. In these months the model reproduces the pattern of the observed influence of the QBO on the EP flux (figure 4.1) within sampling error, with greater upward flux in the high-latitude stratosphere in January in QBO-E and more poleward and downward flux in February. The EP flux differences are somewhat smaller than in ERA-40, however, consistent with the ZMZW differences being smaller.

8.2 Experimental design

A "climatological tropical wind" (ClimEq) control run was performed, which was set up identically to the 240-year pre-industrial control run except that the zonal wind in the tropical stratosphere was nudged towards the ERA-Interim (section 3.2) monthly mean climatology between January 1979 and December 2010 (with the climatology at each model time step calculated by linear interpolation between the middle of each month). ERA-Interim is likely to have a better representation of the stratosphere above 10 hPa than ERA-40 [Simmons *et al.*, 2007b; Dee *et al.*, 2011], so it is used in preference to create target equatorial zonal wind profiles. The climatology of this run is examined in section 8.3. 120 "QBO-E" branch runs of length one month were performed, taking initial conditions at January 1 and Feburary 1 of 60 different years from the ClimEq run (the first two years of this run were not used to allow the model to adjust to the nudging). In these runs the zonal wind in the tropical stratosphere was nudged towards the ERA-Interim climatology plus a typical QBO-E profile. The QBO-E profile was taken as the mean 3D zonal wind anomaly of the 30 months in ERA-Interim with the most negative anomalies in the 5S–5N mean ZMZW at 30 hPa, multiplied by a factor of 3 in order to raise the signal to noise ratio of the vortex response (figure 8.3). The equatorial ZMZW anomaly does not become larger than that in observed QBO-E phases in the time scale of 8 days considered in section 8.4, so the multiplication by 3 just affects the rate at which equatorial winds adopt a QBO-E profile. This method is similar to that used by Garfinkel *et al.* [2012b], but importantly the focus in this experiment is the vortex response at shorter times after nudging towards the QBO-E profile is begun.

Nudging was carried out between 21.25S and 21.25N, and was implemented by subtracting $\alpha(\phi)(u-u^T)\Delta t$ from the change in the zonal wind calculated at the end of each time step at each gridpoint, where u is the zonal wind, u^T is the target nudging profile, ϕ is the latitude and Δt is the model time step. The nudging parameter $\alpha(\phi)$ is given by

$$\alpha(\phi) = \frac{1}{20\,\text{days}}\, e^{-2(\phi/16^\circ)^2}$$

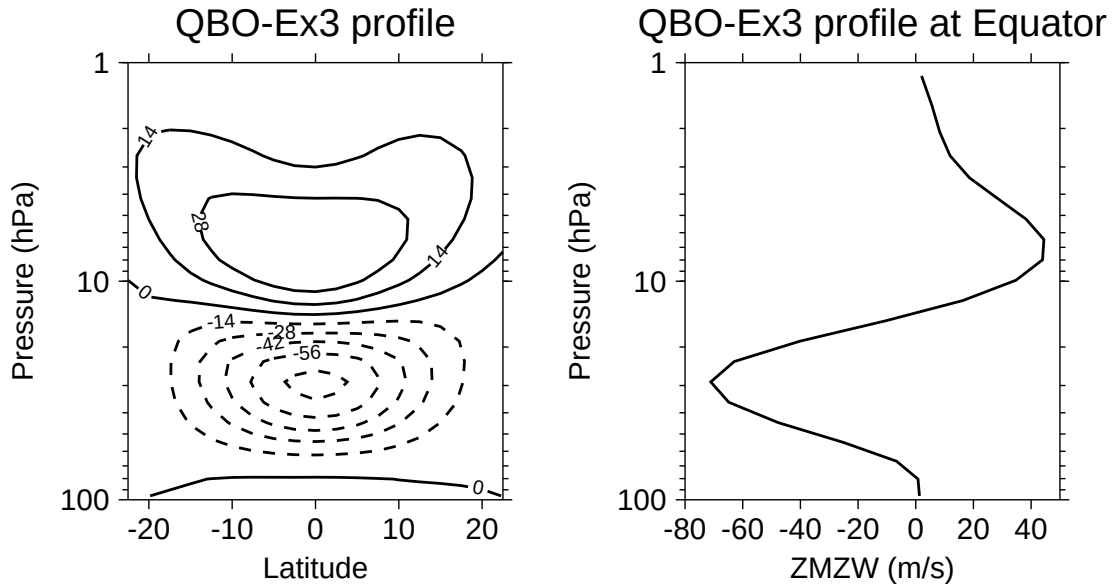

Figure 8.3: Left: the zonal mean of the target "QBO-E times 3" zonal wind profile used in the GCM experiments with nudged equatorial winds. Right: the 5S–5N mean of this profile.

between heights of 17.4–39.1 km (corresponding to pressure range 3.3–84.3 hPa). At one model level below and above this range (16.3 km, 103 hPa and 40.9 km, 2.6 hPa), α was set to one half its value within the range and is zero at other heights. The nudging code was adapted from that developed as part of the United Kingdom Chemistry and Aerosols project [Telford *et al.*, 2008][1]. The vertical profile of the 5S–5N mean ZMZW differences between the QBO-E runs and the ClimEq run is very similar to that in figure 8.3 between about 3–100 hPa over the first 8 days, although the meridional width of the nudged QBO-E winds is only 2/3 that of the target profile due to the weakness of the nudging away from the Equator (not shown) – this would most likely cause the influence on the vortex to be weaker than for a perfect imposed QBO-E profile [O'Sullivan and Young, 1992].

[1]Paul Telford and Mohit Dalvi provided some assistance with implementing nudging in HadGEM2-CCS.

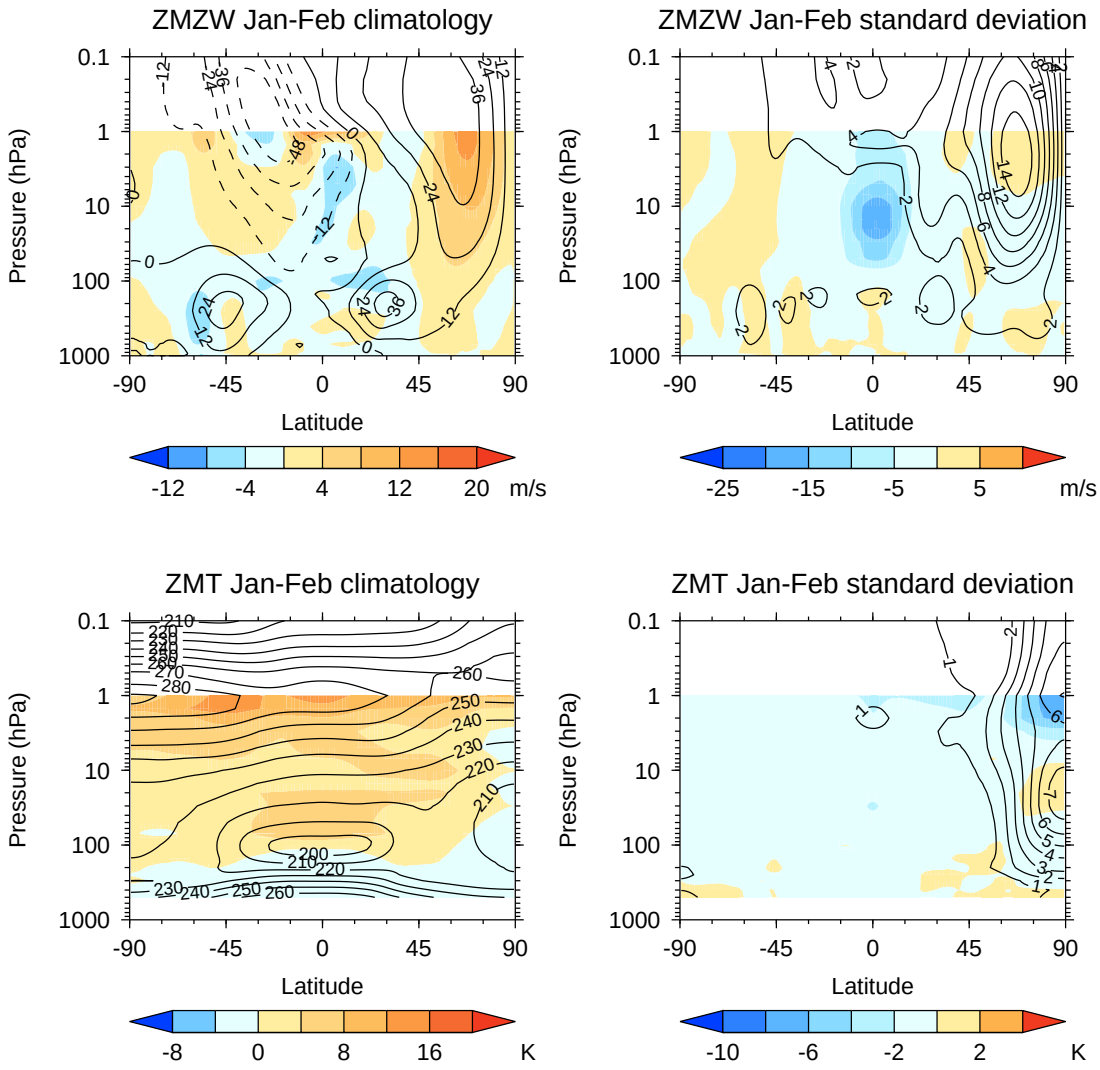

Figure 8.4: (a) shows the ZMZW January–February mean climatology of the ClimEq run (contours) and differences from the ERA-Interim climatology (colours). (b) shows the standard deviation of the model ZMZW (contours) and differences from the ERA-Interim standard deviation (colours). (c) and (d) show the same for ZMT. The model has a realistic extratropical stratosphere, with the QBO eliminated in the tropical stratosphere.

8.3 Climatology of nudged control run

Figure 8.4 shows the January–February mean climatological ZMZW and ZMT and their standard deviations in the ClimEq run from the third year onwards, and the differences in these from those of ERA-Interim. The polar vortex is stronger than in ERA-Interim by up to $\sim$10 ms^{-1} at 10 hPa, but the standard deviation of polar ZMZW is very similar in both. Equatorial ZMZW is up to $\sim$10 ms^{-1} too easterly, and variance in this region is much less due to the nudging eliminating the QBO. The stratosphere is generally warmer by $\sim$5 K except in the Arctic where differences are much smaller and the standard deviation of temperature is similar. This may be due in part to the model having prescribed pre-industrial stratospheric ozone, which is greater than the ozone amount in recent decades. The troposphere is colder than in ERA-Interim, by $\sim$5K below the tropical tropopause. These differences are similar to those in the 240-year pre-industrial model run in the extratropics (not shown). The vortex is stronger by a few metres per second with nudging applied, so the nudging has not greatly changed the extratropical circulation. Therefore the nudged model has a realistic representation of NH mid- and high-latitude circulation and equatorial winds, and temperatures that are reasonably close to the ERA-Interim climatology.

8.4 The transient response of the vortex to QBO-E forcing

Figure 8.5 shows the ensemble mean ZMZW difference between the QBO-E branch runs and the ClimEq run on a time scale of about a week, which is comparable to the mid-stratosphere's dynamical time scale, sub-divided into two-day intervals. The QBO-E pattern of winds is visible in the tropics, with mid-latitude anomalies that arise due to the Coriolis force acting on the QBO meridional circulation. The extratropical winds in the lower stratosphere initially strengthen slightly. However, between days 5–6 following the start of the nudging there is a weakening of the high-latitude upper stratospheric ZMZW, and between days 7–8 there is a weakening of

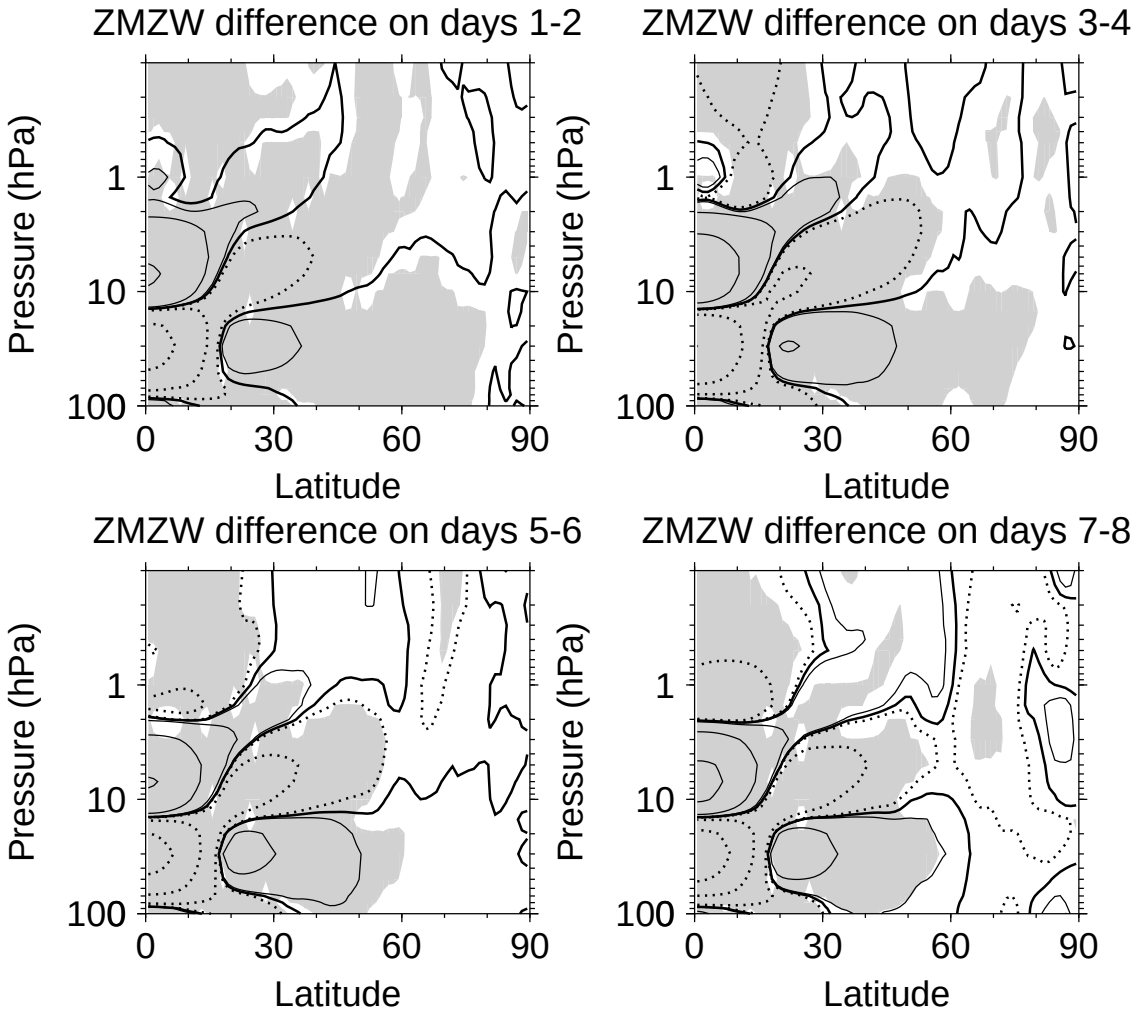

Figure 8.5: Ensemble mean ZMZW differences between branch runs nudged towards the QBO-E equatorial wind profile and the ClimEq run averaged over each two day interval up to 8 days following branching. Contours are at 0, ± 0.02, ± 0.2, ± 2 and $\pm 10\,\mathrm{m/s}$ with negative contours dotted and the zero contour thickened. Shading shows where differences are statistically significant above the 95% level.

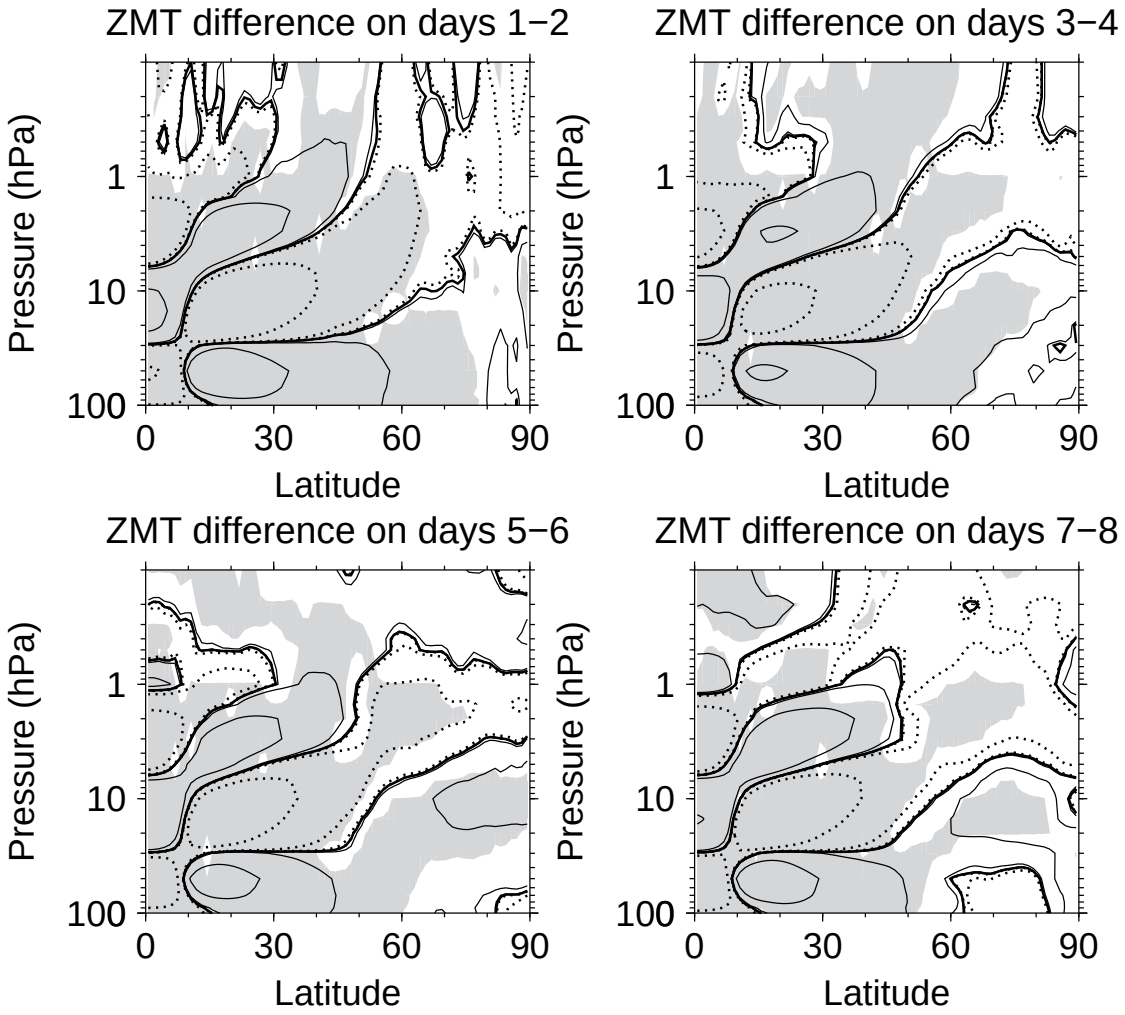

Figure 8.6: As for figure 8.5 but for ZMT. Contours are at $\pm 10^{-4}$, $\pm 10^{-3}$, ± 0.01, ± 0.1 and ± 1 K.

ZMZW throughout the stratosphere north of 60N, though these ZMZW differences are not highly statistically significant.

Statistical significances of the mean differences between the model branch runs nudged to QBO-E and the ClimEq run were calculated using an MC bootstrap technique, in the same way as described in section 6.2. Statistical significances of the differences for other variables were tested in the same way.

Figure 8.6 shows the mean ZMT differences. Again the influences of the QBO and its meridional circulation are initially apparent up to ∼70N. At high latitudes there is

statistically significant warming of the polar region at ∼10 hPa with cooling above on days 5–8. After this time the cooling above ∼2 hPa extends to 20N and is statistically significant above the 95% level for the rest of the duration of the branch runs (not shown). The initial ZMT differences are likely to be due to adiabatic processes, since there is no reason to think the QBO directly affects radiative processes and the model does not include interactive chemistry.

The ZMZW and ZMT differences at high latitudes are qualitatively consistent with the thermal wind relationship, in that warming at high latitudes, with an associated decrease in the meridional ZMT gradient at ∼10 hPa, is concurrent with ZMZW becoming more easterly above this level.

Time series of the ZMZW and ZMT differences show much variation between individual ensemble members, with variation in the sign and magnitudes of the differences. The envelope of the differences grows exponentially in time (not shown) in a similar way to that shown in response to an applied torque in figure 6.1. This may be because nudging towards a QBO-E profile causes changes in small scale structures of the flow which then tend to grow exponentially with time[2]. This is different to the response to an applied torque, which was consistent between ensemble members in the middle stratosphere for about two weeks (section 6.2), and may be due to additional noise resulting from perturbations to the troposphere, and it could also reflect the greater complexity of the way the QBO affects the vortex. This means differences for individual branch runs outside the tropical stratosphere cannot meaningfully be attributed to nudging towards QBO-E. However, taking the mean of the differences over all the branches can reveal meaningful systematic changes.

[2]Using equations 4.2 and 4.3 with index notation and summation convention,

$$\begin{aligned}
\delta\dot{x}_a &= f_a(\mathbf{x}'(t),t) + \mathcal{L}_a(\mathbf{x}'(t),t) - \mathcal{L}_a(\mathbf{x}(t),t) \\
&= f_a(\mathbf{x}'(t),t) + \left.\frac{\partial \mathcal{L}_a}{\partial x_b}\right|_{(\mathbf{x},\,t)=(\mathbf{x}_0,\,0)} \delta\dot{x}_b(0)t + \mathcal{O}(t^2) \\
&= f_a(\mathbf{x}'(t),t) + \left.\frac{\partial \mathcal{L}_a}{\partial x_b}\right|_{(\mathbf{x},\,t)=(\mathbf{x}_0,\,0)} \delta x_b(t) + \mathcal{O}(t^2)
\end{aligned}$$

using $\delta x_b(t) = \delta\dot{x}_b(0)\,t + \mathcal{O}(t^2)$ to get the last line. The first feedback term includes the effect of exponential perturbation growth.

The signal in figures 8.5 and 8.6 loses statistical significance after day 8. The ZMZW and EP flux differences between the branch and ClimEq runs grow approximately linearly with time over the first week, and so the loss of statistical significance is consistent with this signal continuing to grow, but the noise growing more quickly, so that the signal to noise ratio decreases with time. The magnitude of the noise must saturate after some time at the level of the climatological variability. The fact that there is an HT relationship in the 240-year pre-industrial control run implies that the signal to noise ratio would be large enough again after several months for the signal to be statistically significant, implying that the signal would continue to increase in time. Seeing the full evolution of the vortex response to QBO-E forcing in this model would likely require a prohibitive amount of computing resources, and similar experiments with computationally cheaper models may be helpful to understand the transient response fully. The transient response to nudging towards QBO-E presented by Garfinkel *et al.* [2012b] appears to be more statistically significant, but the model they used has unrealistically low vortex variability, with a very low frequency of major SSWs [Richter *et al.*, 2008], and this may be part of the reason for the difference.

Since the envelope of the differences in individual ensemble members grows exponentially in time, it is possible that some ensemble members develop particularly large differences. To test robustness with respect to excluding outliers, a trimmed mean [Wilks, 2006] of the ZMZW and ZMT differences was calculated, where the top and bottom 20% of values were excluded. The results were very similar to the mean of all the data (not shown).

Figure 8.7 shows the mean difference in the residual circulation $(\overline{v}^*, \overline{w}^*)$ between the branch runs and the ClimEq run and the difference in the associated acceleration term in equation 2.13, given by

$$-\overline{v}^* \left[(a\cos\phi)^{-1} (\overline{u}\cos\phi)_\phi - f \right] - \overline{w}^* \overline{u}_z.$$

The cells of the QBO meridional circulation develop over time with a more poleward subtropical residual circulation in the branch runs near 30 hPa and a more equatorward circulation near 7 hPa. On days 1–2 the associated acceleration closely resembles

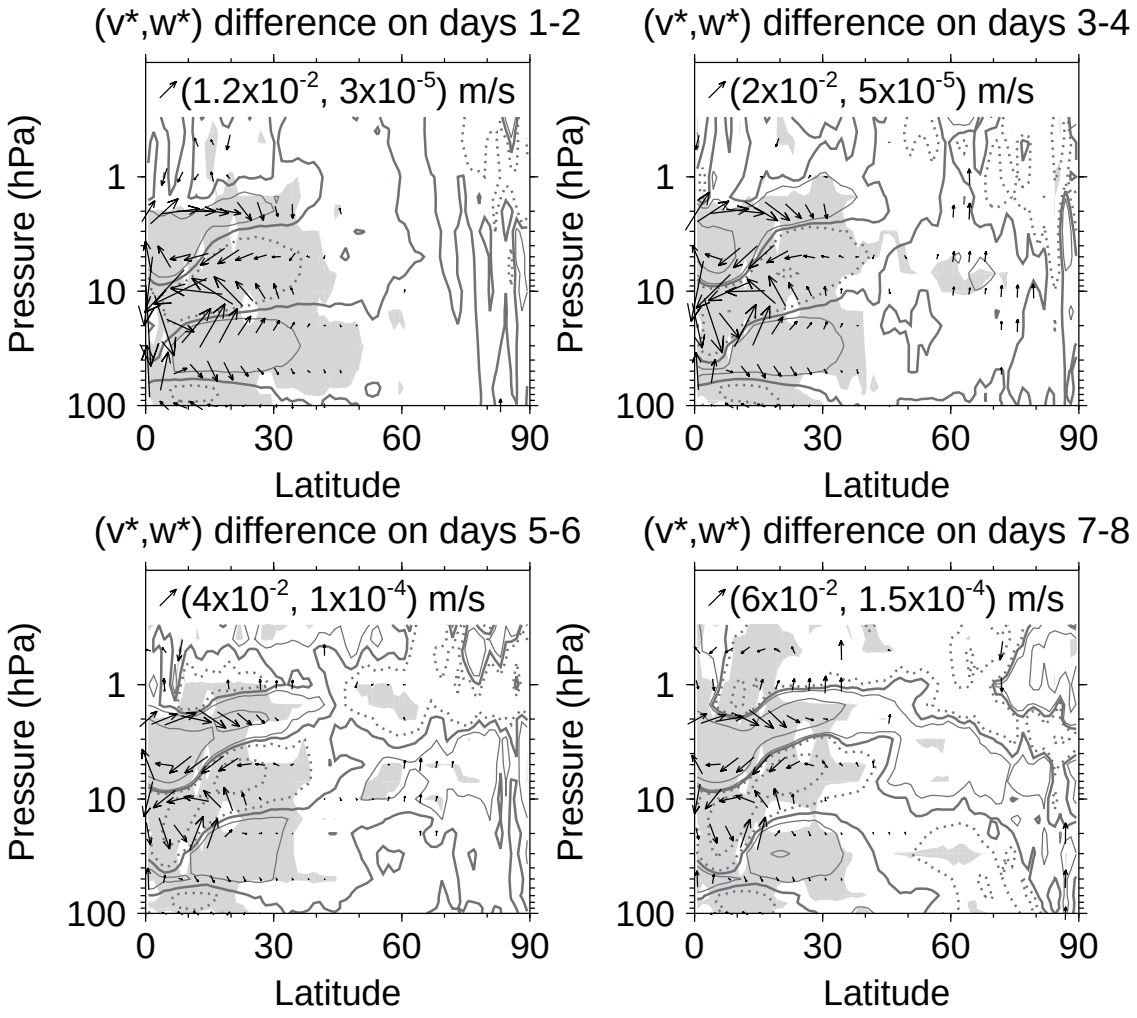

Figure 8.7: Ensemble mean residual circulation differences between branch runs nudged towards the QBO-E equatorial wind profile and the ClimEq run averaged over each two day interval up to 8 days following branching (arrows, shown only where either the $\overline{v}^*$ or $\overline{w}^*$ differences are statistically significant above the 95% level) and the associated zonal acceleration differences (contours, at 0, ± 0.02 and ± 0.1 m/s/day, with negative contours dotted and the zero contour thickened). The residual circulation vectors are shown at pressures 0.5, 1, 2, 5, 10, 20, 50 and 100 hPa and every 3.75° in latitude. A reference arrow is shown in the top left of each plot along with its (v^*, w^*) values. Shading shows where differences in the acceleration term are statistically significant above the 95% level.

the ZMZW differences in figure 8.5 between ∼20–70N in the lower stratosphere and between ∼20–60N in the upper stratosphere, so these differences appear to be caused by the acceleration of the meridional circulation. At later times the acceleration due directly to the QBO meridional circulation is confined equatorward of ∼40N.

Figure 8.8 shows the mean difference in the EP flux and D_F between the branch runs and the ClimEq run. On days 1–2 the EP flux is more equatorward in the lower stratosphere and poleward in the middle and upper stratosphere between ∼15–30N, and the D_F differences are relatively small. Between days 3–8, however, EP flux differences become statistically significant in middle latitudes, and statistically significant poleward of 60N in the middle stratosphere in days 5–6. The EP flux is more poleward in the lower stratosphere and more equatorward in the upper stratosphere. This is associated with increased EP flux convergence in the high-latitude stratosphere during days 5–8 between ∼2–20 hPa north of ∼30N, and therefore this contributes to the weakening of the winds between ∼2–20 hPa and ∼60–90N over days 5–8 (figure 8.5). The EP flux differences are very similar if only zonal wavenumbers 1–3 are considered (not shown), so it is associated with planetary-scale waves. The residual circulation differences are more upward in the mid-stratosphere at this time and partially cancel this contribution, and contribute to weakening of the winds at high latitudes above ∼2 hPa on days 5–8 and below ∼20 hPa on days 7–8 (figure 8.7).

The more poleward EP flux from the tropics to the high latitudes is the signal that Naoe and Shibata [2010] argued should be present if the HT mechanism is correct. These results are thus consistent with the HT mechanism and indicate that increased reflection of eddy zonal momentum flux in QBO-E may directly affect the polar stratosphere. Therefore the HT mechanism cannot be ruled out of playing a part in the HT relationship as argued in recent studies [Naoe and Shibata, 2010; Yamashita *et al.*, 2011; Garfinkel *et al.*, 2012b]. Note that this signal is unlike the QBO-E minus QBO-W composite differences in the EP flux in any calendar month (figure 4.1), illustrating the necessity of examining the transient response to understand the influence of the QBO.

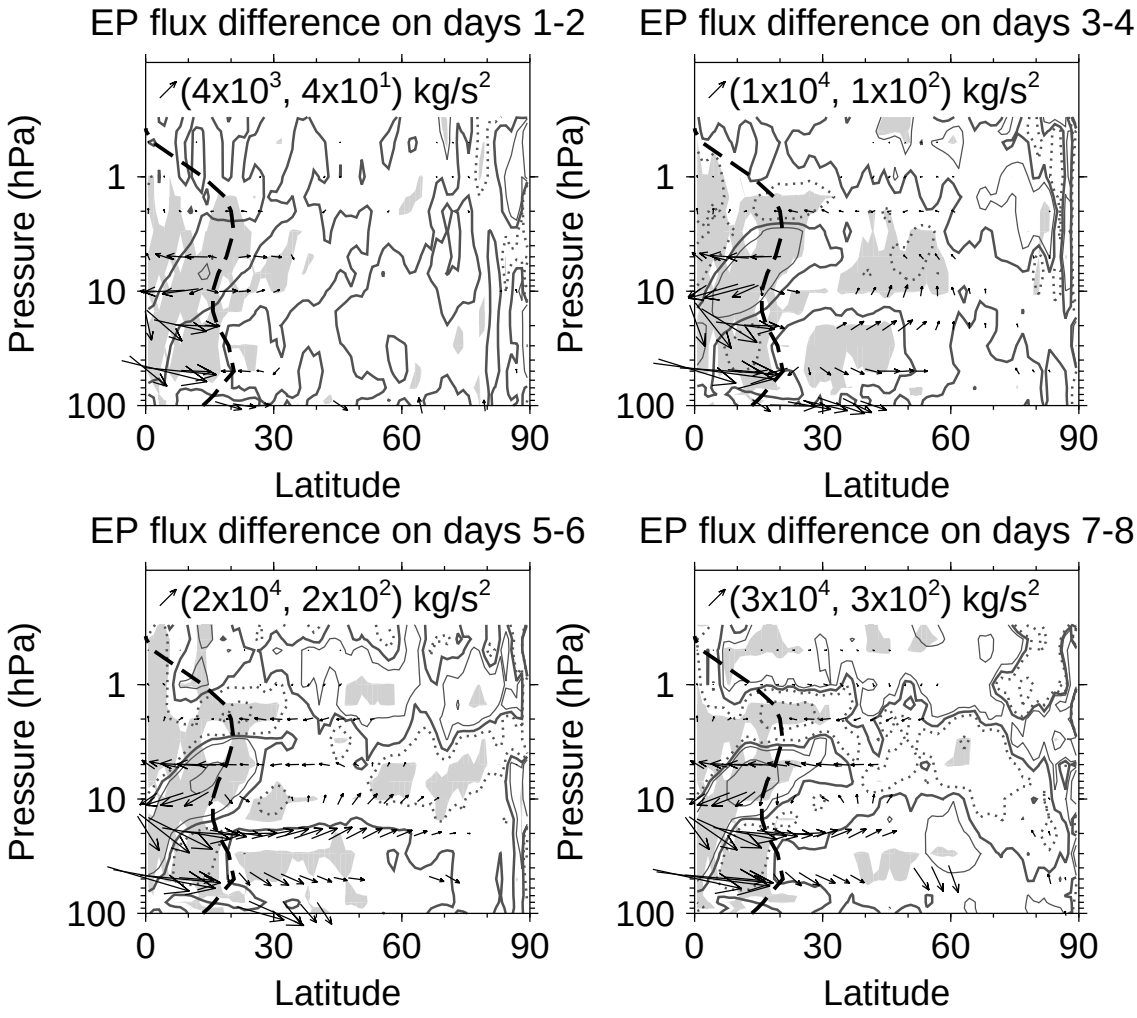

Figure 8.8: Ensemble mean EP flux differences between branch runs nudged towards the QBO-E equatorial wind profile and the ClimEq run averaged over each two day interval up to 8 days following branching (arrows, shown only where either the F^{ϕ} or F^{z} differences are statistically significant above the 95% level) and D_F differences (contours, at 0, ± 0.02 and ± 0.1 m/s/day, with negative contours dotted and the zero contour thickened). The thick dashed line shows the zero wind line in the days 1–8 mean ZMZW in the ClimEq run. The EP flux differences are shown at pressures 0.5, 1, 2, 5, 10, 20, 50 and 100 hPa and every 3.75° in latitude. A reference arrow is shown in the top left of each plot along with its (F^{ϕ}, F^{z}) values. Shading shows where D_F differences are statistically significant above the 95% level.

8.4.1 Role of the tropical middle and upper stratosphere

However, it is not just the tropical lower stratospheric easterly wind acceleration that matters – the westerly acceleration above 10 hPa appears to cause a more equatorward EP flux that decreases its convergence at high latitudes, reducing the QBO influence there.

The changes in the tropical upper stratosphere may also matter. Figure 8.9 shows the EP flux differences using the "acceleration scaling" defined by Gray *et al.* [2003], where the flux components are divided by $\rho\cos(\phi)$, to indicate the size of the acceleration associated with the flux. By days 7–8 there is more poleward flux in the QBO-E runs at 1 hPa and above which appears associated with the tropical ZMZW at these heights becoming more easterly (figure 8.5). Easterly anomalies at these altitudes are also observed in QBO-E in winter [Pascoe *et al.*, 2005]. Although the magnitude of this momentum flux is small, it is associated with EP flux convergence and deceleration of the ZMZW in the mid-latitude upper stratosphere. If the vortex state is sensitive to ZMZW changes in this region then this provides a way for tropical upper stratospheric wind changes to contribute to the HT relationship, as suggested by Gray *et al.* [2001a,b].

8.4.2 Dependence on initial conditions

A further advantage of examining the short-term transient response to a forcing, rather than the response on longer time scales, is that information is available not just about the mean direct effect of the forcing, but also its dependence on the atmospheric state. This allows the dependence on the atmospheric state of the key poleward response of F^ϕ to nudging towards QBO-E that is shown in figure 8.8 to be investigated, which is informative about the conditions under which this effect is strongest.

First, the integral of the F^ϕ response to nudging towards QBO-E (δF^ϕ) through the surface at 45N between 10–100 hPa is defined as

$$\delta F^\phi_{45\mathrm{N}} = \int_{z_1}^{z_2} 2\pi a\, \delta F^\phi \cos(\phi)|_{\phi=45\mathrm{N}}\, \mathrm{d}z,$$

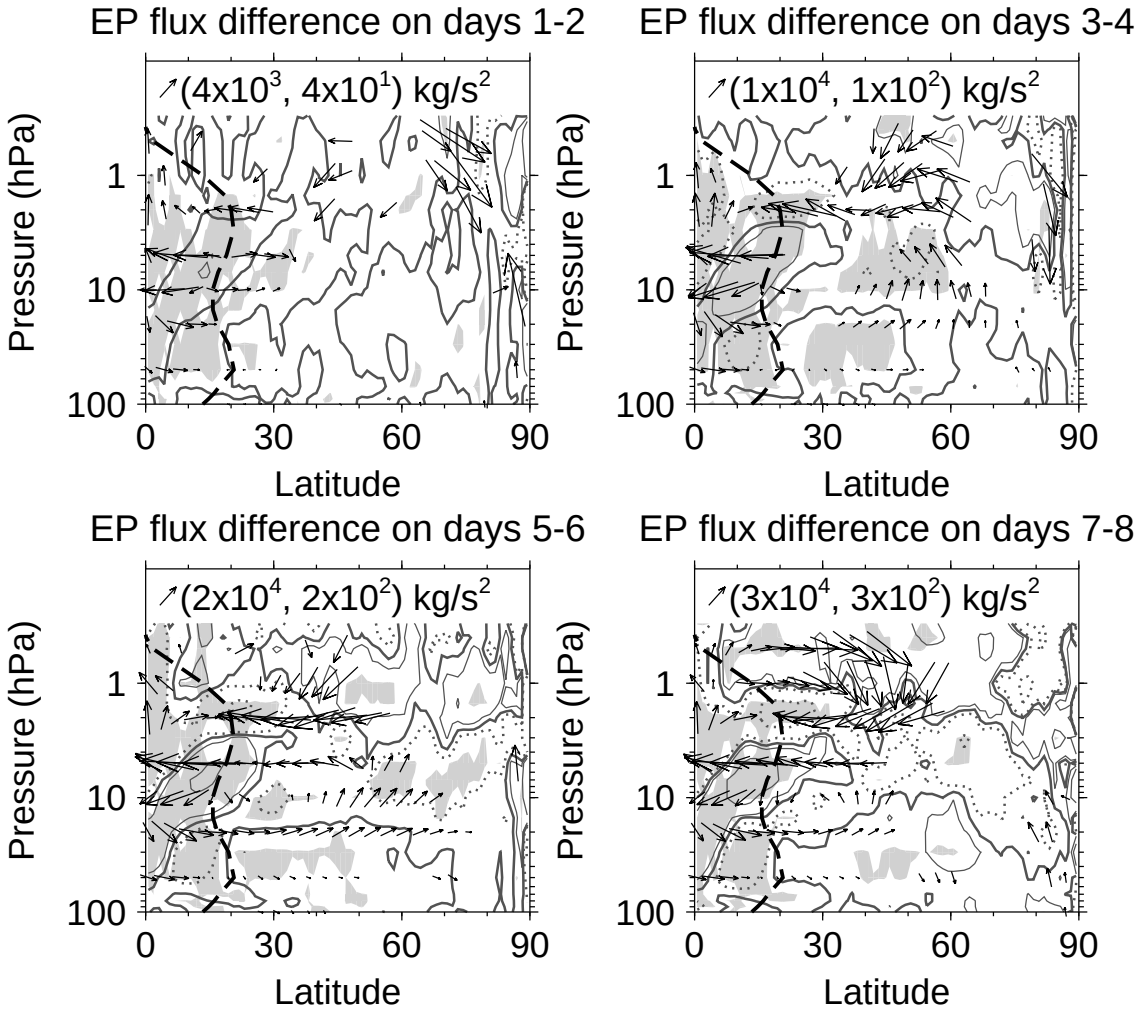

Figure 8.9: As in figure 8.8 but using "acceleration scaling" of the EP flux vectors as defined by Gray *et al.* [2003], which indicates the zonal acceleration associated with the flux. A reference arrow is shown in the top left of each plot along with the (F^{ϕ}, F^{z}) values it would have at the Equator at 10 hPa.

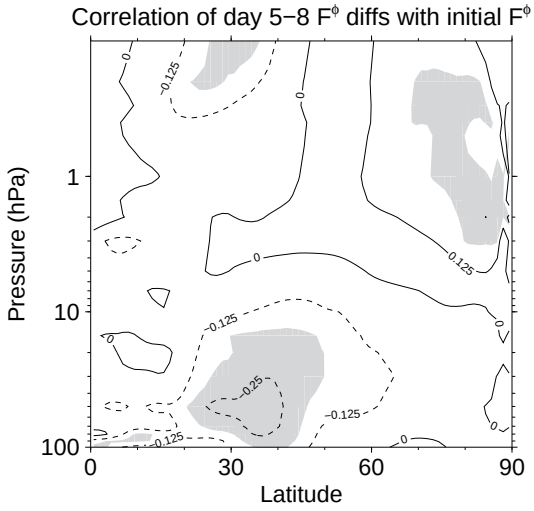

Figure 8.10: The correlation of the day 5–8 mean of $\delta F^{\phi}_{45\mathrm{N}}$ with the day 1 mean of F^{ϕ} in the ClimEq run. Shading indicates where correlations are statistically significant above the 95% level.

where $z = -H\log(p/p_0)$ is log-pressure height, with H=7 km and p_0=1000 hPa, $z_1 = z(p\text{=}100\,\mathrm{hPa})$ and $z_2 = z(p\text{=}10\,\mathrm{hPa})$ and a is the Earth's radius. This definition follows from applying Stokes' theorem to equations 2.5 and 2.6 of Dunkerton *et al.* [1981] to infer the contribution of $\delta F^{\phi}_{45\mathrm{N}}$ to the $\widetilde{D_F}$ response within a closed surface of which the surface at 45N between 10–100 hPa forms a part (where $\widetilde{D_F} = 2\pi a^2\cos(\phi)\nabla\cdot\mathbf{F}$ as in section 6.4).

Figure 8.10 shows the correlation across the branch runs between $\delta F^{\phi}_{45\mathrm{N}}$, averaged over days 5–8, with F^{ϕ} at all latitudes and pressures in the NH stratosphere in the ClimEq run, averaged over the period corresponding to day 1 of each branch run. There is a negative correlation of up to $\sim$−0.3 between $\delta F^{\phi}_{45\mathrm{N}}$ and the initial F^{ϕ} in the mid-latitude lower stratosphere. This is statistically significant above the 95% level between $\sim$25–45N and $\sim$20–100 hPa, according to an MC permutation test used to estimate the probability of obtaining a correlation this large under the null hypothesis that these variables are unrelated, by resampling without replacement the $\delta F^{\phi}_{45\mathrm{N}}$ data 1000 times and finding how often the resulting correlation has a magnitude exceeding that for the actual $\delta F^{\phi}_{45\mathrm{N}}$ data.

Therefore the day 5–8 $\delta F^{\phi}_{45\mathrm{N}}$ is stronger when F^{ϕ} is initially more negative, in other words when the EP flux is more equatorward, as expected if the response is due to increased reflection of this flux. So this provides more supporting evidence that the EP flux response is due to the HT mechanism. There is also a statistically significant positive correlation between $\delta F^{\phi}_{45\mathrm{N}}$ and the initial F^{ϕ} at high latitudes near 1 hPa – this could possibly be due to greater upward wave propagation from the troposphere into the polar upper stratosphere being associated with less equatorward propagation in the lower stratosphere.

8.5 Discussion

These results do not indicate that the EP flux convergence in the high-latitude stratosphere arises due to alternatives to the HT mechanism discussed in section 2.7.2 – the mid-latitude EP flux differences have the opposite sign to those predicted by the suggested mechanisms of Anstey *et al.* [2010], Naoe and Shibata [2010] (figure 2.15) and Yamashita *et al.* [2011] (figure 2.16) (note that these figures show the predicted QBO-W minus QBO-E differences, so they should be reversed in sign for comparison with figure 8.8), and its convergence in mid-latitudes has the opposite sign to that predicted by Garfinkel *et al.* [2012b] (figure 2.17). It is not clear if there are reasons why the signatures of these mechanisms would take more than 8 days to appear, but these results make it seem less likely that any of these mechanisms are dominant. The sub-tropical ZMZW anomalies associated with the QBO meridional circulation may take longer than 8 days to fully develop and affect the EP flux. Due to the extratropical influence on this region it is difficult to deduce the part of these wind anomalies in QBO-E minus QBO-W composite differences that are due directly to the QBO, which would be useful to compare to the sub-tropical ZMZW responses to nudging towards QBO-E shown in figure 8.5.

The greater reflection of EP flux in the lower stratosphere in the QBO-E runs occurs despite the zero wind line (ZWL) shifting both poleward and equatorward on different levels below 10 hPa (not shown). It may therefore be interesting to

examine the contributions to the EP flux differences from waves with non-zero phase velocities, though this is not done here. As remarked in section 8.2, the QBO profile in these simulations is too narrow meridionally, and the ZWL shift may have been more pronounced if this were not the case. It is not clear that the ZWL should be considered fundamental for the HT relationship.

It should also be considered whether the use of nudging in the tropical stratosphere would cause differences between the coupling between the tropics and extratropics in these model runs and in the real stratosphere. Previous studies have used nudging of equatorial winds to examine the HT relationship and have obtained realistic results [e.g. Hamilton, 1998; Garfinkel *et al.*, 2012b], implying that the nudging does not drastically interfere with the coupling between the tropics and extratropics. The nudging would tend to dampen wave activity in the tropical region with time scales longer than a couple of weeks. The most unrealistic effect of the nudging may be to create a QBO profile that is too narrow meridionally compared with that in ERA-Interim, which may weaken the extratropical response [O'Sullivan and Young, 1992]. I see no reason why these effects would qualitatively change the interaction between the tropics and extratropics.

The atmospheric response to an applied forcing at short times following application of the forcing is indicative of the whole forcing mechanism only if the steps of the mechanism unfold on a time scale less than about the system's dynamical time scale, so mechanisms whose early stages are consistent with the results presented above but which also involve subsequent steps cannot be ruled out. These results also do not rule out the meridional circulation having a role through modifying the background state through which the eddy zonal momentum propagates. The experiments also show the effect of nudging towards QBO-E from a state having close to climatological equatorial winds – as the winds approach a full-strength QBO-E state, the vortex response may depend non-linearly on further increases in the strength of the QBO-E profile. This could be explored in a similar way to the method used here, using a control run nudged towards a QBO-E state with branch runs nudged towards a stronger QBO-E state. In the model runs considered here, the equatorial ZMZW differences

do reach the approximate magnitude of the anomalies in the QBO-E phase in observations in the first 8 days, however, allowing the possibility that non-linearity will show an effect, so the effects of non-linearity with respect to the equatorial winds are not expected to greatly change these results.

The ZMZW differences on days 7–8 in the high-latitude upper stratosphere are $\sim$0.05 m/s. This is several times too small for the differences to reach the magnitude of the QBO-E minus QBO-W differences in the pre-industrial control run, of $\sim$5 m/s (figure 8.1(b)), within a few months if the deceleration is constant. This may be due to the nudged QBO being too narrow, or it may indicate that there would be non-linear amplification of the initial vortex response, or that another mechanism would contribute to the deceleration after some time. Therefore these results leave open the question of whether the HT mechanism can account quantitatively for the HT relationship.

The key innovative part of this work has been to examine the stratospheric response to QBO forcing on time scales that are shorter than or of the order of the mid-stratosphere's dynamical time scale of $\sim$1 week, so the response to QBO-E forcing alone before the circulation has evolved is seen. If the EP flux differences are averaged over the first 16 days, as done by Garfinkel *et al.* [2012b], then the poleward EP flux difference from the tropics to high latitudes is no longer apparent. (However, the model used by Garfinkel *et al.* [2012b] does not show a more poleward extratropical EP flux difference in the first 8 days as found in the experiments presented here (C. Garfinkel, *pers. comm.*), so there appears to be model-dependence of this result, perhaps related to their model having a lower spatial resolution.)

8.6 Conclusions

Examination of the first few days of the transient response of the vortex to nudging of equatorial stratospheric winds towards a QBO-E state in the HadGEM2-CCS GCM indicates that the EP flux becomes less equatorwards between the tropics and high latitudes in the lower stratosphere, that there is greater convergence of the EP flux in

the high-latitude stratosphere between about 2–20 hPa, and that the westerly wind in this region decelerates. This is consistent with the hypothesis that more easterly winds in the tropical lower stratosphere cause greater reflection of EP flux towards the polar stratosphere, directly causing deceleration of the westerly winds, as suggested by Holton and Tan [1980]. These results do not show the signatures of the mechanisms suggested by Anstey *et al.* [2010], Naoe and Shibata [2010], Yamashita *et al.* [2011] or Garfinkel *et al.* [2012b], who proposed that the role of the lower stratospheric zero wind line is less important than that of the QBO meridional circulation or the mid-stratospheric zero wind line. Combining these findings with the results of O'Sullivan and Young [1992], Gray *et al.* [2003, 2004] and Naito *et al.* [2003], which indicate that a strong meridional circulation and the shift of the zero wind line in the mid-stratosphere (in the opposite sense to that in the lower stratosphere) are not essential to produce the HT relationship, it is concluded that the mechanism of Holton and Tan [1980] is likely to contribute to the HT relationship. However, the total anomalous poleward EP flux depends on ZMZW changes throughout the depth of the tropical stratosphere, and it is not clear if the role of the zero wind line is fundamental. Contributions have not been ruled out from mechanisms that include multiple steps that unfold on a time scale greater than about a week, which may include those for which the ZMZW anomalies associated with the meridional circulation are important if they take more than 8 days to reach their full strength. The role of feedbacks in setting the magnitude and structure of the response to QBO forcing on long time scales has not been examined, and it was shown in chapters 5 and 6 that these may be important.

Chapter 9

Conclusions and discussion

The principal new findings presented in this thesis are as follows. In chapter 4 the HT relationship on monthly time scales was examined in ERA-40, with a particular emphasis on the response of the zonally asymmetric part of the flow to changes in the QBO phase. It was shown that the response is very similar to the NAM structure, which had been shown previously to be the case for the zonal mean part of the response but not the wave part. It was suggested that this may be a manifestation of a phenomenon seen in other dynamical systems that show the response to arbitrary forcings to be very similar to the leading mode of variability. This prompted a re-evaluation of previous analyses that discussed whether this response is consistent with the various proposed mechanisms for the HT relationship, and it was argued that this response does not put much of a constraint on the mechanism. It was suggested that a useful way of understanding the response of a system to an applied forcing is as a sum of a "direct response", the time-integral of the applied forcing term, and "feedbacks" associated with the state-dependence of the equations of motion. It was shown that, in general, the direct response would be expected to dominate the total response only for a period of time after the forcing is applied that is shorter than about the dynamical time scale. It was argued that therefore the short-term transient response of the vortex to QBO forcing could be much more informative about the mechanism behind the HT relationship than analysis of the long-term response.

In chapter 5, the long-term mean response of the SMM to a set of imposed torques was examined under two types of simulations: those with perpetual January boundary

conditions and those with boundary conditions that followed a more realistic seasonal cycle. In these experiments, tropospheric wave driving was held fixed, so the role of internal stratospheric dynamical processes could be seen clearly. It was found that the response to many different torques placed in the high latitude stratosphere under perpetual January conditions was like the model's NAM in many respects, in line with the suggestion in chapter 4 that feedbacks act to make the stratospheric response to an arbitrary forcing resemble the NAM. Counter-intuitively, placing a positive torque in the high-latitude stratosphere leads to deceleration of the ZMZW there after a couple of weeks. This clearly shows that the feedback of the EP flux associated with resolved planetary waves onto changes in the zonal mean flow greatly affects the long-term mean response, with increased EP flux convergence in the region of the applied torque tending to overcome the torque's direct effect. However, the wintertime composite responses in more realistic simulations with a seasonal cycle were not very NAM-like, indicating that certain conditions need to be satisfied to produce a NAM-like response.

Examination in chapter 6 of the transient response to the torques applied at high latitudes indicated that the feedback of the EP flux was qualitatively similar for both types of boundary conditions. The EP flux initially becomes more upward and convergent in the high-latitude lower and middle stratosphere so that it cancels the direct effect of the torque. This first stage of the feedback seems consistent with the physical explanation that a positive zonal torque acts to increase the meridional PV gradient, and this leads to greater propagation of planetary waves into the region where the torque peaks, where they drive easterly acceleration. This feedback is weaker in experiments with a seasonal cycle, due to the background westerly ZMZW being stronger. Later the EP flux becomes more upward and convergent in the upper stratosphere, causing deceleration of the ZMZW there. This is consistent with the physical explanation that the additional upward-propagating Rossby waves in the middle stratosphere produced in the first stage propagate into the upper stratosphere and transfer easterly momentum there. Correspondingly this momentum flux is also weaker in experiments with a seasonal cycle. Finally EP flux convergence in the polar

stratosphere increases and becomes temporarily larger than the torque, which causes the ZMZW response to become negative here and NAM-like overall in the perpetual January experiments, but not in the seasonal cycle experiments. This indicates that feedbacks associated with large-scale dynamics tend to make the response to an arbitrary forcing AM-like in the stratosphere, but the total response will be AM-like only if the feedback part dominates the response. Simulating the relative strengths of this feedback and the direct response accurately is therefore important in models for producing realistic responses to applied forcings.

The SMM simulates large-scale stratospheric dynamics, but it does not represent tropospheric dynamics, and it uses a very simple Rayleigh friction parameterisation of gravity waves, so feedbacks associated with these processes are not represented. Therefore the simulated responses to applied torques are not expected to be quantitatively accurate, but indicate qualitatively how the large-scale stratospheric dynamics respond.

As well as supporting the hypothesis that feedbacks act in the stratosphere to make responses to arbitrary forcings appear like the AMs, these results help understanding of how feedbacks from large-scale dynamics may act on the direct perturbation to the high-latitude stratosphere caused by the QBO and other external forcings that act on the vortex. However, more investigation of the strength and sign of the feedback is necessary to make firm conclusions about how feedbacks act in the real atmosphere.

In chapter 7, the SMM simulations were used to test several methods proposed in previous studies to gain information about forcing mechanisms using observational data or model runs with realistic forcings. It was found that the only method to robustly give useful information about the known applied torque was to examine the short-term transient response. This method showed potential to yield a lot of information about an applied forcing. (However, it would be expected to be less successful when applied to understand a real forcing mechanism which is likely to be more complicated than an applied torque.) It is correspondingly hard to make inferences using other methods. These findings suggest that experiments designed to investigate, for example, the short-term transient response of the troposphere to

stratospheric perturbations, or the response of the vortex to equatorial heating perturbations associated with the eleven-year solar cycle, may yield new insights into the physics behind these teleconnections.

Finally, in chapter 8 the short-term transient response of the vortex to QBO-E forcing was examined in the HadGEM2-CCS GCM. The most prominent feature of this response is the EP flux becoming less equatorward between the tropics and high-latitudes in the stratosphere between ∼10–100 hPa, with an associated increase in convergence of the flux in the high-latitude stratosphere and deceleration of the ZMZW. This is consistent with the mechanism proposed by Holton and Tan [1980], indicating that this mechanism has a role in the HT relationship. The response did not show the signature of the mechanisms proposed by Anstey *et al.* [2010], Naoe and Shibata [2010], Yamashita *et al.* [2011] or Garfinkel *et al.* [2012b], making it seem less likely that any of these mechanisms plays a dominant role. However, it is possible that the subtropical ZMZW response associated with the QBO meridional circulation may not have reached full strength in the short time considered, so this cannot be ruled out of contributing to the HT relationship. Due to variability in the simulations, it was not possible to study the full transient response to QBO-E forcing in these experiments, and given how important feedbacks were in determining the SMM response to an applied torque discussed above, it seems necessary to do further studies of the transient response to QBO-E forcing in order to fully understand the mechanism behind the HT relationship.

This represents substantial progress on the first issue surrounding the HT relationship identified in chapter 1, namely understanding what physical processes give rise to the relationship. This also gives some insight into the other two issues (understanding which features of the QBO are necessary and which are sufficient to produce the HT relationship, and learning how to improve numerical atmospheric models so that they produce the relationship), by indicating that an easterly ZMZW anomaly in the tropical lower stratosphere alone is sufficient to affect the vortex in a GCM, and that numerical models will probably require a reasonable representation of the QBO, and of the meridional component of EP flux in the lower stratosphere,

to simulate the HT relationship properly. However, the weakness of the HT relationship in HadGEM2-CCS relative to that in ERA-40 (chapter 8) indicates that a good representation of these features is still not sufficient to reproduce the relationship quantitatively, and more work is needed to understand how to improve numerical models so that they simulate the HT relationship well.

9.1 Extending this work

Several lines of experimental investigation could be pursued to extend the progress that has been made by this work. Firstly, it would be desirable to study the full transient response of the vortex to QBO forcing using experiments similar to those discussed in chapter 8. In order to achieve a good signal-to-noise ratio, it would probably be necessary to perform a much larger number of branch runs with QBO-E and QBO-W forcings, and therefore a model that is computationally far cheaper than HadGEM2-CCS would be required. The SMM may be a good candidate for this. The SMM has the additional advantage that there is no source of noise due to the troposphere responding in a chaotic way to perturbations made in the stratosphere, making a high signal-to-noise ratio more achievable. Another candidate model is the Reading Intermediate General Circulation Model (IGCM), which was found to simulate the stratosphere reasonably well by Bell *et al.* [2009]. Experimental results using this model would include feedbacks due to changes in the tropospheric state, making the results more realistic, but probably noisier, than those in the SMM. It may also be interesting to study the transient response in a framework other than that of the TEM equations and to consider the response in 3D, as this may yield more physical insights.

In order to understand which properties of the QBO are necessary and which are sufficient to give rise to the HT relationship, and which properties are the most important, it may be fruitful to perform experiments where the equatorial ZMZW is nudged towards various idealised profiles. These could be designed to systematically test the sensitivity of the vortex response to changing the lower stratospheric

and middle stratospheric equatorial ZMZW and to changing the vertical wind shear between them. One cannot be varied independently of the others in one experiment, but by intelligently combining the results from experiments using selectively chosen profiles, it is likely that conclusions about the importance of different QBO properties could be drawn.

Consider the profiles illustrated in figure 9.1. Profile A is a simple representation of the zonal wind in QBO-E, which would be anticipated to produce a similar vortex response to the QBO-E profile derived from ERA-Interim. Profile B has easterlies throughout the equatorial stratosphere. Profile C has zero lower stratospheric zonal wind, the same mid-stratospheric vertical wind shear as profile A and double the mid-stratospheric westerlies. Profile D has double the lower stratospheric easterlies in profile A, the same strength wind shear and zero mid-stratospheric westerlies.

Assuming that it is the case that the vertical wind shear at the base and top of the QBO profile is unimportant, then if only the equatorial lower stratospheric wind is important, profiles B and D will give rise to vortex weakening but not profile C. If only the mid-stratospheric wind is important, then only profile C will lead to vortex weakening. If only the mid-stratospheric vertical wind shear and the associated meridional circulation is important, then profiles C and D will produce vortex weakening but not profile B. Thus the strength of the vortex weakening observed in response to imposing each profile can be used to assess the importance of each of these factors. Additional wind profiles in which the magnitudes of these properties are varied could be used to quantify the relative importance of each if more than one property is found to matter. It could also be seen to what extent the different properties of the QBO structure combine linearly to influence the vortex.

The effect of increasing the depth of the lower stratospheric equatorial easterlies or westerlies could also be examined using similar experiments. These would test the idea that the equatorial wind influence is especially strong when there is a deep layer of easterlies or westerlies at the Equator [McIntyre, 1982; Dunkerton *et al.*, 1988].

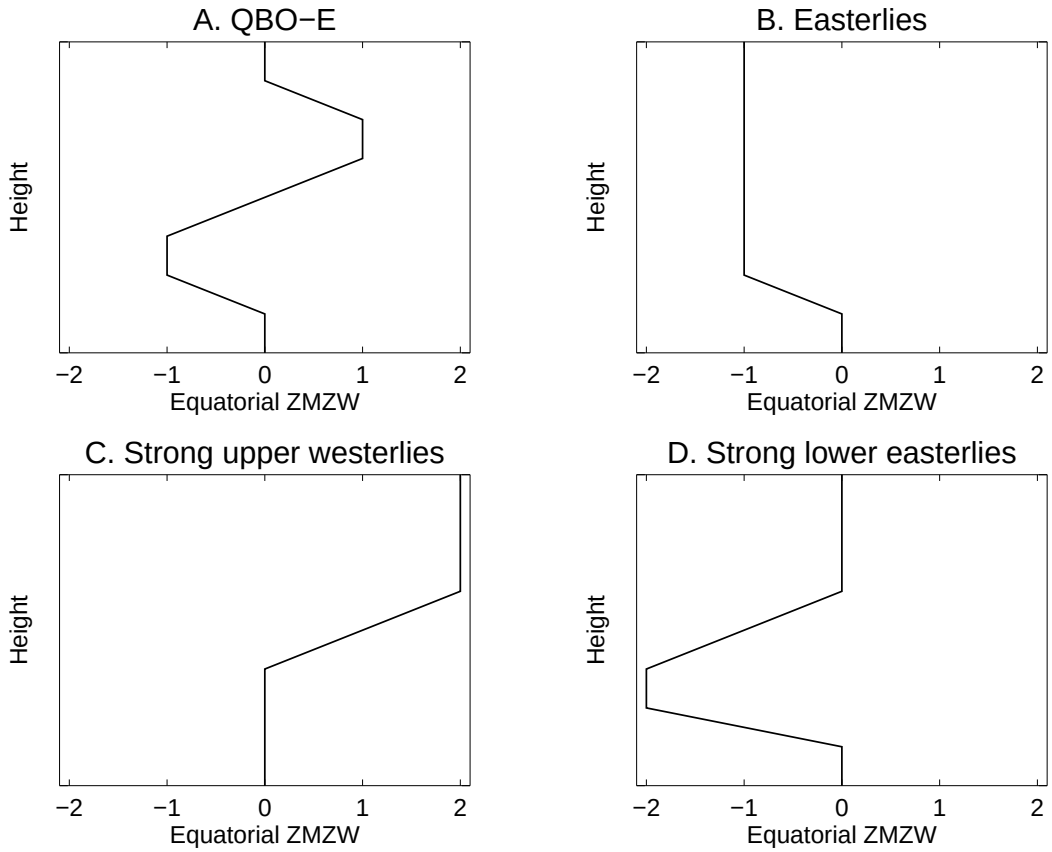

Figure 9.1: Equatorial stratospheric ZMZW profiles that could be used to understand what aspects of the QBO structure are important for the HT relationship. Axis units are arbitrary. Profile A is a simple representation of the QBO-E phase. B has easterlies throughout the equatorial stratosphere, with no mid-stratospheric vertical wind shear. C has mid-stratospheric vertical wind shear and strong upper and middle stratospheric westerlies. D has strong lower stratospheric easterlies and mid-stratospheric vertical wind shear.

Analysing the short-term transient response of the extratropical stratosphere to these idealised profiles may also shed light on the mechanism by which each profile brings about changes in the vortex.

Again, due to the large number of simulations that would need to be performed, it would probably be best to use a computationally cheap model such as the SMM or IGCM for most experiments, and then follow up selected results with a more comprehensive GCM such as HadGEM2-CCS to verify that the results are robust.

A third set of experiments that would be interesting to carry out would be to repeat the experiments of chapters 5 and 6 to investigate the response to a stra-

tospheric torque in a model with an interactive troposphere, such as the IGCM or HadGEM2-CCS. Comparing the results to those in chapters 5 and 6, or to similar experiments in the same model in which the troposphere is nudged so that it does not respond to stratospheric changes, would indicate whether feedbacks from changes in the tropospheric state are important for determining the stratospheric response to forcings on time scales longer than a few weeks. It would also be useful to see if the response to a torque is more NAM-like in model runs with a realistic seasonal cycle than they were in the SMM. This would show whether the NAM-like vortex response to QBO forcing on monthly time scales really is due to the response to arbitrary forcings being NAM-like. It would also be interesting to perform similar experiments with different simple forcings, such as heating perturbations or zonally asymmetric forcings, or use a model with a more realistic representation of gravity waves.

Bibliography

Ambaum MHP. 2010. Significance tests in climate science. *J. Climate* **23**(22): 5927–5932.

Ambaum MHP, Hoskins BJ. 2002. The NAO troposphere-stratosphere connection. *J. Climate* **15**(14): 1969–1978.

Andrews DG. 1987. On the interpretation of Eliassen-Palm flux divergence. *Q. J. R. Meteorol. Soc.* **113**(475): 323–338.

Andrews DG, Holton JR, Leovy CB. 1987. *Middle atmosphere dynamics*. Academic Press.

Andrews DG, Mahlman JD, Sinclair RW. 1983. Eliassen-Palm diagnostics of wave-mean flow interaction in the GFDL "SKYHI" general circulation model. *J. Atmos. Sci.* **40**(12): 2768–2784.

Anstey JA, Shepherd TG. 2008. Response of the northern stratospheric polar vortex to the seasonal alignment of QBO phase transitions. *Geophys. Res. Lett.* **35**(22).

Anstey JA, Shepherd TG. 2013. High-latitude influence of the quasi-biennial oscillation. *Q. J. R. Meteorol. Soc.* In press.

Anstey JA, Shepherd TG, Scinocca JF. 2010. Influence of the quasi-biennial oscillation on the extratropical winter stratosphere in an atmospheric general circulation model and in reanalysis data. *J. Atmos. Sci.* **67**(5): 1402–1419.

Arblaster JM, Meehl GA. 2006. Contributions of external forcings to Southern Annular Mode trends. *J. Climate* **19**(12): 2896–2905.

Balachandran NK, Plumb RA, Suozzo R, Rind D. 1991. The QBO and stratospheric warming - model results. *Journal of Geomagnetism and Geoelectricity* **43**(Part 2, S): 741–757.

Balachandran NK, Rind D. 1995. Modeling the effects of UV variability and the QBO on the troposphere-stratosphere system. Part I:. The middle atmosphere. *J. Climate* **8**(8): 2058–2079.

Baldwin MP, Dunkerton TJ. 1989. Observations and statistical simulations of a proposed solar cycle/QBO/weather relationship. *Geophys. Res. Lett.* **16**(8): 863–866.

Baldwin MP, Dunkerton TJ. 1991. Quasi-biennial oscillation above 10 mb. *Geophys. Res. Lett.* **18**: 1205–1208.

Baldwin MP, Dunkerton TJ. 1998. Quasi-biennial modulation of the Southern Hemisphere stratospheric polar vortex. *Geophys. Res. Lett.* **25**: 3343–3346.

Baldwin MP, Dunkerton TJ. 1999. Propagation of the Arctic oscillation from the stratosphere to the troposphere. *J. Geophys. Res.* **104**: 30 937–30 946.

Baldwin MP, Dunkerton TJ. 2001. Stratospheric harbingers of anomalous weather regimes. *Science* **294**: 581–584.

Baldwin MP, Gray LJ. 2005. Tropical stratospheric zonal winds in ECMWF ERA-40 reanalysis, rocketsonde data, and rawinsonde data. *Geophys. Res. Lett.* **32**(9): L09806, doi:10.1029/2004GL022328.

Baldwin MP, Gray LJ, Dunkerton TJ, Hamilton K, Haynes PH, Randel WJ, Holton JR, Alexander MJ, Hirota I, Horinouchi T, *et al.* 2001. The quasi-biennial oscillation. *Rev. Geophys.* **39**(2): 179–229.

Baldwin MP, O'Sullivan D. 1995. Stratospheric effects of ENSO-related tropospheric circulation anomalies. *J. Climate* **8**(4): 649–667.

Baldwin MP, Stephenson DB, Joliffe IT. 2009. Spatial weighting and iterative projection methods for EOFs. *J. Climate* **22**(2): 234–243.

Baldwin MP, Stephenson DB, Thompson DWJ, Dunkerton TJ, Charlton AJ, O'Neill A. 2003a. Stratospheric memory and skill of extended-range weather forecasts. *Science* **301**(5633): 636–640.

Baldwin MP, Thompson DWJ. 2009. A critical comparison of stratosphere-troposphere coupling indices. *Q. J. R. Meteorol. Soc.* **135**(644): 1661–1672.

Baldwin MP, Thompson DWJ, Shuckburgh EF, Norton WA, Gillett NP. 2003b. Weather from the stratosphere? *Science* **301**(5631): 317–319.

Bell CJ, Gray LJ, Charlton-Perez AJ, Joshi MM, Scaife AA. 2009. Stratospheric communication of El Niño teleconnections to European winter. *J. Climate* **22**: 4083–4096.

Birner T, Williams PD. 2008. Sudden stratospheric warmings as noise-induced transitions. *J. Atmos. Sci.* **65**(10): 3337–3343.

Branstator G, Selten F. 2009. "Modes of variability" and climate change. *J. Climate* **22**(10): 2639–2658.

Butchart N, Clough SA, Palmer TN, Trevelyan PJ. 1982. Simulations of an observed stratospheric warming with quasigeostrophic refractive index as a model diagnostic. *Q. J. R. Meteorol. Soc.* **108**(457): 475–502.

Cagnazzo C, Manzini E. 2009. Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European region. *J. Climate* **22**(5): 1223–1238.

Calvo N, Giorgetta MA, Garcia-Herrera R, Manzini E. 2009. Nonlinearity of the combined warm ENSO and QBO effects on the Northern Hemisphere polar vortex in MAECHAM5 simulations. *J. Geophys. Res.* **114**(D13): D13109, doi: 10.1029/2008JD011445.

Calvo N, Giorgetta MA, Peña Ortiz C. 2007. Sensitivity of the boreal winter circulation in the middle atmosphere to the quasi-biennial oscillation in MAECHAM5 simulations. *J. Geophys. Res.* **112**: D10124, doi:10.1029/2006JD007844.

Calvo N, Marsh DR. 2011. The combined effects of ENSO and the 11 year solar cycle on the Northern Hemisphere polar stratosphere. *J. Geophys. Res.* **116**(D23).

Camp CD, Tung KK. 2007a. Stratospheric polar warming by ENSO in winter: A statistical study. *Geophys. Res. Lett.* **34**(4): L04809, doi:10.1029/2006GL028521.

Camp CD, Tung KK. 2007b. The influence of the solar cycle and QBO on the late-winter stratospheric polar vortex. *J. Atmos. Sci.* **64**(4): 1267–1283.

Chan CJ, Plumb RA. 2009. The response to stratospheric forcing and its dependence on the state of the troposphere. *J. Atmos. Sci.* **66**(7): 2107–2115.

Charlton AJ, O'Neill A, Lahoz WA, Massacand AC. 2004. Sensitivity of tropospheric forecasts to stratospheric initial conditions. *Q. J. R. Meteorol. Soc.* **130**(600): 1771–1792.

Charlton AJ, O'Neill A, Stephenson DB, Lahoz WA, Baldwin MP. 2003. Can knowledge of the state of the stratosphere be used to improve statistical forecasts of the troposphere? *Q. J. R. Meteorol. Soc.* **129**(595): 3205–3224.

Charlton AJ, Polvani LM. 2007. A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *J. Climate* **20**: 449–469.

Charlton-Perez AJ, Baldwin MP, Birner T, Black RX, Butler AH, Calvo N, Davis NA, Gerber EP, Gillett N, Hardiman S, *et al.* 2013. On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *J. Geophys. Res.* **118**: 2494–2505.

Charney JG, Drazin PG. 1961. Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.* **66**(1): 83–109.

Chen G, Zurita-Gotor P. 2008. The tropospheric jet response to prescribed zonal forcing in an idealized atmospheric model. *J. Atmos. Sci.* **65**(7): 2254–2271.

Christiansen B. 1999. Stratospheric vacillations in a general circulation model. *J. Atmos. Sci.* **56**(12): 1858–1872.

Christiansen B. 2003. Evidence for nonlinear climate change: Two stratospheric regimes and a regime shift. *J. Climate* **16**(22): 3681–3690.

Christiansen B. 2010. Stratospheric bimodality: Can the equatorial QBO explain the regime behavior of the NH winter vortex? *J. Climate* **23**(14): 3953–3966.

Cionni I, Eyring V, Lamarque JF, Randel WJ, Stevenson DS, Wu F, Bodeker G, Shepherd TG, Shindell DT, Waugh D. 2011. Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing. *Atmos. Chem. Phys.* **11**(21): 11 267–11 292.

Cnossen I, Lu H, Bell CJ, Gray LJ, Joshi MM. 2011. Solar signal propagation: The role of gravity waves and stratospheric sudden warmings. *J. Geophys. Res.* **116**: D02118, doi:10.1029/2010JD014535.

Cohen NY, Gerber EP, Bühler O. 2013. Compensation between resolved and unresolved wave driving in the stratosphere: Implications for downward control. *J. Atmos. Sci.* **70**: 3780–3898.

Cohen NY, Gerber EP, Bühler O. 2014. What drives the Brewer-Dobson circulation? *J. Atmos. Sci.* Submitted.

Coughlin K, Gray LJ. 2009. A continuum of sudden stratospheric warmings. *J. Atmos. Sci.* **66**: 531–540.

Crooks SA, Gray LJ. 2005. Characterization of the 11-year solar signal using a multiple regression analysis of the ERA-40 dataset. *J. Climate* **18**(7): 996–1015.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P, *et al.* 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**(656): 553–597.

Deser C, Magnusdottir G, Saravanan R, Phillips A. 2004. The effects of North Atlantic SST and sea ice anomalies on the winter circulation in CCM3. Part II: Direct and indirect components of the response. *J. Climate* **17**(5): 877–889.

Dunkerton T. 1978. On the mean meridional mass motions of the stratosphere and mesosphere. *J. Atmos. Sci.* **35**(12): 2325–2333.

Dunkerton T, Hsu CPF, McIntyre ME. 1981. Some Eulerian and Lagrangian diagnostics for a model stratospheric warming. *J. Atmos. Sci.* **38**(4): 819–844.

Dunkerton TJ. 1990. Annual variation of deseasonalized mean flow acceleration in the equatorial lower stratosphere. *J. Meteor. Soc. Japan* **68**: 499–508.

Dunkerton TJ. 1997. The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res.* **102**(D22): 26 053–26 076.

Dunkerton TJ, Baldwin MP. 1991. Quasi-biennial modulation of planetary-wave fluxes in the Northern Hemisphere winter. *J. Atmos. Sci.* **48**(8): 1043–1061.

Dunkerton TJ, Delisi DP, Baldwin MP. 1988. Distribution of major stratospheric warmings in relation to the quasi-biennial oscillation. *Geophys. Res. Lett.* **15**(2): 136–139.

Edmon HJ, Hoskins BJ, McIntyre ME. 1980. Eliassen-Palm cross sections for the troposphere. *J. Atmos. Sci.* **37**(12): 2600–2616.

Edwards JM, Thelen JC, Ingram WJ. 2004. Unified Model documentation paper No 23: The radiation code. Available from http://cms.ncas.ac.uk/wiki/Docs/MetOfficeDocs.

Efron B, Tibshirani RJ. 1993. *An introduction to the bootstrap*. Chapman and Hall: New York.

Eliassen A. 1951. Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophysica Norvegica* **5**: 19.

Fairlie TDA, Fisher M, O'Neill A. 1990. The development of narrow baroclinic zones and other small-scale structure in the stratosphere during simulated major warmings. *Q. J. R. Meteorol. Soc.* **116**(492): 287–315.

Fisher M. 1987. The Met O 20 Stratosphere-mesosphere Model. *Met Office Dynamical Climatology Tech. Note 52* .

Fisher M, O'Neill A, Sutton R. 1993. Rapid descent of mesospheric air into the stratospheric polar vortex. *Geophys. Res. Lett.* **20**(12): 1267–1270.

Frame THA, Gray LJ. 2010. The 11-yr solar cycle in ERA-40 data: An update to 2008. *J. Climate* **23**(8): 2213–2222.

Free M, Seidel DJ. 2009. Observed El Niño–Southern Oscillation temperature signal in the stratosphere. *J. Geophys. Res.* **114**: D23108, doi:10.1029/2009JD012420.

Fritts DC, Alexander MJ. 2003. Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys.* **41**(1).

Garcia RR. 1987. On the mean meridional circulation of the middle atmosphere. *J. Atmos. Sci.* **44**(24): 3599–3609.

Garcia-Herrera R, Calvo N, Garcia RR, Giorgetta MA. 2006. Propagation of ENSO temperature signals into the middle atmosphere: A comparison of two general circulation models and ERA-40 reanalysis data. *J. Geophys. Res.* **111**: D06101, doi:10.1029/2005JD006061.

Garfinkel CI, Butler AH, Waugh DW, Hurwitz MM, Polvani LM. 2012a. Why might stratospheric sudden warmings occur with similar frequency in El Niño and La Niña winters? *J. Geophys. Res.* **117**(D19).

Garfinkel CI, Hartmann DL. 2007. The effects of the quasi-biennial oscillation and the El Nino Southern Oscillation on polar temperatures in the stratosphere. *J. Geophys. Res.* **112**: D19112, doi:10.1029/2007JD008481.

Garfinkel CI, Hartmann DL. 2008. Different ENSO teleconnections and their effects on the stratospheric polar vortex. *J. Geophys. Res.* **113**: D18114, doi: 10.1029/2008JD009920.

Garfinkel CI, Hartmann DL. 2010. Influence of the quasi-biennial oscillation on the North Pacific and El Niño teleconnections. *J. Geophys. Res.* **115**: D20116, doi: 10.1029/2010JD014181.

Garfinkel CI, Shaw TA, Hartmann DL, Waugh DW. 2012b. Does the Holton-Tan mechanism explain how the quasi-biennial oscillation modulates the Arctic polar vortex? *J. Atmos. Sci.* **69**(5): 1713–1733.

Geller MA, Alexander MJ, Love PT, Bacmeister J, Ern M, Hertzog A, Manzini E, Preusse P, Sato K, Scaife AA, *et al.* 2013. A comparison between gravity wave momentum fluxes in observations and climate models. *J. Climate* **26**: 6383–6405.

Gray L. 2010. Equatorial Dynamics. *The Stratosphere: Dynamics, Transport and Chemistry, Geophys. Monogr. Series* **190**: 93–107.

Gray LJ. 2003. The influence of the equatorial upper stratosphere on stratospheric sudden warmings. *Geophys. Res. Lett.* **30**(4): 1166, doi:10.1029/2002GL016430.

Gray LJ, Crooks S, Pascoe C, Sparrow S, Palmer M. 2004. Solar and QBO influences on the timing of stratospheric sudden warmings. *J. Atmos. Sci.* **61**(23): 2777–2796.

Gray LJ, Drysdale EF, Lawrence BN, Dunkerton TJ. 2001a. Model studies of the interannual variability of the northern-hemisphere stratospheric winter circulation: The role of the quasi-biennial oscillation. *Q. J. R. Meteorol. Soc.* **127**(574): 1413–1432.

Gray LJ, Phipps SJ, Dunkerton TJ, Baldwin MP, Drysdale EF, Allen MR. 2001b. A data study of the influence of the equatorial upper stratosphere on northern-hemisphere stratospheric sudden warmings. *Q. J. R. Meteorol. Soc.* **127**(576): 1985–2003.

Gray LJ, Sparrow S, Juckes M, O'Neill A, Andrews DG. 2003. Flow regimes in the winter stratosphere of the Northern Hemisphere. *Q. J. R. Meteorol. Soc.* **129**(589): 925–945.

Gritsun A, Branstator G. 2007. Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem. *J. Atmos. Sci.* **64**(7): 2558–2575.

Gritsun AS, Dymnikov VP. 1999. Barotropic atmosphere response to small external actions: Theory and numerical experiments. *Izvestiya, Atmospheric and Oceanic Physics* **35**(5): 511–525.

Hamilton K. 1993a. An examination of observed Southern Oscillation effects in the Northern Hemisphere stratosphere. *J. Atmos. Sci.* **50**(20).

Hamilton K. 1993b. A general circulation model simulation of El Niño effects in the extratropical Northern Hemisphere stratosphere. *Geophys. Res. Lett.* **20**(17): 1803–1806.

Hamilton K. 1998. Effects of an imposed quasi-biennial oscillation in a comprehensive troposphere-stratosphere-mesosphere general circulation model. *J. Atmos. Sci.* **55**(14): 2393–2418.

Hampson J, Haynes P. 2004. Phase alignment of the tropical stratospheric QBO in the annual cycle. *J. Atmos. Sci.* **61**(21): 2627–2637.

Hampson J, Haynes P. 2006. Influence of the equatorial QBO on the extratropical stratosphere. *J. Atmos. Sci.* **63**(3): 936–951.

Hannachi A, Mitchell D, Gray L, Charlton-Perez A. 2011. On the use of geometric moments to examine the continuum of sudden stratospheric warmings. *J. Atmos. Sci.* **68**(3): 657–674, doi:10.1175/2010JAS3585.1.

Hardiman SC, Butchart N, Hinton TJ, Osprey SM, Gray LJ. 2012. The effect of a well-resolved stratosphere on surface climate: Differences between CMIP5 simulations with high and low top versions of the Met Office climate model. *J. Climate* **25**(20): 7083–7099.

Hartley DE, Villarin JT, Black RX, Davis CA. 1998. A new perspective on the dynamical link between the stratosphere and troposphere. *Nature* **391**(6666): 471–474.

Harvey VL, Hitchman MH. 1996. A climatology of the Aleutian high. *J. Atmos. Sci.* **53**: 2088–2102.

Hasebe F. 1994. Quasi-biennial oscillations of ozone and diabatic circulation in the equatorial stratosphere. *J. Atmos. Sci.* **51**(5): 729–745.

Hauck C, Wirth V. 2001. Diagnosing the impact of stratospheric planetary wave breaking in a linear model. *J. Atmos. Sci.* **58**(11): 1357–1370.

Haynes P. 2005. Stratospheric Dynamics. *Annual Review of Fluid Mechanics* **37**: 263–293.

Haynes PH. 1985. Nonlinear instability of a Rossby-wave critical layer. *J. Fluid Mech.* **161**: 493–511.

Haynes PH. 1989. The effect of barotropic instability on the nonlinear evolution of a Rossby-wave critical layer. *J. Fluid Mech.* **207**: 231–266.

Haynes PH. 1998. The latitudinal structure of the quasi-biennial oscillation. *Q. J. R. Meteorol. Soc.* **124**(552): 2645–2670.

Haynes PH. 2003. Critical layers. *Encyclopedia of Atmospheric Sciences* **2**: 582–589.

Haynes PH, McIntyre ME, Shepherd TG, Marks CJ, Shine K. 1991. On the 'downward control' of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.* **48**(4): 651–679.

Hitchcock P, Shepherd TG, Manney GL. 2013. Statistical characterization of Arctic polar-night jet oscillation events. *J. Climate* **26**(6): 2096–2116.

Hitchman MH, Huesmann AS. 2009. Seasonal influence of the quasi-biennial oscillation on stratospheric jets and Rossby wave breaking. *J. Atmos. Sci.* **66**(4): 935–946.

Holton JR, Austin J. 1991. The influence of the equatorial QBO on sudden stratospheric warmings. *J. Atmos. Sci.* **48**(4): 607–618.

Holton JR, Haynes PH, McIntyre ME, Douglass AR, Rood RB, Pfister L. 1995. Stratosphere-troposphere exchange. *Rev. Geophys.* **33**: 403–439.

Holton JR, Lindzen RS. 1972. An updated theory for the quasi-biennial cycle of the tropical stratosphere. *J. Atmos. Sci.* **29**(6): 1076–1080.

Holton JR, Mass C. 1976. Stratospheric vacillation cycles. *J. Atmos. Sci.* **33**: 2218–2225.

Holton JR, Tan HC. 1980. The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb. *J. Atmos. Sci.* **37**(10): 2200–2208.

Holton JR, Tan HC. 1982. The quasi-biennial oscillation in the Northern Hemisphere lower stratosphere. *J. Meteor. Soc. Japan* **60**: 140–148.

Hu YY, Tung KK. 2002. Tropospheric and equatorial influences on planetary-wave amplitude in the stratosphere. *Geophys. Res. Lett.* **29**(2), doi: 10.1029/2001GL013762.

Hurwitz MM, Newman PA, Oman LD, Molod AM. 2011. Response of the Antarctic stratosphere to two types of El Niño events. *J. Atmos. Sci.* **68**(4): 812–822, doi: 10.1175/2011JAS3606.1.

Ineson S, Scaife AA. 2009. The role of the stratosphere in the European climate response to El Niño. *Nature Geoscience* **2**(1): 32–36.

Ineson S, Scaife AA, Knight JR, Manners JC, Dunstone NJ, Gray LJ, Haigh JD. 2011. Solar forcing of winter climate variability in the Northern Hemisphere. *Nature Geoscience* **4**(11): 753–757.

Jones D, Schneider HR, McElroy MB. 1998. Effects of the quasi-biennial oscillation on the zonally averaged transport of tracers. *J. Geophys. Res.* **103**(D10): 11 235–11 249.

Juckes MN, McIntyre ME. 1987. A high-resolution one-layer model of breaking planetary waves in the stratosphere. *Nature* **328**: 590–596.

Jung T, Barkmeijer J. 2006. Sensitivity of the tropospheric circulation to changes in the strength of the stratospheric polar vortex. *Monthly Weather Review* **134**(8): 2191–2207.

Karoly DJ, Hoskins BJ. 1982. Three dimensional propagation of planetary waves. *J. Meteor. Soc. Japan* **60**: 109–123.

Karpetchko A, Kyrö E, Knudsen BM. 2005. Arctic and Antarctic polar vortices 1957-2002 as seen from the ERA-40 reanalyses. *J. Geophys. Res.* **110**: D21 109.

Kawatani Y, Watanabe S, Sato K, Dunkerton TJ, Miyahara S, Takahashi M. 2010. The roles of equatorial trapped waves and internal inertia-gravity waves in driving the quasi-biennial oscillation. Part I: zonal mean wave forcing. *J. Atmos. Sci.* **67**(4): 963–980.

Killworth PD, McIntyre ME. 1985. Do Rossby-wave critical layers absorb, reflect, or over-reflect? *J. Fluid Mech.* **161**: 449–492.

Kinnersley JS. 1999. Seasonal asymmetry of the low-and middle-latitude QBO circulation anomaly. *J. Atmos. Sci.* **56**(9): 1140–1153.

Kinnersley JS, Tung KK. 1999. Mechanisms for the extratropical QBO in circulation and ozone. *J. Atmos. Sci.* **56**(12): 1942–1962.

Kodera K. 1991. The solar and equatorial qbo influences on the stratospheric circulation during the early Northern-Hemisphere winter. *Geophys. Res. Lett.* **18**(6): 1023–1026, doi:10.1029/90GL02298.

Kodera K. 1993. Quasi-decadal modulation of the influence of the equatorial quasi-biennial oscillation on the north polar stratospheric temperatures. *J. Geophys. Res.* **98**(D4): 7245–7250.

Kodera K. 1995. On the origin and nature of the interannual variability of the winter stratospheric circulation in the Northern-hemisphere. *J. Geophys. Res.* **100**(D7): 14 077–14 087.

Kodera K, Kuroda Y. 2002. Dynamical response to the solar cycle. *J. Geophys. Res.* **107**(D24): 4749, doi:10.1029/2002JD002224.

Kolstad EW, Breiteig T, Scaife AA. 2010. The association between stratospheric weak polar vortex events and cold air outbreaks in the Northern Hemisphere. *Q. J. R. Meteorol. Soc.* **136**(649): 886–893.

Kuai L, Shia RL, Jiang X, Tung KK, Yung YL. 2009. Nonstationary synchronization of equatorial QBO with SAO in observations and a model. *J. Atmos. Sci.* **66**(6): 1654–1664.

Kushner PJ. 2010. Annular modes of the troposphere and stratosphere. *The Stratosphere: Dynamics, Transport and Chemistry, Geophys. Monogr. Series* **190**: 59–91.

Kushner PJ, Held IM, Delworth TL. 2001. Southern Hemisphere atmospheric circulation response to global warming. *J. Climate* **14**(10): 2238–2249.

Kushner PJ, Polvani LM. 2004. Stratosphere-troposphere coupling in a relatively simple AGCM: The role of eddies. *J. Climate* **17**(3): 629–639.

Labitzke K. 1982. On the interannual variability of the middle stratosphere during the northern winters. *J. Meteor. Soc. Japan* **60**: 124–139.

Labitzke K. 1987. Sunspots, the QBO, and the stratospheric temperature in the north polar region. *Geophys. Res. Lett.* **14**(5): 535–537.

Labitzke K. 2005. On the solar cycle-QBO relationship: a summary. *Journal of Atmospheric and Solar-Terrestrial Physics* **67**(1-2): 45–54.

Labitzke K, Loon HV. 1988. Associations between the 11-year solar cycle, the QBO and the atmosphere. Part I: the troposphere and stratosphere in the northern hemisphere in winter. *Journal of Atmospheric and Terrestrial Physics* **50**(3): 197–206.

Lahoz WA. 2000. Northern Hemisphere winter stratospheric variability in the Met. Office Unified Model. *Q. J. R. Meteorol. Soc.* **126**(568): 2605–2630.

Leith CE. 1975. Climate response and fluctuation dissipation. *J. Atmos. Sci.* **32**(10): 2022–2026.

Li D, Shine KP, Gray LJ. 1995. The role of ozone-induced diabatic heating anomalies in the quasi-biennial oscillation. *Q. J. R. Meteorol. Soc.* **121**(524): 937–943.

Limpasuvan V, Hartmann DL. 2000. Wave-maintained annular modes of climate variability. *J. Climate* **13**(24): 4414–4429.

Limpasuvan V, Thompson DWJ, Hartmann DL. 2004. The life cycle of the Northern Hemisphere sudden stratospheric warmings. *J. Climate* **17**: 2584–2596.

Lindzen RS, Holton JR. 1968. A theory of the quasi-biennial oscillation. *J. Atmos. Sci.* **25**: 1095–1107.

Lorenz EN. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**(2): 130–141.

Lu H, Baldwin MP, Gray LJ, Jarvis MJ. 2008. Decadal-scale changes in the effect of the QBO on the northern stratospheric polar vortex. *J. Geophys. Res.* **113**: D10114, doi:10.1029/2007JD009647.

Lu H, Gray LJ, Baldwin MP, Jarvis MJ. 2009. Life cycle of the QBO-modulated 11-year solar cycle signals in the Northern Hemispheric winter. *Q. J. R. Meteorol. Soc.* **135**(641): 1030–1043.

Manzini E, Giorgetta MA, Esch M, Kornblueh L, Roeckner E. 2006. The influence of sea surface temperatures on the northern winter stratosphere: Ensemble simulations with the MAECHAM5 model. *J. Climate* **19**(16): 3863–3881.

Marshall AG, Scaife AA. 2009. Impact of the QBO on surface winter climate. *J. Geophys. Res.* **114**: D18110, doi:10.1029/2009JD011737.

Martin GM, Bellouin N, Collins WJ, Culverwell ID, Halloran PR, Hardiman SC, Hinton TJ, Jones CD, McDonald RE, McLaren AJ, *et al.* 2011. The HadGEM2 family of Met Office Unified Model Climate configurations. *Geoscientific Model Development Discussions* **4**: 765–841.

Matsuno T. 1970. Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *J. Atmos. Sci.* **27**: 871–883.

Matsuno T. 1971. A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.* **28**: 1479–1494.

Matthes K, Langematz U, Gray LL, Kodera K, Labitzke K. 2004. Improved 11-year solar signal in the Freie Universität Berlin Climate Middle Atmosphere Model (FUB-CMAM). *J. Geophys. Res.* **109**: D06101, doi:10.1029/2003JD004012.

Matthewman NJ, Esler JG, Charlton-Perez AJ, Polvani LM. 2009. A new look at stratospheric sudden warmings. Part III: Polar vortex evolution and vertical structure. *J. Climate* **22**(6): 1566–1585.

Maycock AC, Keeley SPE, Charlton-Perez AJ, Doblas-Reyes FJ. 2011. Stratospheric circulation in seasonal forecasting models: implications for seasonal prediction. *Climate Dynamics* **36**(1): 309–321.

McIntyre ME. 1982. How well do we understand the dynamics of stratospheric warmings. *J. Meteor. Soc. Japan* **60**: 37–65.

McIntyre ME, Palmer TN. 1983. Breaking planetary waves in the stratosphere. *Nature* **305**(5935): 593–600.

McIntyre ME, Palmer TN. 1984. The 'surf zone' in the stratosphere. *Journal of Atmospheric and Terrestrial Physics* **46**(9): 825–849.

McLandress C, McFarlane NA. 1993. Interactions between orographic gravity wave drag and forced stationary planetary waves in the winter northern hemisphere middle atmosphere. *J. Atmos. Sci.* **50**: 1966–1990.

Mitchell DM. 2010. Extreme variability of the stratospheric polar vortex. PhD thesis, University of Reading.

Mitchell DM, Charlton-Perez AJ, Gray LJ. 2011a. Characterising the variability and extremes of the stratospheric polar vortices using 2D moments. *J. Atmos. Sci.* **68**: 1194–1213, doi:10.1175/2010JAS3555.1.

Mitchell DM, Gray LJ, Anstey J, Baldwin MP, Charlton-Perez AJ. 2013. The influence of stratospheric vortex displacements and splits on surface climate. *J. Climate* **26**: 2668–2682.

Mitchell DM, Gray LJ, Charlton-Perez AJ. 2011b. The structure and evolution of the stratospheric vortex in response to natural forcings. *J. Geophys. Res.* **116**: D15110, doi:10.1029/2011JD015788.

Mitchell DM, Osprey SM, Gray LJ, Butchart N, Hardiman SC, Charlton-Perez AJ, Watson P. 2012. The effect of climate change on the variability of the Northern Hemisphere stratospheric polar vortex. *J. Atmos. Sci.* **69**(8): 2608–2618.

Naito Y. 2002. Planetary wave diagnostics on the QBO effects on the deceleration of the polar-night jet in the Southern Hemisphere. *J. Meteor. Soc. Japan* **80**(4B): 985–995.

Naito Y, Hirota I. 1997. Interannual variability of the northern winter stratospheric circulation related to the QBO and the solar cycle. *J. Meteor. Soc. Japan* **75**(4): 925–937.

Naito Y, Taguchi M, Yoden S. 2003. A parameter sweep experiment on the effects of the equatorial QBO on stratospheric sudden warming events. *J. Atmos. Sci.* **60**(11): 1380–1394.

Naoe H, Shibata K. 2010. Equatorial quasi-biennial oscillation influence on northern winter extratropical circulation. *J. Geophys. Res.* **115**: D19102, doi: 10.1029/2009JD012952.

Newman PA. 2010. Chemistry and dynamics of the Antarctic ozone hole. *The Stratosphere: Dynamics, Transport and Chemistry, Geophys. Monogr. Series* **190**: 157–171.

Newman PA, Nash ER, Rosenfield JE. 2001. What controls the temperature of the arctic stratosphere during the spring? *J. Geophys. Res.* **106**(D17): 19 999–20 010.

Niwano M, Takahashi M. 1998. The influence of the equatorial QBO on the Northern Hemisphere winter circulation of a GCM. *J. Meteor. Soc. Japan* **76**(3): 453–461.

Norton WA. 2003. Sensitivity of Northern Hemisphere surface climate to simulation of the stratospheric polar vortex. *Geophys. Res. Lett.* **30**(12): 1627, doi: 10.1029/2003GL016958.

O'Neill A, Pope VD. 1988. Simulations of linear and nonlinear disturbances in the stratosphere. *Q. J. R. Meteorol. Soc.* **114**(482): 1063–1110.

O'Neill A, Pope VD. 1993. The coupling between radiation and dynamics in the stratosphere. *Advances in Space Research* **13**(1): 351–358.

O'Neill A, Taylor BF. 1979. Study of the major stratospheric warming of 1976–77. *Q. J. R. Meteorol. Soc.* **105**(443): 71–92.

Osprey SM, Gray LJ, Hardiman SC, Butchart N, Hinton TJ. 2013. Stratospheric variability in Twentieth-Century CMIP5 simulations of the Met Office climate model: High-top versus low-top. *J. Climate* **26**(5): 1607–1625.

O'Sullivan D, Dunkerton TJ. 1994. Seasonal development of the extratropical QBO in a numerical model of the middle atmosphere. *J. Atmos. Sci.* **51**(24): 3706–3706.

O'Sullivan D, Salby ML. 1990. Coupling of the quasi-biennial oscillation and the extratropical circulation in the stratosphere through planetary wave transport. *J. Atmos. Sci.* **47**(5): 650–673.

O'Sullivan D, Young RE. 1992. Modeling the quasi-biennial oscillation's effect on the winter stratospheric circulation. *J. Atmos. Sci.* **49**(24): 2437–2448.

Palmer TN. 1981a. Aspects of stratospheric sudden warmings studied from a transformed Eulerian-mean viewpoint. *J. Geophys. Res.* **86**(C10): 9679–9687.

Palmer TN. 1981b. Diagnostic study of a wavenumber-2 stratospheric sudden warming in a transformed Eulerian-mean formalism. *J. Atmos. Sci.* **38**(4): 844–855.

Palmer TN. 1999. A nonlinear dynamical perspective on climate prediction. *J. Climate* **12**(2): 575–591.

Palmer TN, Weisheimer A. 2011. Diagnosing the causes of bias in climate models - why is it so hard? *Geophysical and Astrophysical Fluid Dynamics* **105**(2-3): 351–365.

Pascoe CL, Gray LJ, Crooks SA, Juckes MN, Baldwin MP. 2005. The quasi-biennial oscillation: Analysis using ERA-40 data. *J. Geophys. Res.* **110**: D08105, doi: 10.1029/2004JD004941.

Pendlebury D, Shepherd TG. 2003. Planetary-wave-induced transport in the stratosphere. *J. Atmos. Sci.* **60**: 1456–1470.

Perlwitz J, Harnik N. 2003. Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection. *J. Climate* **16**(18): 3011–3026.

Plumb RA. 1977. The interaction of two internal waves with the mean flow: Implications for the theory of the quasi-biennial oscillation. *J. Atmos. Sci.* **34**: 1847–1858.

Plumb RA. 1982. Zonally symmetric Hough modes and meridional circulations in the middle atmosphere. *J. Atmos. Sci.* **39**: 983–991.

Plumb RA. 1984. The quasi-biennial oscillation. In: *Dynamics of the Middle Atmosphere*, vol. 1. p. 217.

Plumb RA. 2010. Planetary waves and the extratropical winter stratosphere. *The Stratosphere: Dynamics, Transport and Chemistry, Geophys. Monogr. Series* **190**: 23–39.

Plumb RA, Bell RC. 1982. A model of the quasi-biennial oscillation on an equatorial beta-plane. *Q. J. R. Meteorol. Soc.* **108**(456): 335–352.

Polvani LM, Kushner PJ. 2002. Tropospheric response to stratospheric perturbations in a relatively simple general circulation model. *Geophys. Res. Lett.* **29**(7), doi: 10.1029/2001GL014284.

Polvani LM, Waugh DW, Plumb RA. 1995. On the subtropical edge of the stratospheric surf zone. *J. Atmos. Sci.* **52**(9): 1288–1309.

Randel W, Udelhofen P, Fleming E, Geller M, Gelman M, Hamilton K, Karoly D, Ortland D, Pawson S, Swinbank R, *et al.* 2004. The SPARC intercomparison of middle-atmosphere climatologies. *J. Climate* **17**: 986–1003.

Randel W, Wu F, Swinbank R, Nash J, O'Neill A. 1999. Global QBO circulation derived from UKMO stratospheric analyses. *J. Atmos. Sci.* **56**(4): 457–474.

Read PL, Castrejón-Pita AA. 2012. Phase synchronization between stratospheric and tropospheric quasi-biennial and semi-annual oscillations. *Q. J. R. Meteorol. Soc.* **138**(666): 1338–1349.

Richter JH, Matthes K, Calvo N, Gray LJ. 2011. Influence of the quasi-biennial oscillation and El Nino Southern Oscillation on the frequency of sudden stratospheric warmings. *J. Geophys. Res.* **116**: D20111, doi:10.1029/2011JD015757.

Richter JH, Sassi F, Garcia RR, Matthes K, Fischer CA. 2008. Dynamics of the middle atmosphere as simulated by the Whole Atmosphere Community Climate Model, version 3 (WACCM3). *J. Geophys. Res.* **113**: D08101, doi:10.1029/2007JD009269.

Rind D, Suozzo R, Balachandran NK. 1988. The GISS Global Climate-Middle Atmosphere Model. Part II: Model variability due to interactions between planetary waves, the mean circulation and gravity wave drag. *J. Atmos. Sci.* **45**(3): 371–386.

Ring MJ, Plumb RA. 2007. Forced annular mode patterns in a simple atmospheric general circulation model. *J. Atmos. Sci.* **64**(10): 3611–3626.

Ring MJ, Plumb RA. 2008. The response of a simplified GCM to axisymmetric forcings: Applicability of the fluctuation-dissipation theorem. *J. Atmos. Sci.* **65**(12): 3880–3898.

Ruzmaikin A, Feynman J, Jiang X, Yung YL. 2005. Extratropical signature of the quasi-biennial oscillation. *J. Geophys. Res.* **110**: D11111, doi: 10.1029/2004JD005382.

Sassi F, Kinnison D, Boville BA, Garcia RR, Roble R. 2004. Effect of El Nino–Southern Oscillation on the dynamical, thermal, and chemical structure of the middle atmosphere. *J. Geophys. Res.* **109**(D17): D17108, doi: 10.1029/2003JD004434.

Scaife AA, Butchart N, Warner CD, Swinbank R. 2002. Impact of a spectral gravity wave parameterization on the stratosphere in the Met Office Unified Model. *J. Atmos. Sci.* **59**(9): 1473–1489.

Scaife AA, James IN. 2000. Response of the stratosphere to interannual variability of tropospheric planetary waves. *Q. J. R. Meteorol. Soc.* **126**(562): 275–297.

Scaife AA, Knight JR. 2008. Ensemble simulations of the cold European winter of 2005-2006. *Q. J. R. Meteorol. Soc.* **134**(636): 1647–1659.

Scott RK, Dritschel DG. 2006. Vortex–Vortex Interactions in the Winter Stratosphere. *J. Atmos. Sci.* **63**(2): 726–740.

Scott RK, Haynes PH. 2000. Internal vacillations in stratosphere-only models. *J. Atmos. Sci.* **57**(19): 3233–3250.

Scott RK, Polvani LM. 2004. Stratospheric control of upward wave flux near the tropopause. *Geophys. Res. Lett.* **31**(2): L02 115.

Scott RK, Polvani LM. 2006. Internal variability of the winter stratosphere. Part I: Time-independent forcing. *J. Atmos. Sci.* **63**(11): 2758–2776.

Scott RK, Polvani LM, Waugh DW. 2008. Internal variability of the winter stratosphere. Part II: Time-dependent forcing. *J. Atmos. Sci.* **65**(7): 2375–2388.

Shaw TA, Shepherd TG. 2008. Atmospheric science: Raising the roof. *Nature Geoscience* **1**(1): 12–13.

Shepherd TG. 2000. The middle atmosphere. *Journal of Atmospheric and Solar-Terrestrial Physics* **62**(17): 1587–1601.

Shepherd TG. 2002. Issues in stratosphere-troposphere coupling. *J. Meteor. Soc. Japan* **80**(4B): 769–792.

Shepherd TG. 2007. Transport in the middle atmosphere. *J. Meteor. Soc. Japan* **85B**: 165–191.

Shepherd TG. 2008. Dynamics, stratospheric ozone, and climate change. *Atmosphere-Ocean* **46**(1): 117–138.

Shepherd TG, Semeniuk K, Koshyk JN. 1996. Sponge layer feedbacks in middle-atmosphere models. *J. Geophys. Res.* **101**(D18): 23 447–23 464.

Shepherd TG, Shaw TA. 2004. The angular momentum constraint on climate sensitivity and downward influence in the middle atmosphere. *J. Atmos. Sci.* **61**(23): 2899–2908.

Shindell DT, Schmidt GA. 2004. Southern Hemisphere climate response to ozone changes and greenhouse gas increases. *Geophys. Res. Lett.* **31**(18): L18 209.

Shine KP. 1987. The middle atmosphere in the absence of dynamical heat fluxes. *Q. J. R. Meteorol. Soc.* **113**: 603–633.

Sigmond M, Scinocca JF, Kharin VV, Shepherd TG. 2013. Enhanced seasonal forecast skill following stratospheric sudden warmings. *Nature Geoscience* **6**(2): 98–102.

Simmons A, Uppala S, Dee D. 2007a. Update on ERA-Interim. *ECMWF Newsletter* **111**: 5.

Simmons A, Uppala S, Dee D, Kobayashi S. 2007b. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter* **110**: 25–35.

Simpson IR, Blackburn M, Haigh JD. 2009. The role of eddies in driving the tropospheric response to stratospheric heating perturbations. *J. Atmos. Sci.* **66**(5): 1347–1365.

Simpson IR, Hitchcock P, Shepherd TG, Scinocca JF. 2011. Stratospheric variability and tropospheric annular-mode timescales. *Geophys. Res. Lett.* **38**(20).

Son SW, Gerber EP, Perlwitz J, Polvani LM, Gillett NP, Seo KH, Eyring V, Shepherd TG, Waugh D, Akiyoshi H, *et al.* 2010. Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment. *J. Geophys. Res.* **115**.

Son SW, Lee S. 2006. Preferred modes of variability and their relationship with climate change. *J. Climate* **19**(10): 2063–2075.

Son SW, Polvani LM, Waugh DW, Akiyoshi H, Garcia R, Kinnison D, Pawson S, Rozanov E, Shepherd TG, Shibata K. 2008. The impact of stratospheric ozone recovery on the Southern Hemisphere westerly jet. *Science* **320**(5882): 1486.

Song Y, Robinson WA. 2004. Dynamical mechanisms for stratospheric influences on the troposphere. *J. Atmos. Sci.* **61**(14): 1711–1725.

"SPARC intercomparison of middle-atmosphere climatologies" report. 2002. Available from http://www.sparc-climate.org/publications/sparc-reports/.

Stenchikov G, Hamilton K, Stouffer RJ, Robock A, Ramaswamy V, Santer B, Graf HF. 2006. Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models. *J. Geophys. Res.* **111**: D07107, doi:10.1029/2005JD006286.

Stewartson K. 1978. The evolution of the critical layer of a Rossby wave. *Geophys. Astrophys. Fluid Dyn.* **9**(1): 185–200.

Taguchi M. 2010. Observed connection of the stratospheric quasi-biennial oscillation with El Niño–Southern Oscillation in radiosonde data. *J. Geophys. Res.* **115**(D18): D18 120.

Taguchi M, Hartmann DL. 2006. Increased occurrence of stratospheric sudden warmings during El Niño as simulated by WACCM. *J. Climate* **19**(3): 324–332.

Telford PJ, Braesicke P, Morgenstern O, Pyle JA. 2008. Technical Note: Description and assessment of a nudged version of the new dynamics Unified Model. *Atmos. Chem. Phys.* **8**(6): 1701–1712.

Thompson DWJ, Baldwin MP, Wallace JM. 2002. Stratospheric connections to Northern Hemisphere wintertime weather: Implications for prediction. *J. Climate* **15**: 1421–1428.

Thompson DWJ, Furtado JC, Shepherd TG. 2010. On the tropospheric response to anomalous stratospheric wave drag and radiative heating. *J. Atmos. Sci.* **63**: 2616–2629, doi:10.1175/JAS3771.1.

Thompson DWJ, Solomon S. 2002. Interpretation of recent Southern Hemisphere climate change. *Science* **296**(5569): 895.

Thompson DWJ, Wallace JM. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate* **13**(5): 1000–1016.

Tung KK. 1979. A theory of stationary long waves. Part III: Quasi-normal modes in a singular waveguide. *Monthly Weather Review* **107**(6): 751–774.

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, *et al.* 2005. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **131**(612): 2961–3012.

Wallace JM, Chang FC. 1982. Interannual variability of the wintertime polar vortex in the Northern Hemisphere middle stratosphere. *Meteorological Society of Japan, Journal* **60**: 149–155.

Wallace JM, Panetta RL, Estberg J. 1993. Representation of the equatorial stratospheric quasi-biennial oscillation in EOF phase space. *J. Atmos. Sci.* **50**: 1751–1762.

Warn T, Warn H. 1978. The evolution of a nonlinear critical level. *Studies in Applied Mathematics* **59**: 37–71.

Warner CD, McIntyre ME. 1999. Toward an ultra-simple spectral gravity wave parameterization for general circulation models. *Earth Planets and Space* **51**(7–8): 475–484.

Watson PAG, Gray LJ. 2014. How does the quasi-biennial oscillation affect the stratospheric polar vortex? *J. Atmos. Sci.* **71**(1): 391–409.

Waugh DW. 1997. Elliptical diagnostics of stratospheric polar vortices. *Q. J. R. Meteorol. Soc.* **123**: 1725–1748.

Webster S, Brown AR, Cameron DR, Jones CP. 2003. Improvements to the representation of orography in the Met Office Unified Model. *Q. J. R. Meteorol. Soc.* **129**(591, Part b): 1989–2010.

Wei K, Chen W, Huang R. 2007. Association of tropical Pacific sea surface temperatures with the stratospheric Holton-Tan oscillation in the Northern Hemisphere winter. *Geophys. Res. Lett.* **34**: L16814, doi:10.1029/2007GL030478.

Wilks DS. 2006. *Statistical methods in the atmospheric sciences*. Academic press, San Diego, 2nd edn.

Woollings T. 2008. Vertical structure of anthropogenic zonal-mean atmospheric circulation change. *Geophys. Res. Lett.* **35**(19).

Yamashita Y, Akiyoshi H, Takahashi M. 2011. Dynamical response in the northern hemisphere midlatitude and high-latitude winter to the QBO simulated by CCSR/NIES CCM. *J. Geophys. Res.* **116**: D06118, doi:10.1029/2010JD015016.

Yoden S. 1987a. Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.* **44**(13): 1723–1733.

Yoden S. 1987b. Dynamical aspects of stratospheric vacillations in a highly truncated model. *J. Atmos. Sci.* **44**(24): 3683–3695.

Yoden S, Naito Y, Pawson S. 1996. A further analysis of internal variability in a perpetual January integration of a troposphere-stratosphere-mesosphere GCM. *J. Meteor. Soc. Japan* **74**: 175–188.

Yoden S, Taguchi M, Naito Y. 2002. Numerical studies on time variations of the troposphere-stratosphere coupled system. *J. Meteor. Soc. Japan* **80**(4B): 811–830.