Variability of the polar stratosphere and its influence on surface weather and climate

William J. M. Seviour

Linacre College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

October 2014

Variability of the polar stratosphere and its influence on surface weather and climate

William J. M. Seviour

Linacre College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

October 2014

Research during the last two decades has established that variability of the winter polar stratospheric vortex can significantly influence the troposphere, affecting the likelihood of extreme weather events and the skill of long-range weather forecasts. This influence is particularly strong following the rapid breakdown of the vortex in events known as sudden stratospheric warmings (SSWs). This thesis addresses some outstanding issues in our understanding of the dynamics of stratospheric variability and its influence on the troposphere.

First, a geometrical method is developed to characterise two-dimensional polar vortex variability. This method is also able to identify types of SSW in which the vortex is displaced from the pole and those in which it is split in two; known as displaced and split vortex events. It shown to capture vortex variability at least as well as previous methods, but has the advantage of being easily applicable to climate model simulations. Such an application is desireable because of the relative lack of SSWs in the observational record.

This method is subsequently applied to 13 stratosphere-resolving climate models. Almost all models show split vortex events as barotropic and displaced vortex events as baroclinic; a difference also seen in observational reanalysis data. This supports the idea that split vortex events are caused by a resonant excitation of the barotropic mode. Models show consistent differences in the surface response to split and displaced vortex events which do not project stongly onto the annular mode. However, these differences are approximately co-located with lower stratospheric anomalies. This suggests that a local adjustment to stratospheric potential vorticity anomalies is the mechanism behind the different responses to split and displaced vortex events.

Finally, the predictability of the polar stratosphere and its influence on the troposphere is assessed in a stratosphere-resolving seasonal forecast system. Little skill is found in the prediction of the strength of the Northern Hemisphere vortex at lead times beyond one month. However, much greater skill is found for the Southern Hemisphere vortex during austral spring. This allows for forecasts of interannual changes in ozone depletion to be inferred at lead times much beyond previous forecasts. It is further demonstrated that this stratospheric skill descends with time and leads to an enhanced surface skill at lead times of more than three months.

Contents

List of acronyms and abbreviations

CCMVal-2 Second Chemistry–Climate Validation Activity
CFC Chlorofluorocarbon
CMIP5 Coupled Model Intercomparison Project Phase 5
CP07 *Charlton and Polvani* [2007]
ECMWF European Centre for Medium-Range Weather Forecasts
ENSO El Niño–Southern Oscillation
EOF Empirical orthogonal function
EP Eliassen-Palm
ERA Combination of ERA-40 reanalysis (1958–1978) and ERA-Interim (1979–2010)
GCM General circulation model
GEV Generalised extreme value
GloSea5 Met Office Global Seasonal Forecast System 5
LOOCV Leave-one-out cross-validation
M13 *Mitchell et al.* [2013]
MMM Multi-model mean
MSLP Mean sea-level pressure
NAM Northern annular mode
NAO North Atlantic oscillation
PNA Pacific-North American (pattern)
PSC Polar stratospheric cloud
PV Potential vorticity
QBO Quasi-biennial oscillation
QG Quasi-geostrophic
ROC Reveiver operating characteristic
SAM Southern annular mode
SH Southern Hemisphere
SOI Southern oscillation index
SSW Sudden stratospheric warming
TEM Transformed Eulerian-mean
UTLS Upper-troposphere/lower-stratosphere

CHAPTER 1

Introduction

1.1 Overview and aims

Traditionally the stratosphere was thought to respond passively to tropospheric forcing from below. However, modelling and observational evidence gathered over the last two decades has demonstrated that variability of the winter polar stratosphere can cause significant circulation anomalies at the Earth's surface. This influence has been shown to be particularly strong following the rapid breakdown of the usual westerly winter stratospheric polar vortex; events known as sudden stratospheric warmings (SSWs).

Despite these advances, important issues remain as to the dynamics of SSWs and the stratosphere's influence on the troposphere. Most significantly:

i. The dynamics of SSWs are not fully understood, in particular whether different mechanisms may be responsible for different types of event.

ii. A mechanism for the stratosphere's influence on the troposphere is not well developed and it is not understood why some stratospheric events have different impacts on the troposphere than others.

These are significant long-standing issues, and providing a comprehensive solution is not possible here, but it is hoped that this thesis will go some way to addressing them.

A solution to these issues is not purely of theoretical interest since it is necessary to understand the dynamics of these phenomena in order to represent them realistically in weather and climate prediction models. Indeed, model biases in atmospheric dynamics have been shown to be a major source of uncertainty in seasonal forecasts [*Smith et al.*, 2012] and regional climate change projections [*Shepherd*, 2014]. With an eye on this application, this thesis also aims to assess the representation of the stratosphere and its connection with the troposphere in climate and seasonal forecast models.

The main original contributions of this thesis to the scientific literature are summarised below:

i. In Chapter 3 a new method to diagnose stratospheric polar vortex variability and classify split and displaced vortex events is introduced and tested. This is the first semi-Lagrangian (or vortex-centric) method that can be easily and robustly applied to climate model simulations. Reanalysis data are then used to compare anomalies at the tropopause and the surface following the split and displaced vortex events. Although there may be some significant differences, the relatively short observational record hinders the statistical significance of these results.

ii. In Chapter 4 this method is applied to carry out the first multi-model comparison of split and displaced vortex events and their influence on the troposphere. It is found that there is a wide range of biases in the representation of the stratospheric polar vortex among models, at least some of which may be attributable to differences in vertical resolution. It is also shown that there are consistent differences between the tropospheric response to split and displaced vortex events among the models. These differences, and the large number of events studied, allows some inference of the mechanisms behind the different surface responses.

iii. In Chapter 5 the predictability of the stratospheric polar vortex is assessed in a stratosphere-resolving seasonal prediction system. Little skill is found in the prediction of the strength of the Northern Hemisphere stratospheric polar vortex or the occurrence of split or displaced vortex events on seasonal timescales. However, significant skill is found in the case of the Southern Hemisphere vortex. This enables the skillful prediction of interannual variability in ozone depletion beyond the lead time of previous forecasts. Furthermore, it is demonstrated that the stratospheric skill significantly enhances the skill of tropospheric forecasts several months ahead.

1.2 Relation to published work

Chapter 3 is largely based on a paper written by myself, Daniel Mitchell and Lesley Gray published in *Geophysical Research Letters* [*Seviour et al.*, 2013], although the analysis has been significantly extended and re-written. The work in Chapter 5 was undertaken as part of a CASE studentship with the UK Met Office, and the part which relates to the Southern Hemisphere is based on a paper written by myself, Steven Hardiman, Lesley Gray, Neal Butchart, Craig MacLachlan, and Adam Scaife published in *Journal of Climate* [*Seviour et al.*, 2014]. Additionally, a paper based on Chapter 4 is in preparation and is expected to be submitted in the near future.

In the above papers, all the writing is my own and I carried out all the analysis and produced the figures. However, I am of course very grateful for the constructive comments of my coauthors in the preparation of these papers as well as my reviewers; Harry Hendon (from the Centre for Australian Weather and Climate Research), and three of whom are anonymous.

1.3 Thesis structure

The next chapter introduces the necessary background of the current understanding of the dynamics of the polar stratosphere, including the differences between the Northern and Southern Hemispheres and SSWs. It also reviews the role of dynamics in polar stratospheric ozone depletion, the atmospheric annular modes, and the observational, modelling, and theoretical evidence for the stratosphere's influence on the troposphere. The original results described above are presented in Chapters 3, 4, and 5. Conclusions and possible extensions to the work in this thesis are discussed in Chapter 6.

CHAPTER 2

Background

The stratosphere is the layer of the Earth's atmosphere that lies above the troposphere and is bounded by the tropopause below and the stratopause above. The height of the tropopause varies from about 15 km in altitude in the tropics to 7 km at high latitudes, while the stratopause lies at approximately 50 km. The defining feature of the stratosphere is a temperature gradient increasing with height (in contrast to the troposphere below), caused by the presence of ozone which absorbs ultraviolet radiation[1] and thereby heats the surrounding atmosphere. This temperature gradient makes the stratosphere stable against vertical convection and results in very different dynamical behaviour to the troposphere. In this chapter, our current understanding of the dynamics of the polar stratosphere are reviewed (Section 2.1), as well as the relationship between dynamics and polar stratospheric ozone depletion (Section 2.2), and stratosphere-troposphere coupling (Sections 2.3 and 2.4).

[1]For this reason, the region of the stratosphere with the highest ozone concentrations is often called the "ozone layer".

2.1 Dynamics of the polar stratosphere

2.1.1 Zonal-mean circulation

Each winter the polar region descends into a polar night and the stratosphere cools by infrared radiation to space. This sets up a strong equator-to-pole temperature gradient which increases the vertical zonal wind shear in accordance with the thermal wind balance relation

$$\frac{\partial u_g}{\partial z} = -\frac{R}{fH}\frac{\partial T}{\partial y}, \tag{2.1}$$

where u_g is the geostrophic zonal velocity,[2] $u_g = -f^{-1}\partial Z/\partial y$, Z is geopotential height, f is the Coriolis parameter, $f = 2\Omega \sin\phi$, and R is the specific gas constant. Here, a beta-plane geometry is used such that $f = f_0 + \beta y$, where $f_0 = f(\phi_0)$ and $\beta = 2\Omega a^{-1} \cos\phi_0$, a is the Earth's radius and ϕ_0 is a reference latitude. H is the scale height given by $H = RT_s/g$, where T_s is a reference temperature and g is the acceleration due to gravity. This equation relies on hydrostatic and geostrophic approximations but is approximately satisfied on seasonal timescales. Hence, the meridional temperature gradient results in a region of westerly winds in the winter hemisphere, surrounding the pole; this is known as the *stratospheric polar vortex*.[3]

Figure 2.1 shows zonal-mean zonal wind and temperature averaged over the boreal winter (December–February; DJF) and austral winter (July–August; JJA) using data from 1979–2010 from the ERA-Interim reanalysis (details in Section 3.3.1). In both cases the westerly vortex in the winter hemisphere can be seen along with a local minimum in temperature at the winter pole in the lower stratosphere. Easterly winds are present in the summer hemisphere. The maximum strength of the polar vortex occurs at midlatitudes between 0.1–1 hPa in the mesosphere, and is stronger in the Southern Hemisphere (SH) with a maximum of 90 ms^{-1} than the Northern Hemisphere (NH) with a maximum of 50 ms^{-1}. The winter polar stratosphere is also approximately

[2]That is, the zonal wind balanced by pressure gradient forces.

[3]Alternative names for this that often appear in the literature are *polar night jet* or *polar night vortex*. These names are not used here because the vortex persists outside of the polar night.

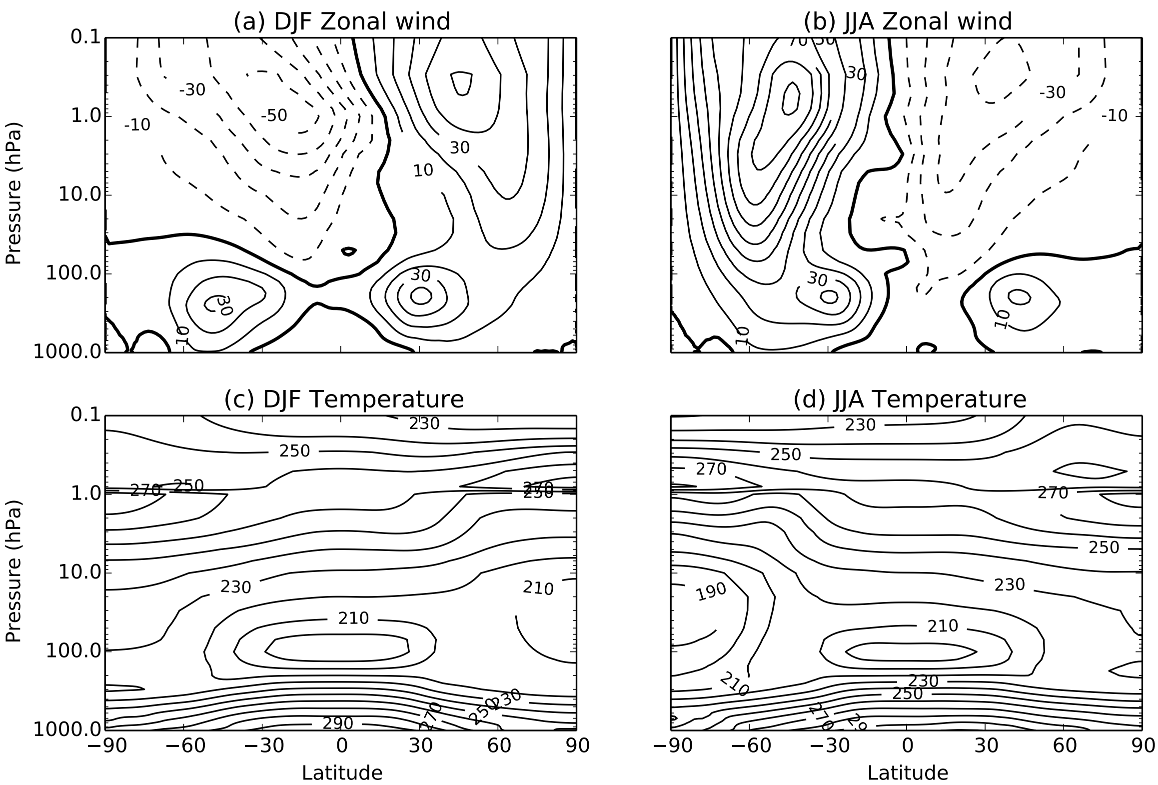

FIGURE 2.1: December-January (DJF) (a,c) and July-August (JJA) (b,d) averages of zonal-mean zonal wind (m s^{-1}) (a,b) and temperature (K) (c,d). Dashed contours represent negative values. Data is from the ERA-Interim reanalysis (1979–2010).

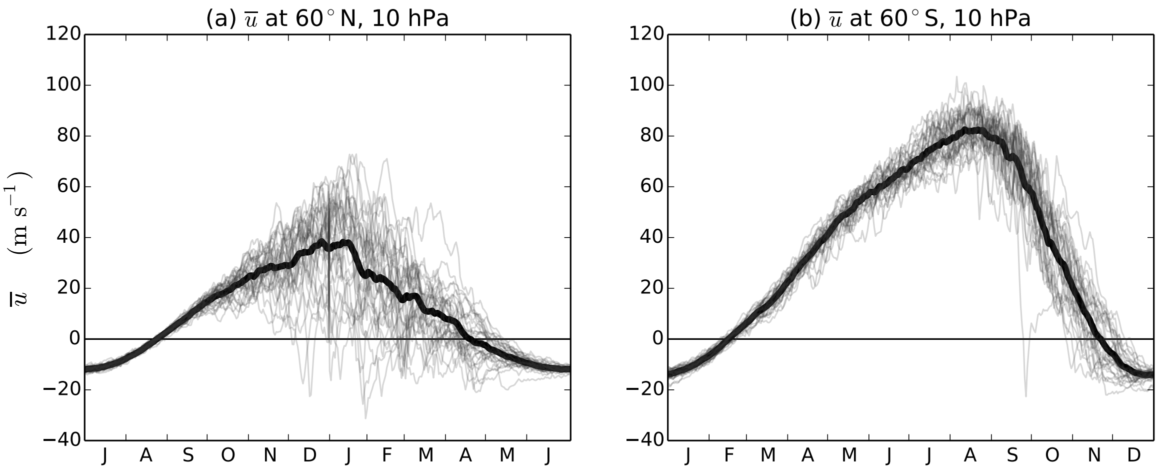

FIGURE 2.2: Seasonal cycle of NH (a) and SH (b) polar vortex strength, measured by $\overline{u}$ at 60°N/S, 10 hPa. The annual mean is shown in a thick black line and individual years in thin grey lines. Both time series are centred on their respective winters. Data is from the ERA-Interim reanalysis (1979–2010).

20 K colder in the SH than the NH.

The maximum strength of the vortex in the stratosphere occurs at approximately 60°N/S with little variation through the depth of the stratosphere. Figure 2.2 shows the annual cycle and variability of zonal-mean zonal wind at 10 hPa 60°N and 60°S. As well as being weaker on average than the SH, the winter NH stratospheric polar vortex can also be seen to be significantly more variable than the SH. There are a number of years in the NH for which $\overline{u}$ becomes negative during the winter, but only one such year in the SH (these events are discussed further in Section 2.1.3). A further clear feature of both NH and SH is that variability during the summer is much less than that during winter. Also, the transition to summer easterlies (known as the final warming) occurs relatively earlier in the seasonal cycle in the NH than the SH. All these observations are due almost entirely to the influence of wave phenomena in the stratosphere, as described in the next section.

2.1.2 Waves in the stratosphere

2.1.2.1 Planetary waves

Large-scale Rossby or planetary waves[4] play a vital role in the dynamics of the extratropical stratosphere. They mostly enter the stratosphere from the troposphere, where they are forced, for example, by air flow around topography, latent heat release, or nonlinear evolution of tropospheric eddies [*Scinocca and Haynes*, 1998]. These large-scale waves approximately satisfy the quasi-geostrophic (QG) approximation of hydrostatically balanced incompressible flow with low Rossby number, $\mathrm{Ro} = U/f_o L \ll 1$, where U and L are characteristic velocity and length scales respectively [*Andrews et al.*, 1987]. Under this approximation and in the absence of friction, the following relation, known as the *quasi-geostrophic potential vorticity equation*, holds:

$$D_g q_g = f_0 \rho_0 \frac{\partial}{\partial z} \frac{\rho_0 Q}{\partial \theta_0 / \partial z} . \tag{2.2}$$

Where

$$D_g \equiv \frac{\partial}{\partial t} + u_g \frac{\partial}{\partial x} + v_g \frac{\partial}{\partial y} , \tag{2.3}$$

and

$$q_g = f_0 + \beta y - \frac{\partial v_g}{\partial x} + \frac{\partial u_g}{\partial y} + \rho_o^{-1} \frac{\partial}{\partial z} \left(\rho_o f_0 \frac{\theta_e}{\partial \theta_0 / \partial z} \right) , \tag{2.4}$$

is the quasi-geostrophic potential vorticity. Here, v_g is the geostrophic meridional velocity, $v_g = f^{-1} \partial Z / \partial x$, Q is the diabatic heating rate, ρ_0 is a reference density and θ_0 is a reference potential temperature, $\theta_0 = T_s(p_s/p)^\kappa$, where $p_s = 1000$ hPa, and $\kappa = R/c_p \approx 2/7$, where c_p is the specific heat capacity of air at constant pressure. θ_e represents the departure from θ_0, and is assumed to be small in the sense that $|\partial \theta_e / \partial z| \ll |\partial \theta_0 / \partial z|$. An important consequence of Equation 2.2 is that q_g is conserved following the geostrophic wind for adiabatic flow ($Q = 0$), and therefore acts as a

[4]Here, as is common in the stratospheric literature, "wave" is taken to mean any deviation from the zonal-mean state.

tracer[5].

In the case of approximately zonal flow $[\overline{u}(y,z),0,0]$, Equation 2.2 can be linearised to give

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)q_g'+v'\frac{\partial \overline{q_g}}{\partial y}=f_0\rho_0\frac{\partial}{\partial z}\frac{\rho_0 Q'}{\partial \theta_0/\partial z}\,, \tag{2.5}$$

where primes represent deviations from the zonal mean (e.g., $q_g=\overline{q_g}+q_g'$). It can be shown that Equation 2.5 supports wave-like solutions, with vertical propagation dependent upon the condition:

$$0<\overline{u}-c<\overline{u}_c\equiv\beta(k^2+l^2+\varepsilon/4H^2)^{-1}\,, \tag{2.6}$$

which is known as the *Charney-Drazin criterion* after *Charney and Drazin* [1961]. Here, c is the wave's zonal phase speed, k and l are the zonal and meridional wavenumbers respectively, and $\varepsilon=f_0^2/N^2$, where $N^2=H^{-1}Re^{-\kappa z/H}\partial\theta_0/\partial z$ is the static stability. In the case of waves whose phase is stationary with respect to the ground ($c=0$), this simplifies to

$$0<\overline{u}<\overline{u}_c\,. \tag{2.7}$$

It is therefore apparent that in order for planetary waves to propagate vertically (such as from the troposphere to the stratosphere), a westerly flow must be present that is not too strong. Additionally, this maximum speed is dependent on wavenumber, such that a lower wavenumber can propagate in a stronger westerly flow. While the assumptions here are not representative of the real atmosphere (such as purely zonal flow, and small deviations from the zonal mean), this criterion does capture the most important features

[5]Throughout most of this thesis, Ertel's potential vorticity, q, is used. This is defined by

$$q=\frac{1}{\rho}\zeta\cdot\nabla\theta\,,$$

where ζ is the absolute vorticity. *Charney and Stern* [1962] showed that when the quasi-geostrophic approximation is valid

$$\left(\frac{\partial q}{\partial s}\right)_{\theta=\text{const.}}\approx\frac{1}{\rho_0}\frac{\partial\theta_0}{\partial z}\left(\frac{\partial q_g}{\partial s}\right)_{z=\text{const.}}\,,$$

where $s=t$, x or y. Hence, a similar conservation law as for q_g applies to q, which is conserved on isentropic (e.g., constant θ) surfaces.

of the relation between zonal assymmetries and the zonal flow, and similar relations can be found for more complex background states [*Andrews et al.*, 1987].

An important consequence of the Charney-Drazin criterion for stratospheric flow is that the strength of the stratospheric polar vortices shown in Figures 2.1 and 2.2 is often sufficient to exclude all but the lowest wavenumbers (typically zonal wavenumbers 1–3; hereafter referred to as 'wave-*n*') from propagating upwards from the troposphere. Hence the length-scale of typical stratospheric zonal asymmetries is much larger than that of the troposphere.

When planetary waves reach a *critical surface*, where propagation is prohibited (for instance, a region where $\overline{u} = c$), the above linear analysis breaks down. In this scenario waves can "break", imparting momentum onto the zonal flow. There is therefore a two-way interaction between the zonal flow and planetary waves; a phenomenon known as *wave-mean flow interaction*. Wave breaking was studied in an idealised two-dimensional model by *Stewartson* [1978] and *Warn and Warn* [1978]. They found momentum to be absorbed in a narrow *critical layer* close to the critical surface, with potential vorticity (PV) contours being irreversibly stretched and mixed in increasingly fine scales; a process known as a 'potential enstrophy cascade' [e.g., *Rhines and Holland*, 1979]. They also showed that the critical layer is initially absorbing, but becomes a reflecting surface after some time. Time varying results from a version of the Stewartson-Warn-Warn model from *Andrews et al.* [1987] are shown in Figure 2.3. Similar wave breaking behaviour was first observed in the real stratosphere by *McIntyre and Palmer* [1983] using isentropic maps of Ertel's potential vorticity, including the irreversible deformation of PV contours of the kind shown in Figure 2.3.

A further effect of wave breaking is the induction of a *residual circulation*, $[0, \overline{v}^*, \overline{w}^*]$, where $\overline{v}^*$ and $\overline{w}^*$ are the transformed Eulerian-mean (TEM) meridional and vertical

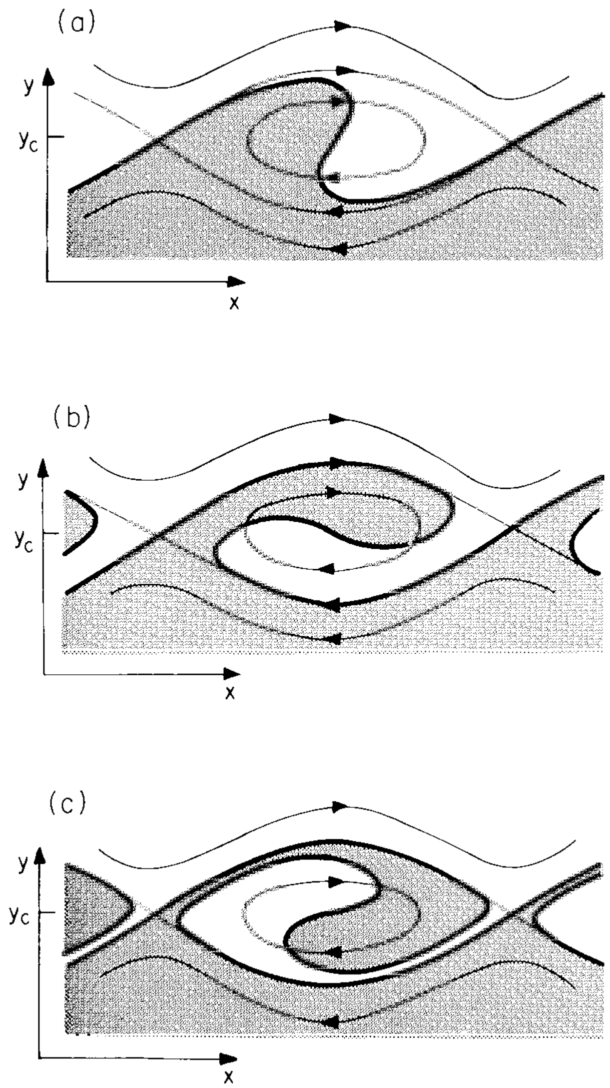

FIGURE 2.3: Stewartson-Warn-Warn time-dependent analytical solution of a Rossby wave non-linear critical layer (time advancing in even steps from (a)-(b)-(c)). The flow is periodic in x and the y scale is greatly exaggerated and the initial critical line was at $y = y_c$. The thin lines indicate streamlines and lens shaped regions of closed streamlines are known as "Kelvin's cats' eyes". The thick line shows the position of the absolute vorticity contour $\zeta = \zeta_c$, that initially lay along $y = y_c$ (in this barotropic model, the quasi-geostrophic potential vorticity, q_g, reduces to ζ). Hence, $\zeta < \zeta_c$ in the stippled region and $\zeta > \zeta_c$ in the unstippled region. In (a) it can be seen that $v > 0$ for most of the stippled region, indicating partial absorption, whereas in (b) and (c) $v \leq 0$ in the stippled region indicating reflection. Figure from *Andrews et al.* [1987].

velocities given in spherical coordinates by

$$\overline{v}^* = \overline{v} - \frac{1}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_o \overline{v'\theta'}}{\partial\overline{\theta}/\partial z}\right), \tag{2.8}$$

$$\overline{w}^* = \overline{w} + \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\frac{\cos\phi\,\overline{v'\theta'}}{\partial\overline{\theta}/\partial z}\right), \tag{2.9}$$

which approximates the Lagrangian-mean circulation under time-averaged conditions [*Andrews and McIntyre*, 1976; *Dunkerton*, 1978; *Holton*, 1990]. Under the TEM formalism, the zonal momentum equation becomes

$$\frac{\partial\overline{u}}{\partial t} + \overline{v}^*\left(\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}(\overline{u}\cos\phi) - f\right) + \overline{w}^*\frac{\partial\overline{u}}{\partial z} = \frac{\nabla\cdot\mathbf{F}}{\rho_o a\cos\phi} + \overline{X} \equiv \overline{\mathscr{F}} \tag{2.10}$$

where $\overline{X}$ represents frictional terms and $\mathbf{F} = [0, F^\phi, F^z]$ is the Eliassen-Palm (EP) flux with components

$$F^\phi = \rho_0 a\cos\phi\left(\frac{\partial\overline{u}}{\partial z}\frac{\overline{v'\theta'}}{\partial\overline{\theta}/\partial z} - \overline{v'u'}\right), \tag{2.11}$$

$$F^z = \rho_0 a\cos\phi\left(\left[f - \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}(\overline{u}\cos\phi)\right]\frac{\overline{v'\theta'}}{\partial\overline{\theta}/\partial z} - \overline{w'u'}\right). \tag{2.12}$$

F can be interpreted as the flux of wave activity [*Andrews et al.*, 1987], and therefore $\nabla\cdot\mathbf{F} < 0$ (convergence) represents a dissipation of wave activity, as is the case in wave breaking. It can be seen that for a steady zonal flow ($\partial\overline{u}/\partial t = 0$) in the absence of wave driving ($\nabla\cdot\mathbf{F} = 0$) or friction ($\overline{X} = 0$), a solution of Equation 2.10 is $\overline{v}^* = 0, \overline{w}^* = 0$. However, in the presence of these forcing terms, a non-zero residual circulation will be induced. Climatologically, this circulation consists of wave-driven poleward and downward motion in the extratropics which is balanced by upwelling in the tropics; a pattern which forms a significant part of the Brewer-Dobson circulation[6]. The downward motion near the poles can be easily seen from Equation 2.10 in the steady case; since $\overline{v}^*$ must become small near the poles (by conservation of mass), and $\partial\overline{u}/\partial z > 0$ in the

[6]Strictly, the Brewer-Dobson circulation represents the meridional transport of tracers, and so also involves two-way mixing (i.e. transport without net transfer of mass) [*Hall and Plumb*, 1994].

polar vortex, if $\nabla \cdot \mathbf{F} < 0$, then $\overline{w}^* < 0$. It is observed that this circulation is strongest in the winter hemisphere due to the fact that more planetary waves can propagate and break in the winter westerly flow than the summer easterly flow (due to the Charney-Drazin criterion). Furthermore, during periods of enhanced wave breaking the residual circulation accelerates and there is more descent and adiabatic heating at high latitudes. This is important in the physical understanding of sudden stratospheric warming events, described in Section 2.1.3.

2.1.2.2 Gravity waves

Gravity waves are another type of atmospheric wave important in the dynamics of the polar stratosphere. These waves owe their existence to buoyancy restoring forces and can be generated by a number of processes such as air flow over topography (orographic gravity waves), convection or frontogenesis (non-orographic gravity waves). As with planetary waves, the differences in the land masses of the two hemispheres leads to orographic gravity wave activity being much greater in the NH. These waves make a net easterly contribution to the winter zonal flow [e.g., *Seviour et al.*, 2012], and so act to enhance the residual circulation. Their typical length scales are much shorter than can be resolved in general circulation models or reanalyses, and so they are usually parametrised, appearing as the term $\overline{X}$ in the zonal momentum equation (Equation 2.10).

Together, the Charney-Drazin criterion and the effects of planetary and gravity wave driving on the zonal flow can expain almost all hemispheric differences seen in Figures 2.1 and 2.2: Greater topography results in more planetary and gravity wave generation in the NH, both of which cause a net deceleration of the westerly polar vortex, thereby causing the NH vortex to be weaker than the SH. This also explains why the NH vortex is warmer than the SH, as the greater NH wave activity induces a stronger residual circulation with enhanced descent and adiabatic warming at high latitudes. Additionally, the strength of the SH vortex is such that it prohibits the vertical propagation of planetary waves from the troposphere throughout much of the winter, meaning that the SH vortex is less variable than the NH. Both hemispheres show very little variability in the summer easterly flow because planetary wave propagation is prohibited in this regime.

2.1.3 Sudden stratospheric warmings

First observed by *Scherhag* [1952] in radiosonde measurements over Berlin, the extreme events visible in Figure 2.2 whereby the winter circulation temporarily becomes easterly[7] are known as sudden stratospheric warmings[8] (SSWs). These events occur approximately 5–7 times per decade in the NH, but only one such event has been observed in the almost 60 year observational record in the SH (in 2002). They are called "warmings" because associated with the circulation reversal is a dramatic increase in temperature; as much as 50 K in the space of a few days in the mid-stratosphere.[9]

Initially these events were thought to result from either solar storms [*Scherhag*, 1952] or baroclinic instability of the stratospheric polar vortex [*Murray*, 1960]. However, *Matsuno* [1970, 1971] proposed a model of SSWs which relies on the influence of tropospherically forced planetary waves. This model (or modifications thereof) remains the most widely accepted dynamical view of SSWs at present. The mechanism proceeds as follows:

i. A packet of enhanced planetary wave activity enters the stratosphere where it reaches a critical surface and breaks. This decelerates the zonal flow over a broader

[7]For a discussion of more precise definitions of SSWs, see Section 3.1.

[8]Following *Butler et al.* [2014] it is suggested that the term *sudden stratospheric warming* is preferable to the common alternative *stratospheric sudden warming*. This is because there are other varieties of stratospheric warming (such as final warmings or Canadian warmings), but not other varieties of atmospheric sudden warming.

[9]A recent study [*Neef et al.*, 2014] has even suggested that SSWs have a sufficiently strong influence on the Earth's angular momentum that they have a detectable influence on the length of the day.

critical layer, and if strong enough causes it to reverse.

ii. Hence a new critical surface is formed at a lower level (where $\overline{u} = 0$), and wave breaking occurs at this level. The process continues as the critical layer descends to the lower stratosphere.

iii. At the same time, wave breaking induces an enhanced residual circulation with greater descent and adiabatic warming at high latitudes. If strong enough, this can act to reverse the meridional temperature gradient, further enhancing the easterly flow by thermal wind balance.

iv. When the critical surface is close to the tropopause, planetary wave activity is essentially prohibited from entering the polar stratosphere. Radiative cooling to space then acts to cool the polar stratosphere and the vortex reforms over a period of approximately 2–4 weeks.

This mechanism considers the effect of planetary waves on the zonal-mean flow. However, it has been observed that SSWs generally occur as either a split or displacement of the vortex, mostly depending (though not exclusively; see Section 3.1, [*Waugh*, 1997]) on whether wave-2 or wave-1 activity is dominant. An example of each of these events is shown in Figure 2.4. *Charlton and Polvani* [2007] and *Matthewman et al.* [2009] studied the dynamics of these two types of events in reanalysis data and noted some differences. Most significantly, split vortex events were observed to occur near-barotropically, with two smaller vortices centred over Canada and Siberia throughout the depth of the stratosphere. On the other hand, displaced vortex events were observed to be more baroclinic, starting first in the upper stratosphere with a vortex centred over Canada, the centre of which rotates westward with height and is centred over Siberia in the lower stratosphere.

This different behaviour of split and displaced vortex events is not accounted for by the *Matsuno* [1970, 1971] model above, and so may suggest that other mechanisms are

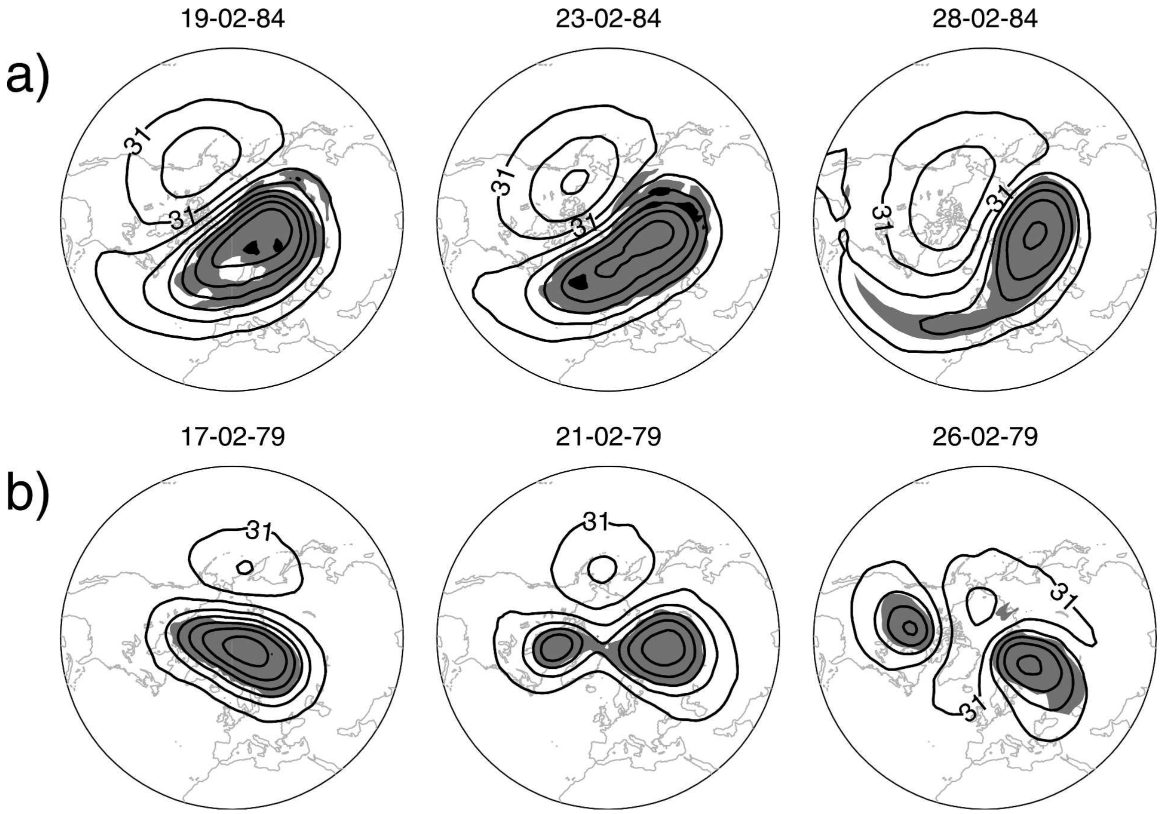

FIGURE 2.4: Polar stereographic plot of geopotential height (contours) on the 10 hPa pressure surface. The contour interval is 0.4 km, and shading shows potential vorticity greater than 4.0×10^{-6}K kg^{-1} m^2 s^{-1}. (a) A vortex displacement type warming that occurred in February 1984. (b) A vortex splitting type warming that occurred in February 1979. Figure from *Charlton and Polvani* [2007].

important in the generation of SSWs. For instance, *O'Neill and Pope* [1988] and *Scott and Dritschel* [2006] have suggested that SSWs can be generated by a cyclone-anticyclone interaction between the Aleutian high and the polar vortex. These studies showed that a smaller anticyclone can act to significantly distort the polar vortex, although such interactions are greatest for circulation ratios higher than are typically found in the polar stratosphere. Other studies have suggested that SSWs can arise through the resonant excitation of normal modes of the stratosphere by planetary waves [*Tung and Lindzen*, 1979]. Significantly, *Plumb* [1981] found that the planetary wave forcing need not be at exactly the resonant frequency of the mode (an occurrence which is probably unlikely), but that the two can be brought to resonance by a process known as nonlinear self-tuning. *Smith* [1989] found behaviour suggestive of this process in simulations of a SSW. More recently, *Esler and Scott* [2005], *Esler et al.* [2006], and *Matthewman and Esler* [2011] have argued that a relevant mode in the case of split vortex events is the barotropic mode of the atmosphere, which may explain the more barotropic nature of split vortex events.[10] *Albers and Birner* [2014] further suggested that gravity waves play an important role in 'tuning' the geometry of the vortex towards resonant excitation.

Further weight is given to these mechanisms which do not rely on strong transient tropospheric forcing by the occurrence of the 2002 SH SSW, since this forcing is much weaker in the SH (several studies of the dynamics of this event can be found in the March 2005 special issue of *Journal of Atmospheric Sciences*). Indeed, *Esler et al.* [2006] provided evidence that this event may have been influenced by resonant excitation of the barotropic mode.

Several studies have also discussed the role of the polar vortex being in a favourable (or 'preconditioned') state prior to SSWs [e.g., *McIntyre*, 1982]. In a simple dynamical model, *Scott et al.* [2004] demonstrated that planetary wave breaking is enhanced by the presence of steep PV gradients at the vortex edge, which are likely to be present in

[10] In an idealised modelling study, *Esler and Matthewman* [2011] also suggested that resonant excitation of the first baroclinic mode may play an important role in the occurrence of displaced vortex events.

an anomalously strong vortex. Indeed, *Limpasuvan et al.* [2004] found evidence for an anomalously strong vortex 30-40 days prior to SSWs, while *Charlton and Polvani* [2007] found this effect to be stronger prior to split vortex events than displaced vortex events.

Overall, while significant advances in understanding the dynamics of SSWs have been made, several uncertainties remain. Tropospherically-driven wave activity is certainly an important factor but the roles (if any) of cyclone-anticyclone interactions or resonance are less certain. Moreover, it is not clear whether different mechanisms may be more or less important in driving split and displaced vortex events.

2.2 Polar stratospheric ozone depletion

The Antarctic ozone hole is a large region of severely depleted ozone concentrations in the lower-mid stratosphere which occurs during the austral spring. During its formation ozone is often be completely destroyed at some altitudes. A similar, but much smaller depletion is observed in the NH [e.g., *Manney et al.*, 1997]. Following its discovery by *Farman et al.* [1985], a chemical and dynamical theory of the ozone hole was rapidly developed which largely attributes this rapid ozone depletion to the presence of anthropogenic chlorofluorocarbon (CFC) compounds [*McElroy et al.*, 1986; *Solomon et al.*, 1986]. This theory is summarised as follows:

i. The strong zonal winds of the stratospheric polar vortex act to confine air over the polar regions, with little mixing with midlatitudes [*Schoeberl and Hartmann*, 1991]. This results in a region of very cold temperatures which allow the formation of polar stratospheric clouds (PSCs; these require temperatures below approximately 195 K to form [*Newman*, 2010]).

ii. Heterogeneous chemical reactions can take place on the surface of PSCs which act to convert 'reservoir' chlorine species such as $ClONO_2$ and HCl into forms that can accelerate ozone depletion, for instance [*Solomon et al.*, 1986]:

$$\mathrm{ClONO_2 + HCl \xrightarrow{PSC} Cl_2 + HNO_3}$$

iii. As sunlight returns to the vortex region in spring, Cl_2 is rapidly photolysed and reacts with oxygen to form ClO. This can then catalytically destroy ozone through a reaction sequence such as the following suggested by *Molina and Molina* [1987]:

$$\begin{aligned}
\mathrm{ClO + ClO} &\xrightarrow{M} \mathrm{ClOOCl} \\
\mathrm{ClOOCl} + h\nu &\longrightarrow \mathrm{ClOO + Cl} \\
\mathrm{ClOO} &\xrightarrow{M} \mathrm{Cl + O_2} \\
2\,(\mathrm{Cl + O_3} &\longrightarrow \mathrm{ClO + O_2}) \\
\text{Net}:\ 2\,\mathrm{O_3} &\longrightarrow 3\,\mathrm{O_2}
\end{aligned}$$

where ν is the frequency of light, h is Planck's constant, and M a third body (necessary for conservation of momentum). Several other reaction sequences are also possible, for instance involving bromine species.

iv. A further effect of PSCs is that their particles fall out of the stratosphere and thereby remove nitrogen compounds (NO_x) from the polar lower stratosphere [*Toon et al.*, 1986]. NO_x compounds are important because they can react with with ClO to form reservior compounds. For instance the reaction

$$\mathrm{ClO + NO_2 \xrightarrow{M} ClONO_2}$$

removes ClO from the ozone-depleting sequence above. Hence a reduction of NO_x due to PSCs leads to an increase in ozone depletion.

v. After some time, radiative heating of the stratosphere is sufficient to prevent the formation on PSCs, so ozone depletion halts. This heating is further accelerated by the increased wave breaking in the polar stratosphere which can take place as

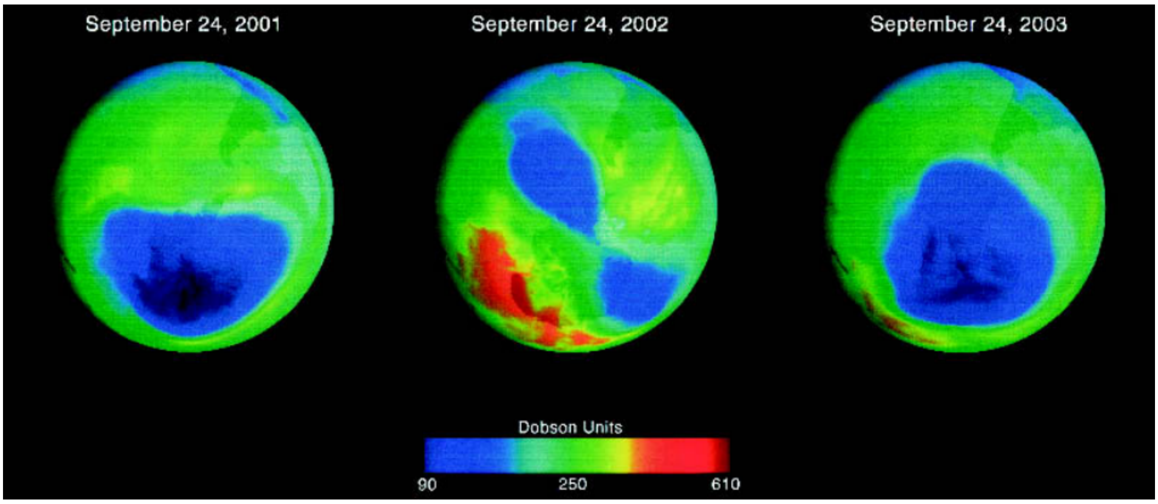

FIGURE 2.5: Comparison of the (middle) first split ozone hole on record (2002) and (left) the Antarctic ozone hole at the same time one year earlier (2001) and (right) one year later (2003). The hole is dark blue and magenta. In 2001, the ozone layer thinning over Antarctica reached 26.5×10^6 km^2 , larger than the size of the entire North American continent. Due to higher Antarctic winter temperatures, the 2002 "hole" seems to be about 40% smaller. In 2003, Antarctic winter temperatures returned to normal and the ozone hole returned to its usual state. Figure from *Shepherd et al.* [2005].

the vortex weakens (due to the Charney-Drazin criterion) and thereby induce an enhanced residual circulation. Following the final breakdown of the vortex and transition to summer easterlies (final warming), the ozone-depleted air is mixed to lower latitudes.

Importantly, the stratospheric dynamics described in Section 2.1, play an important role in ozone depletion. As discussed in Section 2.1.2.1, wave breaking in the stratosphere acts to drive a residual circulation, with descent and adiabatic warming over the pole. Enhanced descent and warming over the pole acts to inhibit the formation of PSCs which are necessary for the above heterogeneous chemical reactions which cause ozone depletion. A stronger meridional circulation also acts to transport more tropical ozone-rich air to the polar regions, further acting to increase ozone concentrations. Another mechanism in which wave breaking acts to inhibit ozone depletion is by actively stripping away filaments of ozone-depleted air from the polar vortex [*Waugh et al.*, 1994]. It is this greater wave activity in the NH then that explains why the extent of ozone depletion is much less in the NH than the SH.

In the extreme event of SSWs, the ozone hole can be severely disrupted. This can be seen in Figure 2.5, where a clear split of the ozone hole is visible during the 2002 SH SSW, which contrasts with the more zonally symmetric distributions seen at the same times in 2001 and 2003. The magnitude of the 2002 ozone hole can also be seen to be reduced; a result of an enhanced residual circulation causing warmer stratospheric temperatures. The importance of this link between dynamics and chemistry for predicting interannual variability in ozone depletion is discussed further in Sections 5.4.2 and 5.5.

2.3 Annular modes

Before discussing the influence of the stratosphere on the troposphere, it is necessary to introduce the concept of the annular modes which are often analysed in studies on this topic (as well as in this thesis). The northern and southern annular modes (NAM and SAM) are the leading modes of large-scale variability in the two hemispheres [*Baldwin and Dunkerton*, 1999; *Thompson and Wallace*, 2000; *Thompson et al.*, 2000; *Limpasuvan and Hartmann*, 1999, 2000]. They are commonly defined to be the leading empirical orthogonal function (EOF)[11] of extratropical monthly-mean geopotential height calculated at each pressure surface [e.g., *Baldwin and Dunkerton*, 1999], with an index of the respective principal component or the projection of daily data onto the EOF pattern. *Baldwin and Thompson* [2009] introduced an alternative using zonal-mean geopotential height, which is therefore less computationally expensive to calculate. The annular modes at the surface are also often calculated from mean sea-level pressure [e.g., *Gong and Wang*, 1999], where they may also be referred to as the Arctic and Antarctic oscillation (although in this thesis, the terms *surface NAM/SAM* are used).

Figure 2.6 shows linear regressions of zonal-mean geostrophic winds and lower

[11]EOFs are the eigenvectors of the spatially weighted covariance matrix of a variable. EOF analysis is also known as principal component analysis [*Wilks*, 2006].

tropospheric geopotential height on the NAM and SAM indices as defined by *Thompson and Wallace* [2000]. It can be seen that the near-surface NAM structure is more zonally asymmetric than the SAM, with centres of action located over the Atlantic and Pacific oceans, and the SAM has a more clearly 'annular' structure. The vertical structure of the NAM and SAM appears near barotropic, although with a slight poleward tilt with height.[12] The magnitude of the zonal wind signature also increases with height, reaching a maximum in the upper troposphere/lower stratosphere in the SH, and the mid-stratosphere in the NH.

Despite this near-barotropic appearance, annular mode variability has quite different physical interpretations in the troposphere and stratosphere. Stratospheric variability is associated with approximately zonally symmetric strengthening and weakening of the stratospheric polar vortex, while tropospheric variability is more closely associated with meridional shifts in the eddy-driven jets [*Limpasuvan and Hartmann*, 1999]. Furthermore, stratospheric annular mode variability is largely confined to winter, while tropospheric variability has a much less pronounced seasonal cycle.

It can be seen from Figure 2.6(d) that the Atlantic centre of action of the NAM resembles the familiar North Atlantic Oscillation (NAO) pattern. Indeed, the surface NAM and NAO have been observed to be highly correlated [*Ambaum et al.*, 2001]. This has led to some debate as to whether the NAM represents a physical mode of variability of which the NAO is just a regional manifestation, or whether the NAM is simply a statistical artifact of more regional variability. This point is addressed in Section 4.4.2.

2.4 Stratosphere-troposphere coupling

So far this chapter has dealt with the dynamics of the stratosphere as responding passively to tropospheric forcing from below. Indeed, this was the dominant view until the last two decades [e.g., *Andrews et al.*, 1987]. However, observational evidence

[12] *Thompson and Woodworth* [2014] have recently argued that the barotropic and baroclinic aspects of the SAM can be viewed as two independent modes of variability.

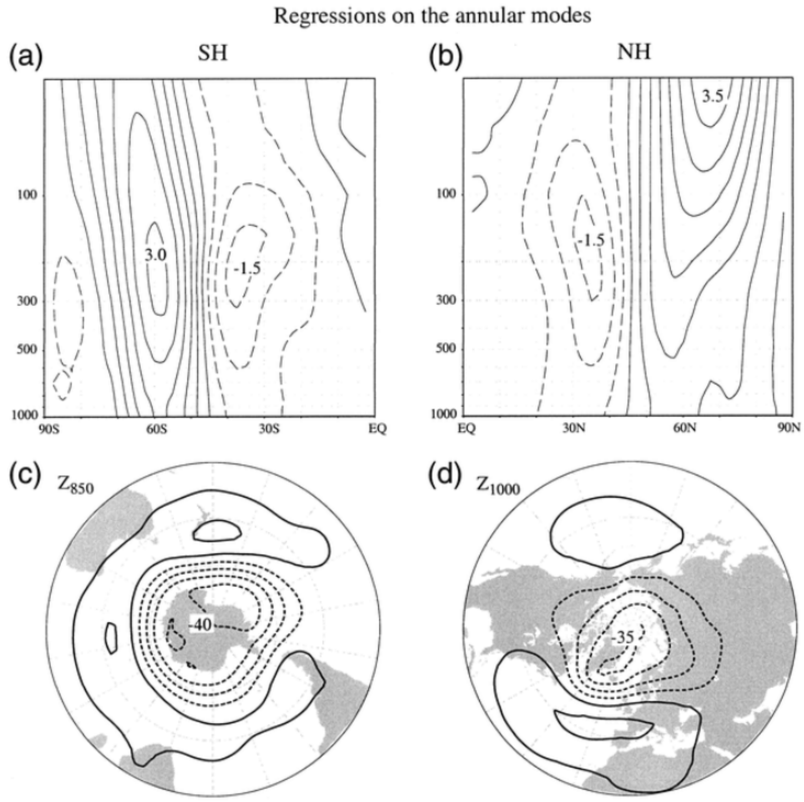

FIGURE 2.6: (top) Zonal-mean geostrophic wind and (bottom) lower-tropospheric geopotential height regressed on the standardised indices of the annular modes (the AO and its SH counterpart) based upon monthly data, Jan 1958–Dec 1997. Left panels are for the SH, right panels are for the NH. Units are m s^{-1} (top) and m per std dev of the respective index time series (bottom). Contour intervals are 10 m $(-15, -5, 5, \ldots)$ for geopotential height and 0.5 m s^{-1} $(-0.75, -0.25, 0.25)$ for zonal wind. Figure from *Thompson and Wallace* [2000].

supported by modelling studies and some theoretical arguments have now provided evidence that variability of the polar stratosphere can significantly influence tropospheric weather and climate. This evidence is discussed below.

2.4.1 Observational evidence

The accumulation of observational evidence for a two-way dynamical link between the polar stratosphere and the troposphere began in the early 1990s. *Kodera et al.* [1990] found that the strength of the NH upper stratospheric polar vortex during December was correlated with the strength of the tropospheric eddy-driven jet in February. However, they did not investigate the mechanism for this relation, and suggested it may be radiative. Further evidence was provided by *Nigam* [1990] who found barotropic and baroclinic modes of the zonal-mean zonal wind which vary coherently in the troposphere and stratosphere. This analysis was extended by *Baldwin et al.* [1994] who found significant correlations between stratospheric and tropospheric EOFs of daily NH geopotential height (patterns which would now be referred to as annular modes). However, they found that the strongest correlations occurred with the troposphere leading, and while suggesting that the stratosphere may exert some influence on the troposphere, they concluded that the direction of causality was mostly upwards.

These studies were followed by several others which looked at the co-variability of the stratospheric and tropospheric annular modes, and demonstrated links between the strength of the polar vortex and surface temperature and sea level pressure patterns [*Perlwitz and Graf*, 1995; *Thompson and Wallace*, 1998; *Baldwin and Dunkerton*, 1999]. However, it was *Baldwin and Dunkerton* [2001] who first demonstrated that the stratosphere-troposphere link is particularly strong following extreme weakenings or strengthenings of the stratospheric polar vortex. Figure 2.7 shows a time series of the NAM during the winter of 1998–1999, which includes two such weak vortex events (in December and February), characterised by a negative stratospheric NAM (these events

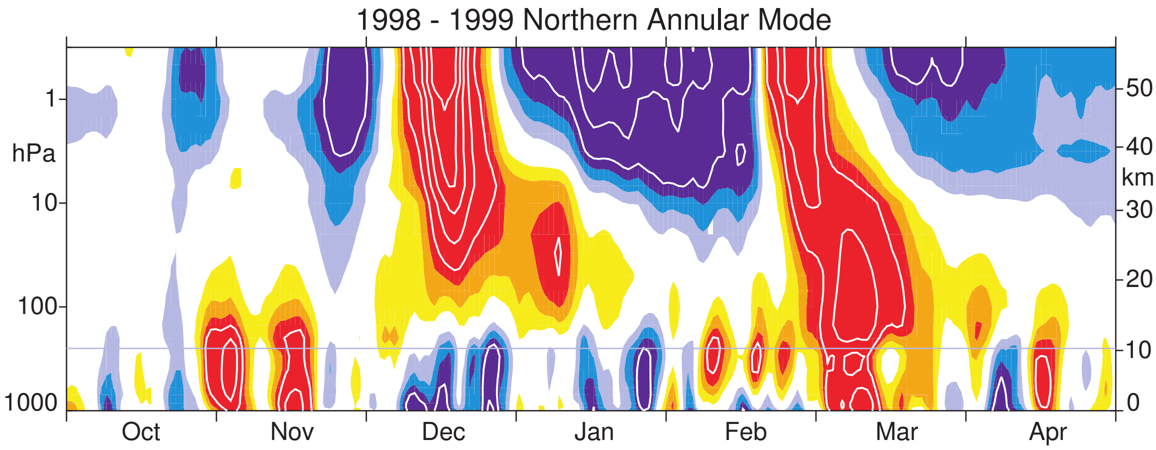

FIGURE 2.7: Time-height development of the northern annular mode during the winter of 1998–1999. The indices have daily resolution and are nondimensional. Blue corresponds to positive values (strong polar vortex), and red corresponds to negative values (weak polar vortex). The contour interval is 0.5, with values between −0.5 and 0.5 unshaded. The thin horizontal line indicates the approximate boundary between the troposphere and the stratosphere. Figure from *Baldwin and Dunkerton* [2001].

are highly correlated with SSWs). It is apparent that the NAM signal appears to descend in the case of the latter event (but not the former) and consistent anomalies remain in the troposphere for approximately one month. This timescale is much longer than the usual timescales of tropospheric NAM variability [e.g., *Simpson et al.*, 2011] but is more representative of the time taken for the stratosphere polar vortex to recover following a SSW.

Most observational studies of stratosphere-troposphere coupling have focused on the NH because of the greater stratospheric variability there. However, *Thompson and Solomon* [2002] argued that a long-term strengthening of the SH stratospheric polar vortex resulting from ozone depletion may be causing trends in the tropospheric SAM through a dynamical link. This was supported by *Thompson et al.* [2005], who performed an analysis similar to *Baldwin and Dunkerton* [2001], finding long-lived tropospheric SAM anomalies following strengthenings and weakenings of the SH stratospheric polar vortex (although these have a much smaller magnitude to the NH equivalents). They also found a strong tropospheric SAM signal following the 2002 SH SSW.

More recent studies of tropospheric anomalies following stratospheric events have focused on more localised extreme weather events in addition to the annular modes. Several such studies have found associations between weakenings of the NH stratospheric polar vortex and an increased likelihood of extreme cold events over North America, northern Europe and eastern Asia [*Thompson et al.*, 2002; *Kolstad et al.*, 2010; *Tomassini et al.*, 2012]. These extreme cold events are often linked with persistent 'blocking' weather patterns,[13] and a number of studies have investigated the association between stratospheric variability and blocking. Although *Taguchi* [2008] found no statistically significant change in blocking frequency during periods before or after SSWs, several studies have asserted an upwards link, with blocking events preceding a weakened polar vortex [*Quiroz*, 1986; *Andrews et al.*, 1987; *O'Neill et al.*, 1994]. More recent studies have further investigated whether blocking in particular locations precedes split or displaced vortex events, although these have reached conflicting conclusions [*Martius et al.*, 2009; *Woollings et al.*, 2010; *Castanheira and Barriopedro*, 2010]. A downwards link between stratospheric variability and the likelihood of blocking events was suggested by *Kodera and Chiba* [1995]. *Woollings et al.* [2010] and *Davini et al.* [2014] have recently found more evidence for this link, including stratosphere-leading relationships between stratospheric variability and high-latitude blocking which may be associated with the stratospheric impact on the tropospheric NAM. Overall however, the nature of link between stratospheric variability and blocking remains uncertain and is a topic of active research.

It has been observed that some SSW events, while appearing to have similar magnitudes in the stratosphere, have very different signatures in the troposphere [e.g., *Baldwin and Dunkerton*, 2001; *Tomassini et al.*, 2012] (See also the two events shown in Figure 2.7). This issue was addressed by *Nakagawa and Yamazaki* [2006] and *Mitchell et al.* [2013], who compared tropospheric anomalies following SSWs dominated by wave-1/wave-2 activity and displaced/split vortex events respectively (these two

[13] These are characterised by a persistent anticyclonic anomaly that causes a reversal in the upper-tropospheric meridional gradient of a quantity such geopotential height [e.g., *Tibaldi and Molteni*, 1990].

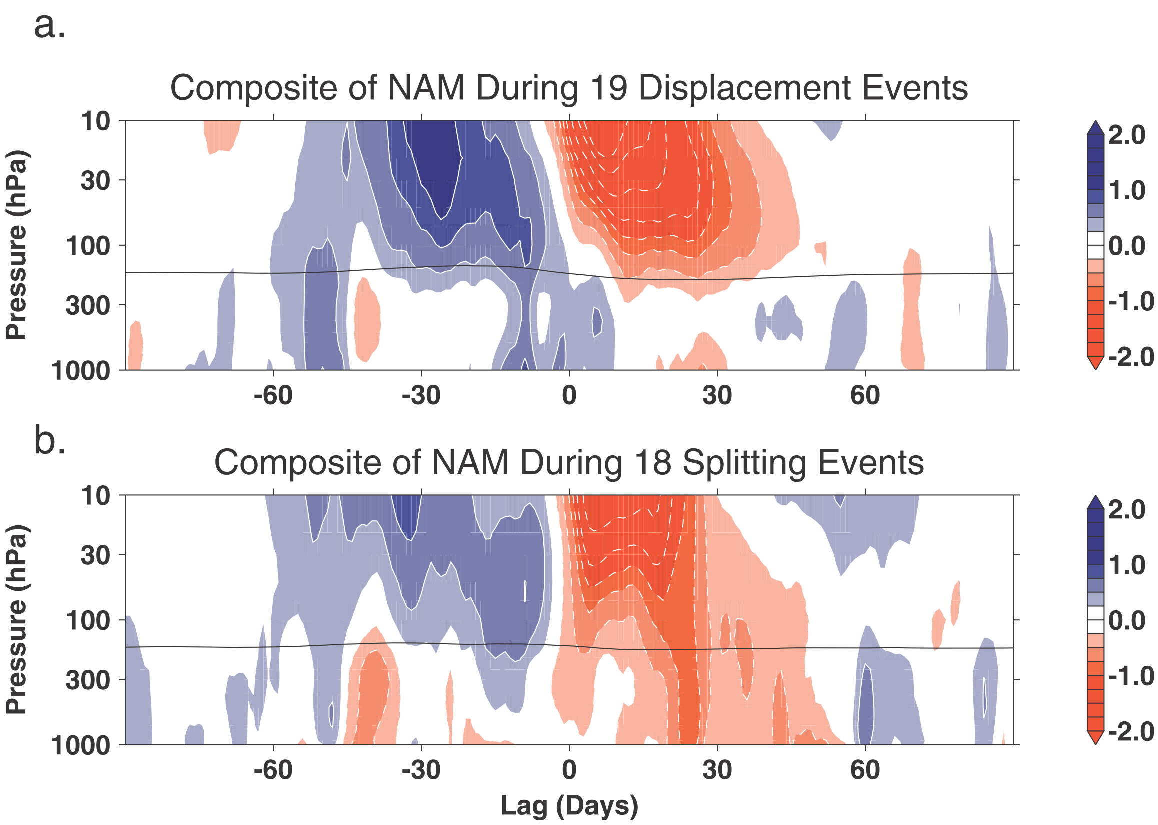

FIGURE 2.8: Composites of the time–height evolution of the NAM during (a) 19 vortex displacement events and (b) 18 splitting events. The horizontal line is a composite of the thermal tropopause level for the two types of event. Lag 0 shows the onset of an event as measured at 10 hPa. Contour intervals are 0.25 and the region between −0.25 and 0.25 is unshaded. Figure from *Mitchell et al.* [2013].

classifications are related but not identical; see *Waugh* [1997], Section 3.3.5). These studies found stronger stratospheric anomalies following wave-2/split vortex SSWs than following wave-1/displaced vortex SSWs.[14] This result is illustrated in Figure 2.8, which shows composites of the NAM following the split and displaced vortex events identified by *Mitchell et al.* [2013]. It can be seen that tropospheric anomalies are stronger following split vortex events and persist for up to two months, while the NAM signal following displaced vortex events appears not to descend below the tropopause even though its stratospheric magnitude is greater. *Mitchell et al.* [2013] went on to link these different NAM signals to an increase in high-latitude blocking following split vortex events, and a weaker blocking signal following displaced vortex events. On the other hand *Charlton and Polvani* [2007], using a different classification method, found little difference in the tropospheric signals following split and displaced vortex events. This discrepancy between these studies highlights the importance of the classification method of split and displaced vortex events (or of tropospheric variability); an issue which is addressed in Chapter 3.

It is not only mid-winter stratospheric variability which has been suggested to influence tropospheric weather. *Hardiman et al.* [2011] found differences in NH springtime sea-level pressure anomalies following two types of stratospheric final warming; more radiatively driven final warmings where the transition to easterlies happens first in the upper stratosphere and descends with time, and more dynamically driven final warmings where the transition to easterlies happens first in the mid-stratosphere.

A significant limitation of the above observational studies, which employ composite or correlation analysis, is that they cannot in themselves demonstrate *causality*. For example, there may be the possibility that a third factor (such as sea-surface temperatures) affects both the troposphere and stratosphere separately. To address this, a number of

[14]Using the method developed in Chapter 3, of the two weak vortex events illustrated in Figure 2.7, the first is identified as a displaced vortex event and the second as a split vortex event. Their tropospheric NAM may therefore be expected to be different, although no firm conclusions should be reached from the study of just two events.

modelling studies have been carried out which measure the tropospheric response to some imposed stratospheric perturbation. These are described in the next section.

2.4.2 Modelling evidence

Modelling investigations into a possible stratospheric influence on the troposphere predate the observational evidence described above. In a general circulation model (GCM) study, *Boville* [1984] showed that changes to upper tropospheric and lower stratospheric zonal mean zonal winds have a significant effect on mid-troposphere wave fields, although the wind changes imposed in his model were larger than those in reality. Using more realistic wind variations, *Jacqmin and Lindzen* [1985] found the troposphere to be largely insensitive to changes in the state of the stratosphere. However, several more recent studies using models with a more realistic representation of stratospheric dynamics, have found a consistent tropospheric response to an imposed stratospheric torque [e.g., *Polvani and Kushner*, 2002; *Norton*, 2003; *Taguchi*, 2003]. This tropospheric response is found to resemble the negative phase of the NAO or NAM following a weakening of the vortex, and as such is consistent with the observational results discussed in the previous section.

Figure 2.9 shows the results from another such study by *Jung and Barkmeijer* [2006], who imposed a weakening of the stratospheric polar vortex using an adjoint method. Differences of zonal-mean zonal wind at four time periods following the imposition of this forcing are shown for the composite of 60 forecasts. It can be seen that anomalies are initially confined to the stratosphere but descend to the troposphere after approximately 20 days. *Jung and Barkmeijer* [2006] also found surface anomalies associated with this forcing to resemble the negative phase of the NAO, though with the southern node shifted slightly eastwards. Interestingly, they found an almost opposite surface response when a strengthening of the stratospheric polar vortex was imposed. This indicates that the surface response may be linear, potentially providing some

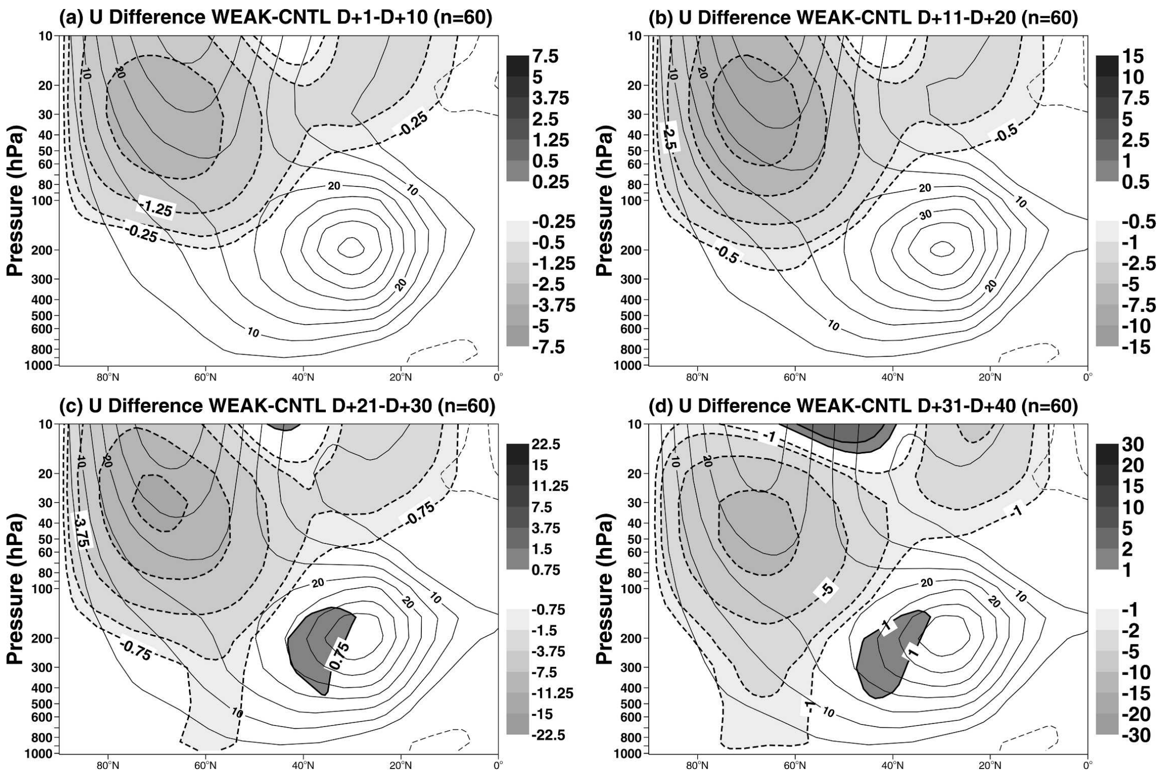

FIGURE 2.9: Difference of average zonal-mean zonal winds (shading in m s^{-1}) between the weak polar vortex (WEAK) and the control experiment (CNTL) for 10-day averages: (a) D +1 to D +10, (b) D +11 to D +20, (c) D +21 to D +30, and (d) D + 31 to D +40. Shown is the average over 60 different cases (40-day integrations). Notice that the contour interval for the differences changes linearly with the forecast range. Also shown are zonal-mean zonal winds from the control integration (contour interval is 5 m s^{-1}). Figure from *Jung and Barkmeijer* [2006].

information about the mechanism responsible for the response, as discussed in the next section.

This tropospheric response to an imposed torque was shown to also operate on long time scales by *Scaife et al.* [2005a]. They found that when a long-term acceleration of the stratospheric polar vortex was imposed from the 1960s–1990s, in line with observations, their model more faithfully reproduced the long-term trend towards a more positive phase of the NAO over this time. It is however, important to note that this does not necessarily imply that stratospheric changes were driving the trend in reality.

In contrast to the studies above, which have calculated the tropospheric response to an imposed stratospheric forcing, *Simpson et al.* [2011] removed the stratospheric influence by damping the polar stratosphere to a climatological seasonal cycle. In doing so, they found that the effect of the stratosphere is to lengthen SAM timescales[15] during the austral late spring/early summer, and to lengthen NAM timescales during the boreal winter-spring.

A criticism of studies with an imposed damping or torque is that they may not be simulating a balanced, realistic, or physical state (for instance, the imposed torque may not conserve angular momentum). As such, several studies have taken place using free-running GCMs. *Plumb and Semeniuk* [2003] found using a simple free-running stratosphere-only model that descending annular mode-like signals of the kind shown in Figure 2.7 can be simulated only by changing the forcing at the lower boundary. They therefore conclude that observations of this kind do not necessarily indicate a downwards influence from the stratosphere. However, *Gerber et al.* [2008] undertook a similar study, comparing relatively simple GCM simulations with varying representations of the stratosphere. They altered the strength and variability of the stratospheric polar vortex independently of the troposphere by varying the stratospheric lapse rate, finding that simulations with a more variable polar vortex also had longer timescales of the tropospheric NAM (a result similar to that of *Simpson et al.* [2011], although using a free-running model). This lengthened timescale, in turn, was related to the long-lived tropospheric anomalies following SSWs. Using the same model, *Gerber et al.* [2009] performed a series of ensemble forecasts of model-simulated SSW events. They compared events in which the initial tropospheric NAM was negative and positive, finding a negative NAM to be much more likely following the SSW event in both cases. Hence, they conclude that the stratosphere does indeed exert a significant downwards influence on the troposphere in their model.

A further class of model investigation has been to compare models with a well-resolved stratosphere with models with a coarser stratospheric resolution; known as 'high-top' and 'low-top' models respectively [*Huebener et al.*, 2007; *Sigmond et al.*, 2008;

[15]The SAM/NAM timescale is defined as the lag (in days) for its autocorrelation to fall by $1/e$.

Cagnazzo and Manzini, 2009; *Sassi et al.*, 2010; *Scaife et al.*, 2011a; *Charlton-Perez et al.*, 2013] (discussed in greater detail in Section 4.1). These have all indicated a more realistic representation of tropospheric variability and long-term trends is achieved with improved vertical resolution. Additionally, in a case study of the cold European winter 2005–2006, *Scaife et al.* [2008] found increased blocking activity in a model with greater stratospheric resolution. On the other hand, in a multi-model comparison of blocking, *Anstey et al.* [2013] found a stronger relationship between blocking frequency and upper-troposphere/lower-stratosphere vertical resolution than with model lid height. Overall, large biases remain in models' representation of blocking. A difficulty of these studies is that there are often several differences between high- and low-top models besides their vertical resolution (such as their parametrisation schemes), so it is difficult to pin down any differences between the simulations to a stratospheric influence alone. *Hardiman et al.* [2012] attempted to address this issue by comparing model simulations which differ only in their vertical resolution above 15 km. They found the climatology and trends of surface temperature to be largely insensitive to the increased vertical resolution, although stronger surface anomalies following SSWs and a more realistic trend in the NAO were found in the high-top model.

Several studies have investigated the influence of the stratosphere on the skill of medium-range forecasts (also discussed in Section 5.1). These have demonstrated improvements in the medium-range predictive skill of high-top relative to low-top models [*Marshall and Scaife*, 2010; *Roff et al.*, 2011], and enhanced predictability when forecasts are initialised at the time of anomalously negative stratospheric NAM or SSWs [*Kuroda*, 2008; *Mukougawa et al.*, 2009; *Sigmond et al.*, 2013]. These studies therefore further indicate a downwards influence of stratospheric variability on the troposphere.

2.4.3 *Mechanisms*

The results described in the previous two sections provide strong evidence that the stratosphere does indeed exert a significant influence on tropospheric weather and climate. Even so, may uncertainties about the exact nature of the influence remain; observational studies are always open to the criticism that they do not demonstrate causality, and modelling studies that they contain large model biases or are simulating an unrealistic scenario. A coherent picture of stratosphere-troposphere coupling also requires a physical understanding of the mechanism by which this takes place. Several such mechanisms have been proposed, and are briefly described below:

Radiative effects. It is well known that stratospheric (particularly lower-stratospheric) temperature increases following SSWs lead to increasing downwelling longwave radiation entering the troposphere. *Ramanathan* [1977] argued that this warming could reduce the available potential energy accessible to tropospheric eddy activity. Similarly, the stratospheric cooling associated with ozone depletion leads to a reduction in downwelling longwave radiation entering the troposphere [*Forster and Shine*, 1997]. *Grise et al.* [2009] used a radiative transfer model to asses the impact of these radiation changes, finding a significant impact on surface polar temperatures, although they did not assess the circulation response. Indeed, relatively few studies have assessed the impact of radiative processes in stratosphere-troposphere coupling in the light of more recent observations. Despite this, it is unlikely that radiative effects play a large role in the observed stratosphere-troposphere coupling since relatively realistic coupling can be simulated in a model lacking interactive ozone or a detailed radiation scheme.

Baroclinic instability. The growth rate of baroclinic eddies is related to the vertical shear in zonal wind throughout the troposphere [*Eady*, 1949]. Some studies have suggested that lower stratospheric zonal wind anomalies which penetrate into the upper troposphere affect this vertical wind shear, and so the growth of baroclinic eddies [*Wittman et al.*, 2007; *Chen and Zurita-Gotor*, 2008]. Hence, we may expect reduced tropospheric eddy activity at mid-to-high latitudes following SSWs, due to weaker lower stratosphere/upper troposphere zonal winds. *Scaife et al.* [2011a] also used this relationship to demonstrate that an increase in European winter storminess and equatorward shift in the North Atlantic storm track projected under climate change are consistent with a weakening and equatorward shift of the stratospheric polar vortex. It is less clear whether the change in baroclinic eddy activity is sufficient to lead to the annular mode signals of the kind shown in Figure 2.7, although tropospheric eddies have been demonstrated to be important in driving annular mode variability [*Limpasuvan and Hartmann*, 1999].

Downward control. Under steady-state or time-mean conditions, the 'downward control' principle of *Haynes et al.* [1991] shows that the streamfunction, ψ, of the TEM residual circulation is given by

$$\psi(\phi, z) = \int_z^\infty \left(\frac{\rho_0 a \overline{\mathscr{F}} \cos^2\phi}{\partial \overline{m}/\partial \phi} \right)_{\phi=\phi(z')} \mathrm{d}z', \tag{2.13}$$

where $\overline{\mathscr{F}}$ is the total wave driving term from Equation 2.10, and $\overline{m} = a\cos\phi(\overline{u} + a\Omega\cos\phi)$ is the angular momentum per unit mass. Strictly, the integration is along a line of constant angular momentum, but this is approximated as vertical (an approximation which breaks down near the equator). Hence, it can be seen that under these conditions the residual circulation at a given altitude depends only on the wave drag above that altitude.

This relation therefore shows that circulation induced by stratospheric wave drag extends to the Earth's surface, although it should be noted that the above assumptions (particularly steady-state) do not strictly hold in the real atmosphere. *Thompson et al.* [2006] suggested that this induced residual circulation (or 'balanced response') is sufficient to explain the observed stratosphere-troposphere coupling. Others, however, have argued that the observed annular mode-like tropospheric response cannot be generated in the zonal-mean framework of downward control, necessitating feedbacks involving tropospheric eddies [*Kushner and Polvani*, 2004; *Song and Robinson*, 2004].

Planetary wave reflection/refraction. The critical surface formed in the stratosphere during SSW events acts to reflect upward-propagating planetary waves, as described in Section 2.1.2.1. *Perlwitz and Harnik* [2003] argued that these reflected planetary waves re-enter the troposphere and affect the tropospheric wave structures, leading to the observed tropospheric response. *Shaw et al.* [2010] have further suggested that strong two-way coupling exists in the presence of both a vertical reflecting surface and a strong meridional PV gradient, which act to create a wave guide.

Other studies have suggested that it is the refraction of planetary waves at lower levels, in the upper-troposphere/lower stratosphere (UTLS), that is most important for communicating stratospheric anomalies to the surface [*Limpasuvan and Hartmann*, 2000; *Hartmann et al.*, 2000]. Indeed, *Chen and Robinson* [1992] showed that planetary wave propagation is very sensitive to small changes in this region (also discussed by *Haynes* [2005]).

Local adjustment to PV anomalies. Under the QG approximation, PV anomalies, q', can be related to geopotential height anomalies, Z', through

$$q' = \mathscr{L}(Z'), \tag{2.14}$$

where $\mathscr{L}$ is a linear Laplacian-like operator [*Charney and Stern*, 1962]. It follows that the geopotential anomalies (and with QG approximations, other dynamical fields) associated with a given PV anomaly can be derived by inverting this operator. The geopotential height anomalies associated with a given PV anomaly can be thought to be localised in that they decay with a typical vertical and horizontal scales H and L,

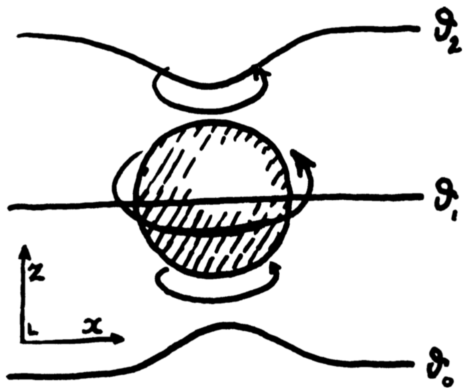

FIGURE 2.10: Schematic of the bending of isentropic surfaces (labelled θ_0, θ_1, and θ_2) toward a positive potential vorticity anomaly. The arrows represent winds associated with the potential vorticity anomaly, becoming weaker away from the anomaly. Figure from *Ambaum and Hoskins* [2002].

which are the scale height and horizontal scale of the QG flow respectively[16]. In the case of lower-stratospheric PV anomalies, their influence may therefore extend to the troposphere. *Hartley et al.* [1998] and *Black* [2002] performed such PV-inversions to calculate the tropospheric effect of stratospheric PV anomalies. They found stratospheric PV anomalies to contribute significantly to anomalies at the tropopause, which extend to the Earth's surface.

It can be argued that PV-inversion does not constitute a 'mechanism' since it is not a time-dependent relationship and does not imply a direction of causality. However, *Ambaum and Hoskins* [2002] developed a physical mechanism analogous to PV-inversion, as follows. Anomalies of PV act to bend isentropic surfaces [*Hoskins et al.*, 1985], as illustrated in Figure 2.10. At high latitudes, the tropopause lies approximately on a potential vorticity surface, also the potential temperature of the tropopause will be conserved for adiabatic changes (Section 2.1.2.1). Hence, the tropopause will also bend in the presence of a stratospheric potential vorticity anomaly; moving upwards for a positive anomaly, and downwards for a negative anomaly. This deformation of the tropopause will then lead to hydrostatic and geostrophic adjustment of the

[16]It follows from this that PV-inversion is global in the sense that knowledge of the PV field everywhere is needed in order to fully determine the other dynamical fields.

tropospheric column below, which can be thought of in terms of the conservation of angular momentum, with a stretched tropospheric column leading to an increase of vorticity and lower pressure. It is therefore be expected that a strengthening of the stratospheric polar vortex would be associated with negative sea-level pressure anomalies over the pole, and a weakening with positive pressure anomalies, which is consistent with the observed annular mode relationships (Section 2.4.1).

Importantly, *Ambaum and Hoskins* [2002] noted a stratosphere-leading time-lagged relationship between the strength of the stratospheric polar vortex, the tropopause height and the NAO, with the stratosphere acting to integrate the NAO. This therefore represents a time-dependent and causal mechanism for stratosphere-troposphere coupling.

The above list is not exhaustive but represents the most prominent proposed mechanisms. It is clear that several of these mechanisms are not mutually exclusive and so it is likely that more than one is at work in the real atmosphere. Furthermore, a difficulty in distinguishing the relative importance of the different mechanisms is that they are largely consistent with observations, particularly in the tropospheric annular mode response to changes in stratospheric polar vortex strength. It is therefore important to identify situations in which mechanisms make different predictions and test these against reality and numerical simulations.

One such situation is in the response to zonally asymmetric stratospheric anomalies. For instance, the planetary wave reflection and downward control mechanisms are largely sensitive only to zonal-mean quantities. Hence, it might be expected that under these mechanisms the tropospheric response to two equal zonal-mean stratospheric anomalies would be the same, even if the anomalies have different zonal asymmetries. On the other hand, the tropospheric response via adjustment to PV anomalies would be expected to be local to the stratospheric anomalies.

It is the aim of Chapters 3 and 4 to diagnose zonally asymmetric stratospheric variability, specifically split and displaced vortex events, in observations and model simulations. Part of the motivation for this analysis is then to study the tropospheric response in order to test the relative importance of the above mechanisms.

CHAPTER 3

A geometrical description of vortex variability

Much of the work contained in this chapter is based upon *Seviour et al.* [2013], published in *Geophysical Research Letters*, although the analysis presented here has been significantly extended.

3.1 Introduction

A quantitative description of stratospheric polar vortex variability is desirable for a number of reasons; it allows for the comparison of different studies, observational data sets, and model simulations, as well as permitting robust definitions of extreme events. Traditional methods to quantify vortex variability have been based on zonal-mean diagnostics, such as the zonal-mean zonal wind [e.g., *Andrews et al.*, 1987]. This was motivated both by the simplicity of these diagnostics and the physical reasoning that the strength of the zonal flow controls the propagation of planetary waves [*Charney and Drazin*, 1961, Section 2.1.2]. *McInturff* [1978] provided the first quantitative definition of SSWs[1] (referred to in that text as "major stratospheric warmings") using zonal mean quantities as below.

[1] In the literature, this is often called "the WMO definition", although at least two different definitions can be attributed to the WMO [*Butler et al.*, 2014].

> *A stratospheric warming can be said to be major if at 10 mb or below the latitudinal mean temperature increases poleward from 60 degrees latitude and an associated circulation reversal is observed (i.e., mean westerly winds poleward of 60° latitude are succeeded by mean easterlies in the same area).*

A number of variations of this definition have since appeared in the literature. Most commonly, the temperature gradient criterion has been neglected and/or zonal wind reversals at a particular latitude (usually 60°N) used instead of the stricter criterion of a reversal everywhere poleward of 60°N [e.g., *Labitzke and Naujokat*, 2000; *Christiansen*, 2001; *Reichler et al.*, 2012].

Although the reversal of zonal-mean zonal wind is physically relevant for the propagation of planetary waves, the choice of 60°N and 10 hPa in the definition of SSWs is less physically significant. Indeed, different numbers of SSWs are identified if these locations are varied. *Butler et al.* [2014] found that a greater number of events are identified if the threshold is located either equatorward or poleward of 60°N. Some studies have aimed to avoid this sensitivity to spatial location by quantifying vortex variability through empirical orthogonal function (EOF) analysis, using fields over a larger area. This includes the Northern Annular Mode (NAM) (calculated either from the three-dimensional geopotential height field [*Baldwin and Dunkerton*, 2001] or zonal-mean geopotential height [*Baldwin and Thompson*, 2009]), EOFs of zonal wind [*Limpasuvan et al.*, 2004], and vertical profiles of polar cap-averaged temperature [*Kuroda*, 2004]. SSW events are then defined by a threshold in the principal component of the relevant EOF.

As it has become increasingly recognised that SSWs generally occur as either split or displaced vortex events, studies have aimed to objectively distinguish these two types of event. Commonly this has been achieved through Fourier decomposition of the zonal wave structure. For instance, *Nakagawa and Yamazaki* [2006] defined SSWs through a polar temperature criterion and then split these events into two groups depending on whether the 150 hPa Eliassen-Palm (EP) flux prior to the events was dominated by zonal wavenumber one or two. *Charlton and Polvani* [2007] (hereafter CP07) introduced a new classification method, which does not rely on Fourier decomposition; first they identified events using the traditional wind reversal at 60°N, 10 hPa criterion, then they calculated the circulation around the two largest contours of relative vorticity on the vortex edge. If these two contours have a circulation ratio of 2:1 or lower the event is classified as a split, and all other events are automatically classed as displacements.

Both Fourier decomposition of the zonal wave structure and the method of CP07 rely on an Eulerian framework, with fields analysed at a fixed spatial location. *Waugh* [1997] first applied two-dimensional moment diagnostics (otherwise known as elliptical diagnostics) to the stratospheric polar vortex to provide an alternative semi-Lagrangian (or vortex-oriented) framework. These diagnostics are calculated by fitting an ellipse to a contour and then determining its properties such as the centre, orientation, aspect ratio, and area (a further diagnostic, excess kurtosis–a measure of the 'peakedness' of the distribution–was introduced by *Matthewman et al.* [2009]). This allows the movement and elongation of the vortex to be quantified. *Waugh* [1997] also compared these diagnostics to the traditional Fourier decomposition. He showed that wave-1 and 2 amplitudes relate most strongly to the displacement and elongation of the vortex respectively, however, these relationships were not found to be strong, with correlations of daily values less than 0.5. These weak relationships were attributed to the fact that planetary wave propagation can be affected by changes in the meridional PV gradient, even if the vortex shape and location are fixed. Furthermore, the wave-1 amplitude depends to some extent on the elongation of the vortex as well as the location of the centre (and similarly for the wave-2 amplitude). He concluded that it is difficult to extract quantitative information about the shape and location of the vortex based on wave amplitudes alone, highlighting the advantages of the moment diagnostics.

Hannachi et al. [2010] then applied a hierarchical clustering algorithm to daily values of the area, centroid latitude, and aspect ratio diagnostics and found that the vortex falls preferably into three clusters corresponding to undisturbed, split, and displaced states. These groupings were used by *Mitchell et al.* [2013] (hereafter M13) to identify split and displaced vortex events; if the vortex remained in the split or displaced cluster for at least five consecutive days it was classified as the corresponding event. Significantly, as discussed in Section 2.4.1, M13 demonstrated that split vortex events penetrated deep into the troposphere and resulted in significant surface anomalies, while anomalies associated with displaced vortex events do not descend far below the tropopause. This is in agreement with *Nakagawa and Yamazaki* [2006] who found tropospheric anomalies to be larger following SSWs with dominant wave 2 amplitude, however, it contrasts with CP07, who found little difference in the tropospheric impact of split and displaced vortex events. This highlights the potential importance of the method of classification of split and displaced vortex events in any study.

In this chapter we wish to develop a method for the classification of split and displaced vortex events with the following properties:

- It is based on vortex moment diagnostics.
- It can be easily applied to a range of data sets, including climate model simulations.
- It is computationally inexpensive.

The motivation for the use of moment diagnostics includes their advantages in quantifying the shape and location of the vortex, as noted above. This, in turn, is desirable because the location of the vortex near the tropopause may be important for understanding the regional tropospheric effect of stratospheric anomalies [e.g., *Ambaum and Hoskins*, 2002, Section 2.4.3]. Previous calculations of vortex moment diagnostics have been based on the distributions of quasi-conservative tracers such as PV on isentropic surfaces [*Mitchell et al.*, 2011] or long-lived tracer (e.g., N_2O) concentrations [*Waugh*, 1997]. These quantities have strong meridional gradients allowing for clear determination of the vortex edge [*Nash et al.*, 1996]. Unfortunately, many climate models do not output PV or tracer concentrations, and these are often computationally expensive or impractical to calculate. As such, we wish to develop a method which uses geopotential height, a variable which is output by all contemporary climate models. This effort will also allow us to test the robustness of the result of M13 regarding the different surface impacts of split and displaced vortex events using a semi-independent classification method and extended data set.

The remainder of this chapter is structured as follows. The next section introduces the necessary theoretical background for the calculation of moment diagnostics. Section 3.3 describes the methods used for the classification of split and displaced vortex events, and compares these events with those determined by M13 and CP07. Section 3.4 contrasts the surface impacts of split and displaced vortex events calculated using the new method and discusses potential mechanisms behind any differences.

3.2 Vortex moment diagnostics

The moments, M_n, of a one-dimensional distribution can be classified by their order, n, and provide familiar parameters. These are the area under the distribution (0th order), mean (1st order), variance (2nd order), skewness (3rd order), and kurtosis (4th order), given by

$$M_n = \int_S x^n f(x)\,\mathrm{d}x\,, \tag{3.1}$$

where S represents the extent of the distribution, $f(x)$, to be integrated over. The extension of this for a two-dimensional distribution is straightforwardly

$$M_{nm} = \iint_S x^n y^m f(x,y)\,\mathrm{d}x\mathrm{d}y\,, \tag{3.2}$$

where the order of the moment is now defined as $m+n$, meaning it is possible to have different diagnostics with the same order (e.g., M_{01}, M_{10}). Although these diagnostics can be further extended to three dimensions, this has been demonstrated to be highly computationally expensive [*Li and Ma*, 1994], and would require assumptions about the lower and upper bounds of the vortex region. We therefore calculate two-dimensional moment diagnostics for the stratospheric polar vortex on quasi-horizontal surfaces. We use two variables; geopotential height ($f(x,y)=Z(x,y)$) on the 10 hPa pressure level, and potential vorticity ($f(x,y)=q(x,y)$) on the 850 K potential temperature (isentropic) surface, which lies close to 10 hPa. Following *Waugh* [1997], the calculation of moment diagnostics is simplified by transforming the spherical data $q(\phi,\lambda)$ and $Z(\phi,\lambda)$, where ϕ is latitude and λ longitude, to Cartesian coordinates using the polar stereographic projection

$$x=\frac{\cos\lambda\cos\phi}{1\pm\sin\phi}, \quad y=\frac{\pm\sin\lambda\cos\phi}{1\pm\sin\phi}, \tag{3.3}$$

where the positive sign is used in the NH and negative in the SH.

In order to calculate moment diagnostics for the stratospheric polar vortex we must first isolate the vortex region by defining the vortex edge. Different methods have previously been used for this calculation; *Waugh and Randel* [1999] used the mean PV at the maximum of the mean meridional PV gradient, while *Matthewman et al.* [2009] defined the vortex edge on a daily basis, using the average value of PV poleward of 45°N nine days before the onset of a SSW (their SSWs were defined by zonal-mean zonal wind reversal, as in CP07). A more complex method due to *Nash et al.* [1996], starts by transforming PV to 'equivalent latitude' [*Butchart and Remsberg*, 1986] coordinates, before defining the vortex edge as the position of the largest gradient in a plot of PV against equivalent latitude. This method was applied in *Mitchell et al.* [2011] to calculate the vortex edge.

None of the three methods outlined above are found to be appropriate for the present study. We wish to directly compare the PV and geopotential height-derived moments, but the methods of *Waugh and Randel* [1999] and *Nash et al.* [1996] rely on meridional gradients in PV and so may not be transferable to geopotential height.

Furthermore, the method of *Matthewman et al.* [2009] is impractical because we wish to define the events from the moment diagnostics, so will not know their dates before calculation. Instead, we pick a simple definition; PV (q_b) or geopotential height (Z_b) on the vortex edge is defined as the value of the December-March (DJFM) mean at 60°N for the NH and the June-September (JJAS) mean at 60°S for the SH. This is seen to lie close to contours defined by the above methods, and results are insensitive to small changes in the latitude chosen.

Having defined the vortex edge, we extend the method *Matthewman et al.* [2009] to isolate the vortex region by introducing a transformed PV field, $\hat{q}$, given by

$$\hat{q}(x,y) = \begin{cases} q(x,y) - q_b & \text{if } q(x,y) > q_b\,, \\ 0 & \text{if } q(x,y) \leq q_b\,, \end{cases} \tag{3.4}$$

and conversely for geopotential height

$$\hat{Z}(x,y) = \begin{cases} Z(x,y) - Z_b & \text{if } Z(x,y) < Z_b\,, \\ 0 & \text{if } Z(x,y) \geq Z_b\,. \end{cases} \tag{3.5}$$

By substituting $f(x,y) = \hat{q}(x,y)$ or $f(x,y) = \hat{Z}(x,y)$ in equation 3.2 it is then possible to calculate the moment diagnostics. The zeroth order moment diagnostic, M_{00} can be used to define the 'equivalent area', A_{eq} [*Matthewman et al.*, 2009], as

$$A_{\text{eq}} = \frac{M_{00}}{q_b} \quad \text{or} \quad A_{\text{eq}} = \frac{M_{00}}{Z_b}\,, \tag{3.6}$$

depending on whether PV or geopotential height based diagnostics are calculated. Because $M_{00} \approx Aq$, where A is the vortex area, the equivalent area can be considered a measure of both vortex strength and area. The first order moment diagnostic can be used to calculate the vortex centroid,

$$(\bar{x}, \bar{y}) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right)\,. \tag{3.7}$$

In order for higher order moment diagnostics to be useful, the moment equation (3.2), must be transformed to the *centralised moment* form [*Hall*, 2005]. This calculates moments relative to the vortex centroid, and is given by

$$J_{mn} = \iint_S f(x,y)(x-\bar{x})^n(y-\bar{y})^m \, \mathrm{d}x\mathrm{d}y \,. \tag{3.8}$$

Two useful parameters can be derived from the second-order centralised moment diagnostics, the vortex orientation, ψ (defined as the angle between the major axis of the ellipse and the x-axis) and the aspect ratio, r (defined as the ratio of the lengths of the major to minor axes), given by

$$\psi = \frac{1}{2}\tan^{-1}\left(\frac{2J_{11}}{J_{20}-J_{02}}\right), \tag{3.9}$$

$$r = \left|\frac{(J_{20}+J_{02}) + \sqrt{4J_{11}^2 + (J_{20}-J_{02})^2}}{(J_{20}+J_{02}) - \sqrt{4J_{11}^2 + (J_{20}-J_{02})^2}}\right|^{1/2} . \tag{3.10}$$

Using the area, centroid, orientation, and aspect ratio, the *equivalent ellipse* can be uniquely defined. Figure 3.1 shows the equivalent ellipse calculated from both PV and geopotential height fields over a 16-day period centred on a displaced vortex event (classified using the method in Section 3.3). It can be seen that the equivalent ellipse provides a qualitatively good fit to the vortex, although this is less good in Figures 3.1(c,f) when the vortex becomes less elliptical and filamentation occurs. Greater fine-scale structure and filamentation is visible in the PV field due to its quasi-conservative properties, however reasonable agreement can be seen between the PV and geopotential height ellipses.

Equivalent ellipses for an example of a split vortex event are shown in Figure 3.2. It can be seen that after the vortex has separated the equivalent ellipse becomes less physically significant, as it spans the two vortices. *Matthewman et al.* [2009] introduced the 4th order moment diagnostic, “excess kurtosis”, in order to identify splits of the polar vortex; it is given by

$$\kappa_4 = M_{00}\frac{J_{40}+2J_{22}+J_{04}}{(J_{20}+J_{02})^2} - \frac{2}{3}\left[\frac{3r^4+2r^2+3}{(r^2+1)^2}\right] . \tag{3.11}$$

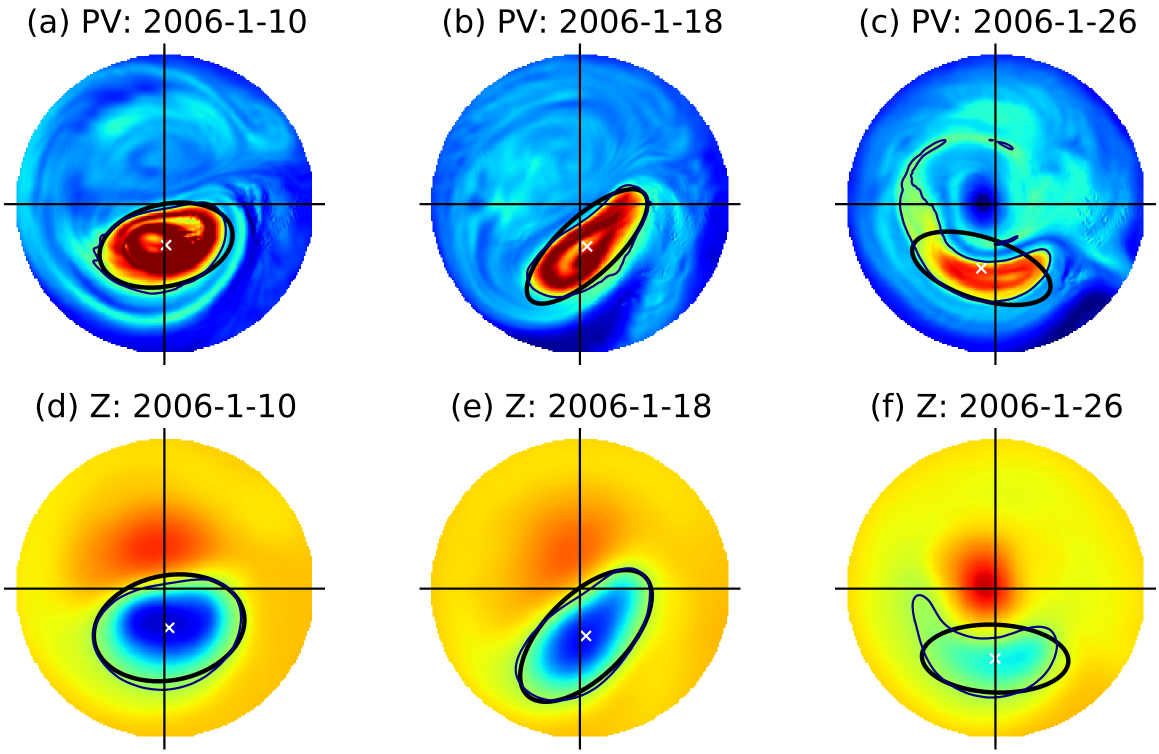

FIGURE 3.1: PV on the 850 K θ surface (a,b,c) and geopotential height at 10 hPa (d,e,f) 8 days before (a,d), at onset (b,e), and 8 days following the onset (c,f) of a displaced vortex event. Contours of q_b and Z_b are shown in thin black lines, the equivalent ellipse in a thick dark line, and its centroid with a white cross. Data are transformed to Cartesian coordinates with a polar stereographic projection.

This has the property of being negative for a vortex with a "pinched" shape, zero for a perfectly elliptical vortex, and positive for a vortex with a strong central core. When negative kurtosis was detected *Matthewman et al.* [2009] split the PV field into two regions along the minor axis of the equivalent ellipse and re-calculated moment diagnostics for the vortices in these regions separately.

In this study we do not make use of the excess kurtosis or calculate separate diagnostics for split vortices for three reasons. First, as a 4th order diagnostic it is a highly skewed variable, making its use in event classification problematic (this was also found by *Hannachi et al.* [2010]). Second, this procedure is more computationally expensive, requiring about three times the number of calculations during split vortex events. Third, kurtosis is highly sensitive to horizontal resolution [*Mitchell et al.*, 2011], and so may not be a suitable diagnostic in the comparison of climate models with different resolutions. Hence, we calculate single moment diagnostics even when the vortex has split, but bear in mind that these may not represent the properties of any real vortex.

Code for the calculation of moment diagnostics using the method described in this section is available from `https://github.com/wseviour/vortex-moments`. Additionally, a comparison of aspect ratio and centroid latitude with zonal wavenumber amplitudes is given in Section 3.3.5.

3.3 Data and methods

3.3.1 Reanalysis data

For the analysis in this chapter NH winter daily-mean data for December-March (DJFM) are employed from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalyses. The ERA-40 data set [*Uppala et al.*, 2005] is used from 1958-1978 and ERA-Interim [*Dee et al.*, 2011] from 1979-2009. The combination of these two data sets is chosen in order to maximise the total number of years entering the analysis (ERA-40

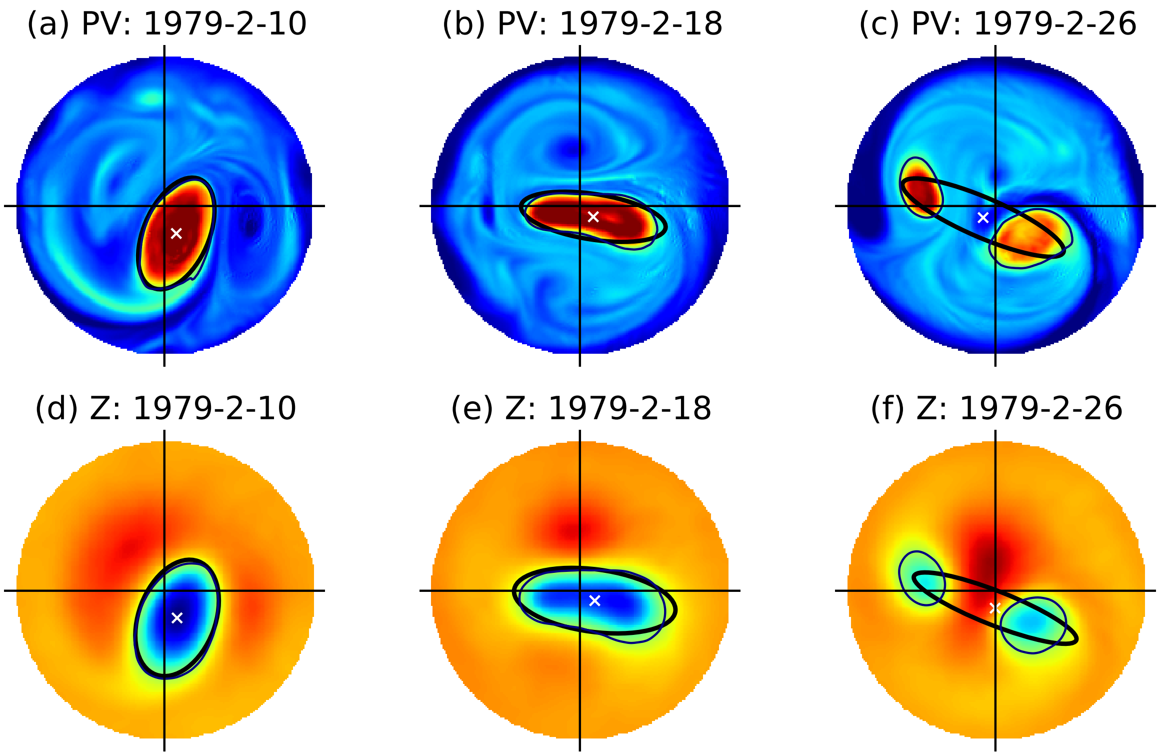

FIGURE 3.2: As Figure 3.1 but for a split vortex event.

runs only to 2002), as well as to compare results from the more recent ERA-Interim with previous studies using only ERA-40, such as *Charlton and Polvani* [2007] and *Mitchell et al.* [2013].

ERA-Interim is similar to ERA-40 but uses a four-dimensional variational data assimilation system (4D-Var) as opposed to the 3D-Var system used in ERA-40. It also has higher horizontal and vertical resolution, improved humidity analysis, model physics, data quality control, bias handling and other improvements as noted in *Simmons et al.* [2007]. The majority of observational data for the stratosphere entering both reanalyses are from radiosonde and satellite measurements. It is important to note that in the pre-satellite era (1958-1971) observations in the stratosphere were much more sparse, leading to greater errors in reanalyses during this time [*Uppala et al.*, 2005].

A number of studies have evaluated the stratospheric circulation in ERA-40 and ERA-Interim against other observations or reanalyses. *Randel et al.* [2004] found ERA-40 to closely match measurements of the zonal stratospheric circulation derived from radiosonde, rocketsonde and lidar measurements. *Karpetchko et al.* [2005] found that the representation of the polar vortices in ERA-40 agrees well with the NCEP/NCAR reanalysis, and CP07 demonstrated that this also holds for the occurrence of SSWs. *Seviour et al.* [2012] showed that the strength of the stratospheric meridional mean stratospheric circulation in ERA-Interim agrees well with previous reanalysis, but that the residual vertical velocity is more smoothly represented.

In order to perform a consistent analysis across the two data sets, ERA-Interim data is linearly interpolated to the lower resolution ERA-40 ($1.125° \times 1.125°$) Gaussian grid. PV is also interpolated from pressure levels to the 850 K isentropic surface (which lies close to 10 hPa), as this quantity has the property of being conserved under adiabatic flow. Both in the calculation of the vortex edge (climatological mean q or Z at 60°N/S) and the moment diagnostics themselves, no clear jumps were seen between ERA-40 and ERA-Interim data sets. As such, the two are considered together with no bias corrections. For the remainder of this thesis, this combined ERA-40 and ERA-Interim data set is referred to as *ERA*.

3.3.2 Moment diagnostic calculation

In order to calculate the moment diagnostics, the values of PV and geopotential height on the vortex edge (q_b and Z_b) must first be determined. These are the 60°N DJFM/60°S JJAS mean values of PV at 850 K (q_{850}) and 10 hPa geopotential height (Z_{10}) respectively. They are found to be $q_b = 460$ PVU (1 PVU $= 10^{-6}$ Km^{2}kg^{-1}s^{-1}) and $Z_b = 30.2$ km for the NH, and $q_b = 618$ PVU and $Z_b = 29.0$ km for the SH. Using these values the moment diagnostics are calculated from ERA data for 1958–2009 using the method described in Section 3.2.

As discussed in Section 3.2 the excess kurtosis diagnostic is not used in the present analysis. In the interests of simplicity, only the aspect ratio and centroid latitude diagnostics are used to identify events, and the centroid longitude, orientation and

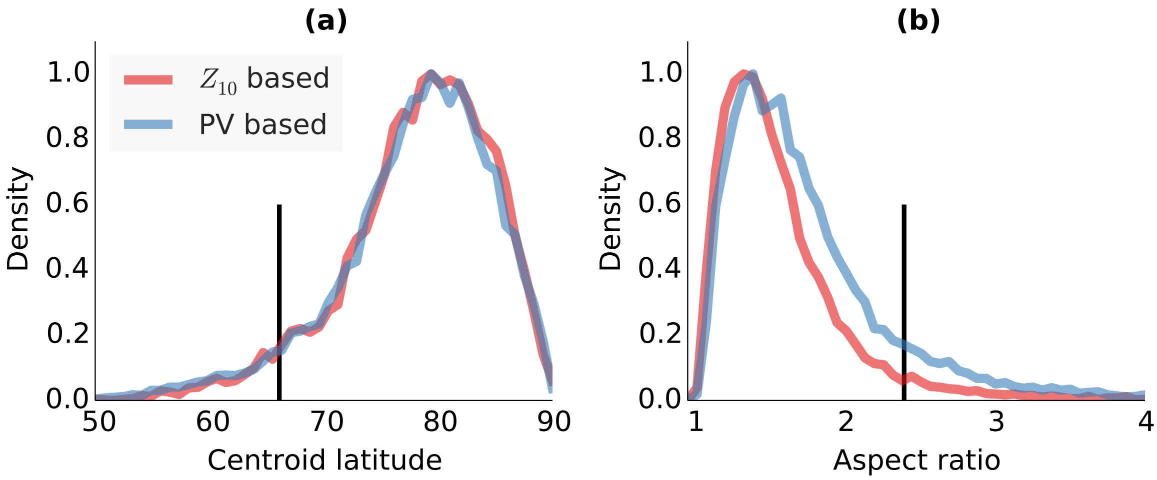

FIGURE 3.3: Distributions of the December-March centroid latitude (a) and aspect ratio (b), of the NH stratospheric polar vortex over 1958–2009. Diagnostics are calculated from geopotential height at 10 hPa (Z_{10}) and potential vorticity at 850 K (PV). Thresholds of 66°N in centroid latitude and 2.4 in aspect ratio are used to define events, and are indicated by the black vertical lines.

equivalent area are not used. The aspect ratio and centroid latitude are the most intuitive diagnostics for this purpose, with a high aspect ratio and poleward centroid latitude expected during split vortex events, and a low aspect ratio and equatorward centroid latitude expected during displaced vortex events.

Figure 3.3 shows the distributions of these two quantities calculated from q_{850} and Z_{10} for the NH vortex. The centroid latitude distributions are almost identical, with a peak near 80°N which is in agreement with previous studies [*Waugh and Randel*, 1999; *Mitchell et al.*, 2011]. The aspect ratio distributions have a similar shape, with a peak at about 1.3, but the PV based diagnostic has a larger tail. This is because the PV field contains more small-scale filamentary structures than geopotential height (e.g. Figures 3.1 and 3.2), making high aspect ratios more likely. As well as having similar distributions, the time series of the PV and geopotential height derived diagnostics (not shown) are significantly correlated, with correlation coefficients of 0.9 for daily centroid latitude and 0.6 for aspect ratio. Overall, these results suggest that geopotential height-derived moment diagnostics are appropriate for the identification of split and displaced vortex events.

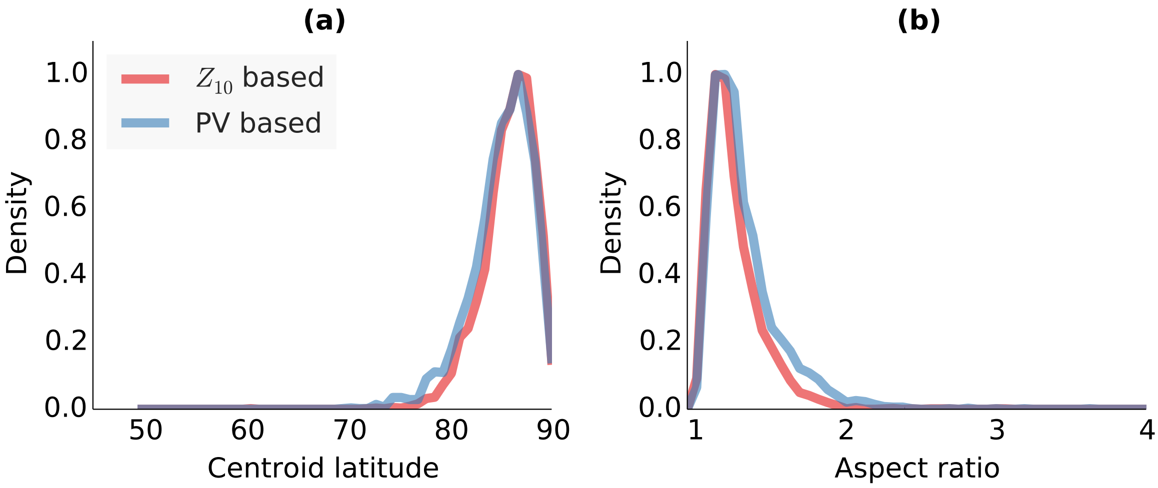

FIGURE 3.4: As Figure 3.3 but for moment diagnostics calculated for the SH stratospheric polar vortex over June-September.

Figure 3.4 shows the same distribution as Figure 3.3, but for the SH vortex aspect ratio and centroid latitude. As in the case of the Northern Hemisphere, the geopotential height and PV-based distributions have very similar shapes, with the PV-based aspect ratio having a slightly larger tail. Comparing the Northern and SH distributions it can be seen that there is much less variability in both aspect ratio and centroid latitude in the Southern Hemisphere. This is because of the reduced planetary wave propagation into the SH stratosphere, in turn a result of lesser forcing from orography and land-sea temperature contrasts. The peak in the Southern Hemisphere centroid latitude is at about 86°S; the same as that found by *Waugh and Randel* [1999].

A result of this reduced SH vortex variability is that only one SSW has been observed in the SH (discussed further in Chapter 5). The rest of this chapter relates to the classification and impacts of split and displaced vortex events and so focuses only on the NH. However, it should be noted that all the methods below can also be applied to the SH.

3.3.3 Event identification

Previous attempts to identify SSW events have used a clustering method [*Coughlin and Gray*, 2009; *Hannachi et al.*, 2010]. These methods attempt to classify the vortex state

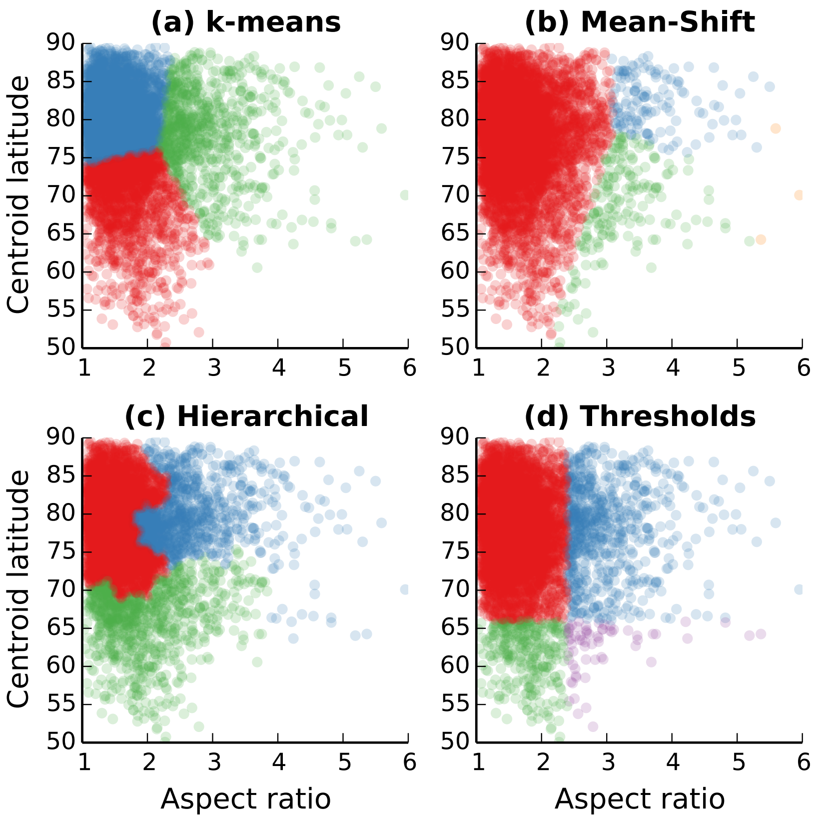

FIGURE 3.5: Three clustering algorithms and a threshold division applied to the moment diagnostics in centroid latitude-aspect ratio space. For the K-means and hierarchical algorithms three clusters were specified. The mean-shift algorithm determined the number of clusters to be 4.

for each day into a number of groups, which may be specified beforehand or determined by the clustering algorithm. Individual days within the same cluster should be physically similar, while those in different clusters distinct. More precisely, clustering aims to maximise the between-cluster variance while minimising the within-cluster variance. In the case of the stratospheric polar vortex, clusters may represent, for instance, stable, split, and displaced states. Events are then typically defined by the vortex persisting in a particular cluster for a number of days.

A large number of clustering algorithms exist, and some may be more appropriate than others for certain uses. Here, three different algorithms are applied to the moment diagnostics in centroid latitude-aspect ratio space, and their outcomes shown in Figure 3.5(a,b,c). Details of the three algorithms are given below:

(a) *K*-means clustering requires the number of clusters, K, to be specified beforehand (in Figure 3.5, $K = 3$). The algorithm begins by randomly selecting K data points to be the centroids of the initial clusters, all other data points are assigned to the cluster with the nearest centroid. Having assigned the initial cluster membership, the algorithm proceeds as follows:

1. Compute the centroids (the vector means), $\overline{\mathbf{x}}_k$ of each cluster.

2. Calculate the qdistance between the current data point, $\mathbf{x}_i$, and each of the K $\overline{\mathbf{x}}_k$s. (Various distance measures can be used; in Figure 3.5(a), the Euclidean distance is used).

3. If $\mathbf{x}_i$ is not in the group with the closest mean then reassign it to that group, otherwise repeat step 2 for $\mathbf{x}_{i+1}$.

This is repeated until a full cycle through each $\mathbf{x}_i$ produces no reassignments. An advantage of this method is that it is computationally efficient, but the major disadvantage is that the number of clusters must be pre-determined. Several methods exist to estimate the ideal number of clusters, which generally have the aim of finding the best compromise between minimising within-cluster variance and maximising between-cluster variance. *Coughlin and Gray* [2009] applied K-means clustering to several variables representing the stratospheric polar vortex. They used the method of *silhouette values* [*Rousseeuw*, 1987] to determine the ideal number of clusters to be two (representing stable and disturbed vortex states). However, three clusters has been imposed in Figure 3.5 in order to attempt to identify stable, split, and displaced states.

(b) Mean-shift clustering aims to discover 'blobs' in a data set. It works by updating candidates for centroids to be the mean of the points within a given region. That is, given a candidate centroid $\mathbf{x}_i$ for iteration t, the candidate is updated according to

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{m}(\mathbf{x}_i^t), \tag{3.12}$$

where $\mathbf{m}$ is the mean shift vector. This is calculated as

$$\mathbf{m}(\mathbf{x}_i) = \frac{\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} K(\mathbf{x}_j - \mathbf{x}_i)\mathbf{x}_j}{\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} K(\mathbf{x}_j - \mathbf{x}_i)}, \tag{3.13}$$

where $K(\mathbf{x}_j - \mathbf{x}_i)$ is a kernel function which determines the weight of nearby points. Typically, and in Figure 3.5(b), a Gaussian kernel is used, $K(\mathbf{x}_j - \mathbf{x}_i) = e^{-c\|\mathbf{x}_j - \mathbf{x}_i\|^2}$. $N(\mathbf{x}_i)$ represents the set of points for which $K(\mathbf{x}_i) \neq 0$. This shifting is repeated until $\mathbf{m}$ converges. Following this calculation, the candidates are then filtered to remove near duplicates. The greatest advantage of this method is that it automatically sets the number of clusters, so no prior assumptions about the data set are required. A disadvantage is that it requires multiple nearest neighbour searches during each iteration, and so may not be scalable to large data sets. In Figure 3.5(b) the number of clusters was determined to be four (because the fourth cluster is very small it is not easily visible in Figure 3.5(b), but represents points with high aspect ratio).

(c) Hierarchical clustering proceeds by calculating a series of nested clusters. To begin with, all data points are considered each as a separate cluster and then at each iteration the nearest two clusters are merged. There are a number of methods to identify the distance between clusters when those clusters consist of more than one member. Following *Hannachi et al.* [2010] the *complete-linkage* method is used here, defining the distance as the largest distance between members in the two groups. As with the K-means clustering, the number of clusters desired must be pre-determined, otherwise the algorithm will run to completion with a single cluster consisting of all data points. Again, many methods exist to determine the optimum number of clusters. *Hannachi et al.* [2010] used the gap statistic method [*Tibshirani et al.*, 2001] with vortex area, centroid latitude, and aspect ratio moment diagnostics, and found a slight preference for three clusters. As such, three clusters are used in Figure 3.5(c).

Figure 3.5 demonstrates that the three clustering methods produce very different results. As well as the size and extent of the clusters, there is also disagreement between this and past studies on the optimum number of clusters; *Coughlin and Gray* [2009] found two clusters using a silhouette values method, *Hannachi et al.* [2010] found three clusters using the gap statistic, while the mean-shift algorithm applied here produces four clusters. Further sensitivity tests were performed by randomly removing 1% of the data and re-calculating the clustering. It was found that very different clusters were calculated with this small alteration to the data, suggesting that these clusterings may not be robust if applied to different data sets, such as climate model simulations. The likely reason for this sensitivity is that the data itself is not highly clustered; as can be seen in Figure 3.3 no clear bi-modality is present. Rather, it is more appropriate to view the split and displaced vortex states as the tails of a distribution rather than distinct clusters or regimes.

For the reasons above, clustering methods are deemed inappropriate for the present study, and a simpler, more robust, thresholds-based method is introduced. Days with an aspect ratio > 2.4 (11% of all days) or a centroid latitude $< 66°$N (5% of all days) are classified as split and displaced states respectively. A small number of days lie beyond both thresholds, and these are classified as a mixed state (1% of all days). The vast majority of days (83%) lie in the stable state, where neither threshold is exceeded. The choice of thresholds is somewhat subjective but the results presented below are not sensitive to the exact choice of threshold. They were chosen to give a similar frequency of split and displaced vortex events (identified using the method below) as CP07 and M13.

Mitchell et al. [2011] found that above certain thresholds the aspect ratio and centroid latitude follow a generalised Pareto distribution, which is used to model extreme values [*Coles*, 2001]. Both thresholds chosen here lie beyond these extreme value thresholds of their respective distributions (these were found to be 2.3 for aspect ratio and 72°N for centroid latitude). Some theoretical motivation for the aspect ratio threshold can also be provided by the theoretical stability of an idealised elliptical vortex.

Love [1893] found that the Kirchoff ellipse (an elliptical patch of uniform vorticity in a quiescent fluid) is linearly unstable if the aspect ratio exceeds 3. The aspect ratio threshold of 2.4 used here lies below this limit, and so under this idealised model it might expect that some split vortex events do not display a full separation into two vortices.

Having classified each day into these four groups (split, displaced, mixed, and stable), a persistence criterion is introduced in order to identify split and displaced vortex *events*. A displaced vortex event requires the centroid latitude to remain equatorward of 66°N for 7 days or more, while a split vortex event requires the aspect ratio to remain higher than 2.4 for 7 days or more. A mixed event is identified if both thresholds are exceeded for 7 days or more. The onset date is defined as the day that the appropriate threshold is first exceeded, and to ensure that no events are counted twice, these onset dates are required to be spaced at least 30 days apart, chosen to reflect radiative timescales in the lower stratosphere [*Newman and Rosenfield*, 1997]. Using this method with geopotential height data, 17 displaced and 18 split vortex events (listed in Table 3.1) are identified over the 52 winters, an average of 7 per decade (no mixed events were identified). This frequency lies between the values of CP07 (6 per decade) and M13 (8 per decade). Although data is restricted to DJFM in this analysis, no measures are taken to exclude early final warmings which may occur in late March. This is motivated by the fact that these are highly dynamically driven events which may have significant impacts on the troposphere [*Hardiman et al.*, 2011]. The events defined here may therefore include some which would traditionally be classed as final warmings (i.e. the zonal-mean zonal wind does not return to westerly after the event). For this reason, these events are not referred to as SSWs, but simply as split and displaced vortex events.

Figures 3.6 and 3.7 show geopotential height at the peak of each of the split and displaced vortex events. The peak is defined as the day with the maximum aspect ratio or minimum centroid latitude in the two weeks following the onset date of split and

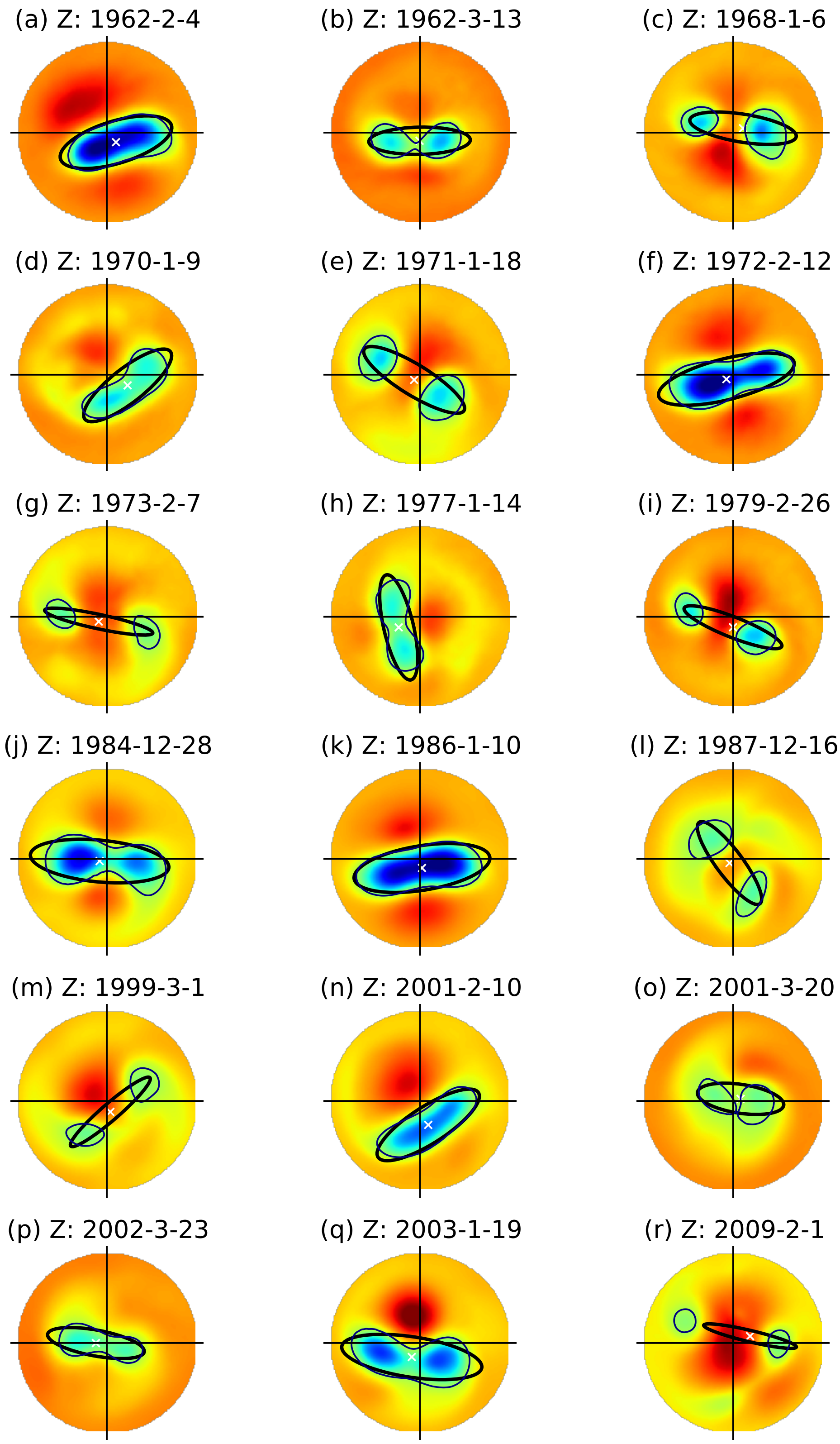

FIGURE 3.6: 10 hPa geopotential height at the peak of each of the 18 split vortex events identified in ERA. The peak is defined as the day with the largest aspect ratio during the two weeks following the onset date. The vortex edge is shown as a thin black contour, the equivalent ellipse the thick black contour and its centroid as a white cross. Data are transformed to Cartesian coordinates with a polar stereographic projection.

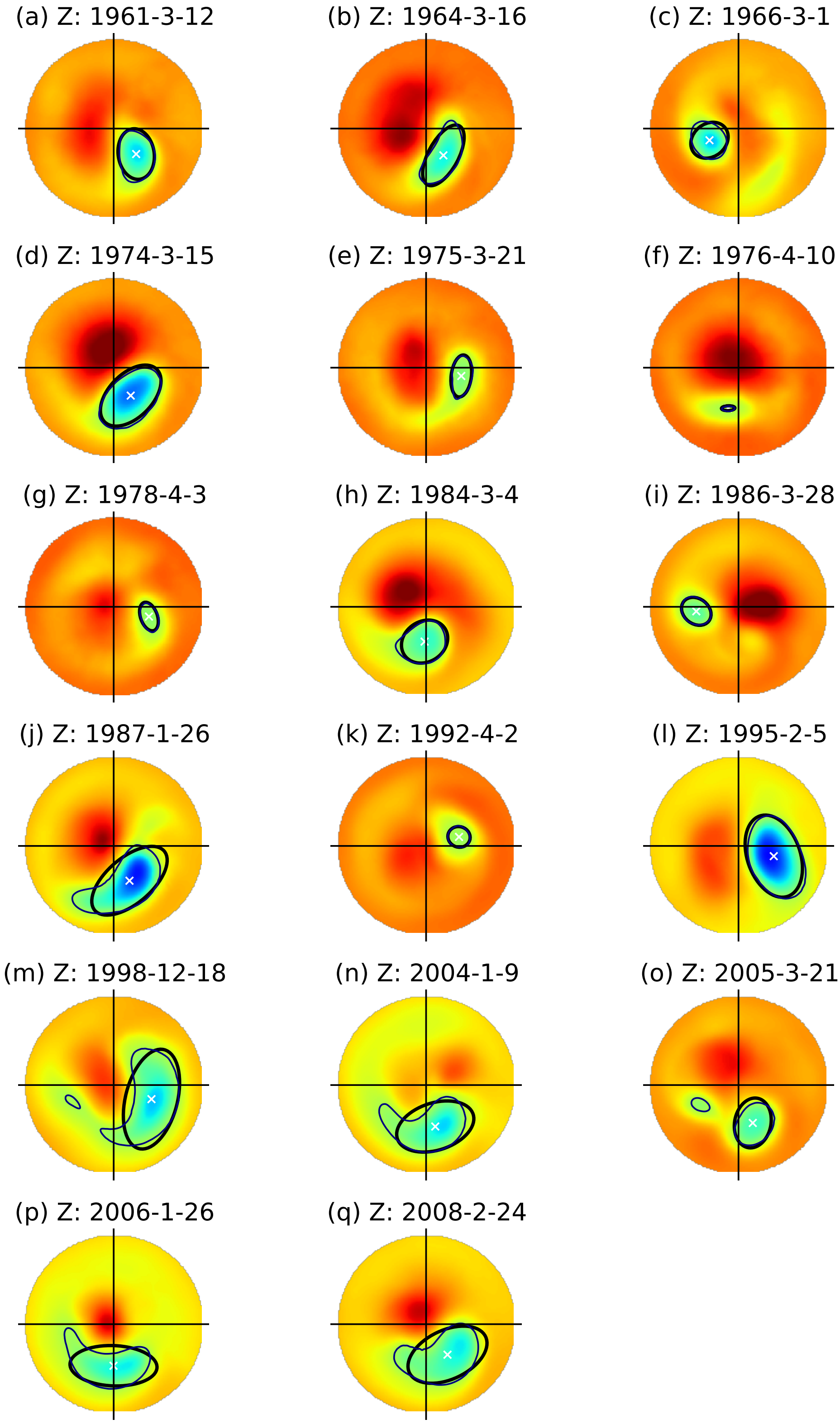

FIGURE 3.7: As Figure 3.6 but for the 17 displaced vortex events identified in ERA.

displaced vortex events respectively. Almost all of the split vortex events show two clearly separated vortices or a pinched vortex shape, which is approximately symmetrical about the North Pole. Two exceptions are Figures 3.6(a) and (k), in which the vortex is highly elliptical but not clearly split. Figure 3.6(n) shows an event with a highly elliptical vortex that is also somewhat displaced from the pole, indicating that it has some displaced nature. The majority of split vortex events are seen to occur along the 90°E-90°W axis, in line with the climatological wave-2 pattern [*Andrews et al.*, 1987]. Figure 3.6(h) shows an exception to this, with an orientation orthogonal to the majority of events.

The displacement events mostly show a smaller and weaker vortex, owing to the fact that they are more common later in winter (see Figure 3.8). Some events, particularly those occurring in late March, are also likely to be events which would traditionally be defined as final warmings. It can be seem that the majority of displacement events occur in the direction of the 0-90°E quadrant, again in line with the climatological wave-1 pattern. However, there are some exceptions to this, for instance Figures 3.7 (c) and (i), which show a westward-displaced vortex.

3.3.4 Comparison with CP07 and M13

The split and displaced vortex events identified using the above method are now compared with those of the CP07 and M13 methods. Table 3.1 identifies those events which do not coincide with the events of CP07 and M13, where 'coincide' indicates events within 10 days and of the same type. Of the 35 events identified, 16 were found not to coincide with events of CP07 (10 displacement and 6 split). Six events were found not to coincide with those of M13 (3 displacement and 3 split), although this comparison only covers the 28 events from 1958-2002, as it was not possible to reproduce the M13 method over the longer period studied here because of the difficulties with hierarchical clustering discussed in Section 3.3.3. Just two completely new events were identified

No.	Event onset	Event type	ΔT_{10} (K)	$\overline{u}_{10}$ (m s^{-1})
1*	1961-3-9	D	10.2	2.7
2*	1962-1-30	S	1.9	38.9
3*	1962-3-7	S	-1.0	16.9
4*	1964-3-15	D	11.9	1.3
5†	1966-2-26	D	2.5	-5.9
6	1967-12-29	S	13.0	19.4
7†	1970-1-5	S	8.5	-4.0
8	1971-1-15	S	10.8	-1.7
9*	1972-2-4	S	-1.6	33.6
10†	1973-2-4	S	7.3	-6.6
11*	1974-3-12	D	5.3	-4.8
12*	1975-3-16	D	7.6	-8.0
13*	1976-3-31	D	8.2	-13.3
14†	1977-1-7	S	7.6	-5.5
15*†	1978-3-25	D	2.5	-9.3
16	1979-2-18	S	5.6	-0.4
17	1984-2-25	D	11.6	-4.4
18	1984-12-25	S	15.0	-1.7
19*	1986-1-7	S	3.4	29.9
20*	1986-3-21	D	9.1	-12.2
21	1987-1-20	D	8.3	-7.7
22	1987-12-10	S	9.8	-3.0
23*	1992-3-22	D	7.6	-4.4
24*†	1995-2-2	D	5.6	7.7
25	1998-12-15	D	8.2	8.1
26	1999-2-24	S	6.6	-12.7
27	2001-2-7	S	5.2	-7.2
28*	2001-3-15	S	-6.8	12.1
29*	2002-3-21	S	-1.5	5.1
30	2003-1-17	S	6.1	16.8
31	2004-1-2	D	5.8	-4.8
32*	2005-3-11	D	3.1	-5.0
33	2006-1-17	D	4.2	-14.3
34	2008-2-18	D	4.6	2.3
35	2009-1-18	S	13.2	16.9

TABLE 3.1: A summary table of displaced (D) and split (S) vortex events, identified from 10 hPa geopotential height data from 1958–2009. ΔT_{10} represents the mean area-weighted 50°-90°N cap temperature anomaly at 10 hPa calculated 5 days either side of the event onset date. $\overline{u}_{10}$ represents $\overline{u}$ at 60°N and 10 hPa averaged over the same period. Asterisks (*) represent those numbers that do not coincide (i.e. within 10 days and of the same type) with events defined by CP07 and daggers (†) events which do not coincide with events of M13.

(i.e. not coinciding with either CP07 or M13); these are the displaced vortex events with onset dates 1978-3-25 and 1995-2-2.

Table 3.1 also shows polar cap averaged 10 hPa temperature anomalies (ΔT_{10}), averaged 5 days either side of the event to give a measure of the event magnitude. The events of CP07 show a larger average anomaly than events identified with the current method, although the two are not statistically significantly different: CP07 average 8.6 K [6.1, 10.9] split and 7.8 K [5.5, 9.9] for displaced vortex events, while the current method averages 5.7 K [3.0, 8.3] for split and 6.8 K [5.5, 8.2] for displaced vortex events (numbers in square brackets represent the 95% uncertainty range, calculated using a bootstrap test). It can be seen that while the vast majority of events show positive values of ΔT_{10} (i.e. warming), four events show negative values. All of these events are also identified by M13, and they attributed the negative values to the presence of a strong, cold vortex prior to the event. Zonal-mean zonal wind at 60°N and 10 hPa ($\overline{u}_{10}$), averaged over the same period is also shown in Table 3.1. The majority of events show negative values, in line with the traditional wind reversal criterion, although some show positive values. Again this results from a strong vortex prior to these events, as well as the fact that the new method detects some events with a distorted but strong vortex (seen in Figures 3.6 and 3.7).

The seasonal distribution of split and displaced vortex events identified by the current method (Z_{10}), M13, and CP07, is shown in Figure 3.8. In all three methods split vortex events are more frequent in early-mid winter, with a peak in January. For displaced vortex events, both the current method and M13 show a skew towards events occurring later in winter. However, there is less similarity with the CP07 distribution of displaced vortex events. CP07 indicates an approximately flat distribution throughout winter, and many fewer displaced vortex events overall. It should be noted that the seasonal distribution of split vortex events from the moment based methods does not arise from the underlying climatology of aspect ratio, which remains approximately

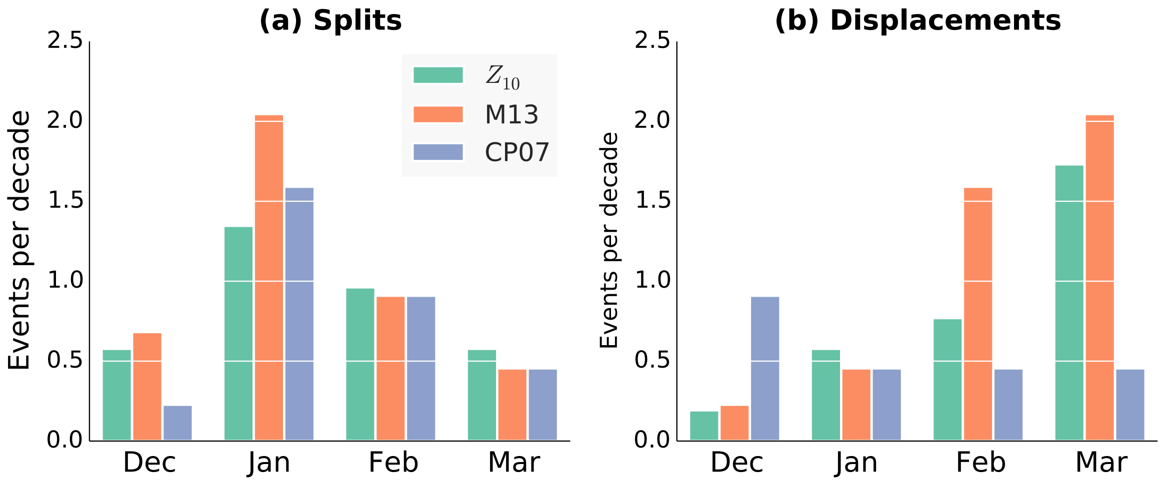

FIGURE 3.8: Histogram of the seasonal distribution of displaced and split vortex events, from the new geopotential height-based method (Z_{10}), M13 and CP07.

constant throughout winter (e.g., Figure 4.2). The centroid latitude does however, show a small equatorwards trend throughout winter, which may to some extent account for the seasonal distribution of displaced vortex events [*Mitchell et al.*, 2011].

Figure 3.9 compares the average shape of the stratospheric polar vortex following the split and displaced vortex events identified by the three methods. Composites of PV in the mid-stratosphere (850 K) are shown averaged 5 days following each event. For the split vortex events, the new method (Z_{10}) method clearly shows two separated vortices, one centred over Canada and the other over Siberia. For the M13 events the split vortex composite shows the vortex stretched across the same 90°W-90°E line, although not as clearly split, while the composite for the CP07 events looks very different. This has a weak vortex centred over Canada, with the other over Northern Europe in a similar location to the composite for displaced events. All three composites for displaced events show a vortex centred over Northern Europe, but this extends most westward in the CP07 composite, suggesting that there may be some contamination from misdiagnosed split vortex events.

Overall, Figure 3.9 demonstrates that the new method succeeds (in a composite sense) in identifying displaced and split vortex events at least as well as the methods

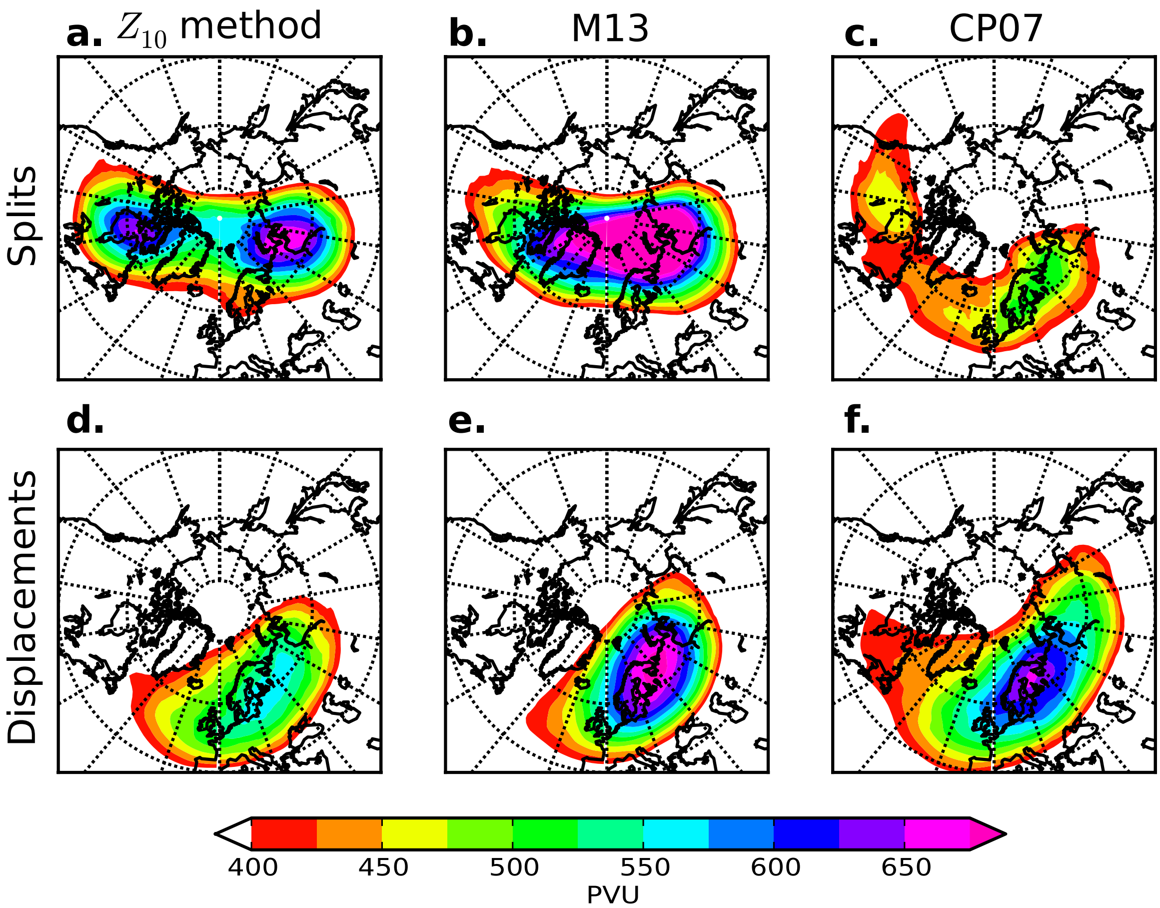

FIGURE 3.9: Composites of potential vorticity at the 850 K isentropic surface from the ERA reanalysis over 1958–2009. Composites are taken over the 5 days following the onset date of split vortex events (a,b,c) and displaced vortex events (d,e,f). The new (Z_{10}) method (a,b) is compared with that of M13 (b,e) and CP07 (c,f).

of M13 or CP07. When comparing the three methods, CP07 is the clear outlier. This is most likely because the CP07 approach employs a zonal-mean threshold which cannot accurately capture some extreme events (as discussed in M13).

3.3.5 Comparison with zonal wave amplitudes

Many studies have characterised stratospheric polar variability by its zonal wave structure [e.g, *Randel*, 1988; *Yoden et al.*, 1999; *Nakagawa and Yamazaki*, 2006; *Bancalá et al.*, 2012]. It is therefore instructive to compare this wave analysis with the moment diagnostics developed above. Here this is carried out for the displaced and split vortex events shown in Figures 3.1 and 3.2 respectively. A similar comparison was shown by *Waugh* [1997] and *Waugh and Randel* [1999], and the results here are consistent with their findings.

Zonal wavenumber decomposition is carried out by taking the Fourier transform of the 60°N, 10 hPa geopotential height field over all longitudes. The amplitude of wave-n on a given day is then given by the modulus of the nth Fourier component on that day. In ERA data, the amplitude of DJFM wave-2 is, on average, about 30% that of wave-1, and wave-3 13% of wave-1, indicating a Charney-Drazin filtering of zonal wavenumbers, as discussed in Section 2.1.2.1.

The correlation of geopotential height-derived daily aspect ratio and the wave-2 amplitude over DJFM is 0.30 and that of centroid latitude and wave-1 amplitude −0.22, both of which are statistically significant at the 99% level. Figure 3.10 illustrates these relationships for the example split and displaced vortex events previously shown in Figures 3.1 and 3.2 . In the case of the 1979 split vortex event the wave-2 amplitude peaks approximately 5 days before the peak of the aspect ratio. Wave-2 amplitude is also more variable than aspect ratio in early winter, although the two are correlated at this time. In the case of the 2006 displaced vortex event, the difference is even greater. The wave-1 amplitude peaks about three weeks before the centroid latitude, and is actually anomalously small at the peak of centroid latitude. Before and after the event, the wave-1 amplitude and centroid latitude are highly correlated.

A result of these differences is that not all split vortex events are defined as wave-2 warmings and not all displaced vortex events as wave-1 warmings. For example, the split vortex events with onset dates 1973-2-7, 1987-12-10, and 1999-2-24 are classified as wave-1 warmings by *Bancalá et al.* [2012]. However, the structure of the vortex appears clearly split in these three examples (see Figures 3.6 (g), (l) and (m)), highlighting the differences in these classifications.

The physical reasons for these differences are investigated in Figure 3.11. This compares the vortex structure at the peak wave amplitude and peak aspect ratio/centroid latitude for the two events. For the 1979 split vortex event, the vortex appears split at both times but there is a greater separation between the vortices at the time of maximum aspect ratio. The wave-2 amplitude is greatest at the earlier time, however, because the vortices (particularly the eastern vortex) are stronger. Similarly for the 2006 event, at the time of maximum wave-1 amplitude the vortex is not displaced far from its average position but its strength means that the wave-1 amplitude is greater than the much more displaced vortex found later. Generally, the sensitivity of wave amplitudes to vortex strength means that wave amplitudes may actually decline during periods of intense wave breaking due to the weakening of the vortex, even if that vortex is significantly distorted.

Overall, these results show that as the vortex departs from zonal symmetry linear wave theory breaks down and changes in the wave-1 and wave-2 amplitudes cannot be simply interpreted as changes in the position and elongation of the vortex respectively.

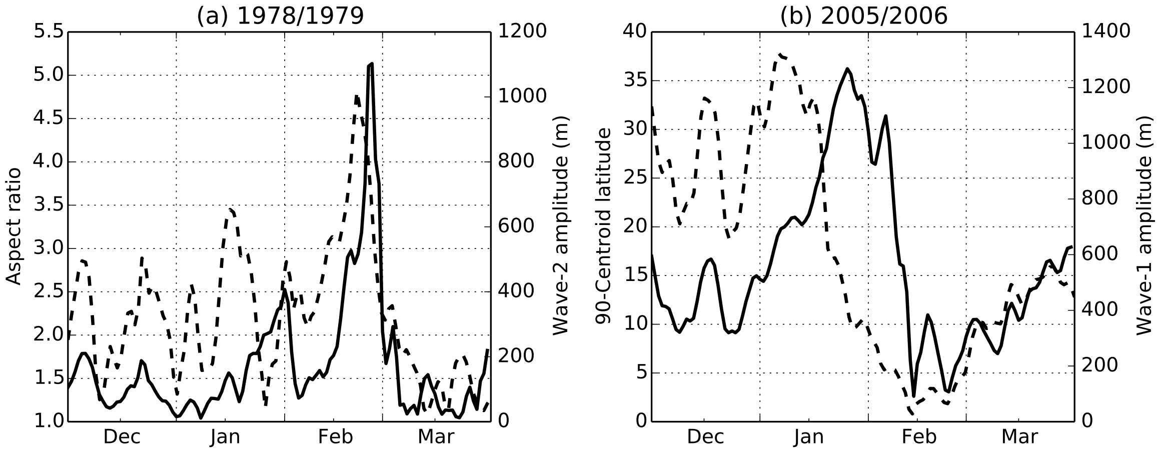

FIGURE 3.10: (a) Aspect ratio (solid line) and wave-2 amplitude (dashed line) over the winter 1978-1979. (b) Centroid latitude (solid line) and wave-1 amplitude (dashed line) over the winter 2005-2006. Centroid latitude is expressed as its deviation from the North Pole.

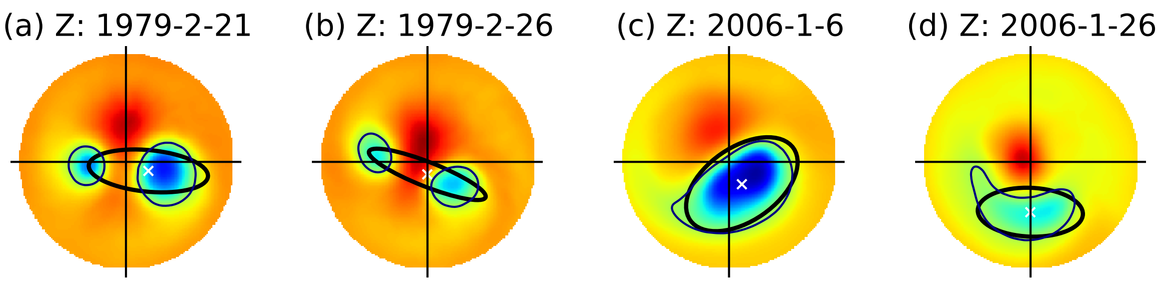

FIGURE 3.11: 10 hPa geopotential height on the day of maximum wave-2 amplitude (a) and maximum aspect ratio (b) for the 1979 split vortex event; and maximum wave-1 amplitude (c) and minimum centroid latitude (d) for the 2006 displaced vortex event. The vortex edge contour, equivalent ellipse, and centroid latitude are shown as Figure 3.1.

3.4 Stratosphere-troposphere coupling

3.4.1 Tropospheric response

Having verified that this new method identifies split and displaced vortex events as skillfully as previous methods, it is now possible to study their influence on the troposphere. This is motivated by the result of M13 who, as discussed in Section 2.4.1, found tropospheric anomalies to be larger following split vortex events that displaced vortex events. Figure 3.12(a,b) shows time-height composites of the NAM over the 90 days following split and displaced vortex events. Here the method of *Baldwin and Thompson* [2009] is used to define the NAM as the leading empirical orthogonal function (EOF) of daily wintertime (November-April) zonal mean geopotential height anomalies poleward of 20°N. The anomalies are calculated by subtracting the seasonal cycle which has been smoothed with a 90-day low-pass filter. The daily NAM anomalies are then determined by projecting daily geopotential anomalies onto the leading EOF patterns. Finally, the NAM is normalised so that the time series at each level has unit variance.

In agreement with M13, it can be seen that the tropospheric NAM is more negative during the 60 days following split vortex events than displaced vortex events. Also similar to M13 is the fact the vertical evolution for the two events greatly differs, with split vortex events occurring almost instantaneously throughout the depth of the atmosphere and displaced vortex taking almost two weeks to propagate through the stratosphere. The near-barotropic nature of split vortex events suggests that resonant excitation of the barotropic normal mode [*Esler and Scott*, 2005] may be an important influence in this case.

The difference in the NAM composites (split minus displaced) is shown in Figure 3.12(c). Statistical significance of this difference is calculated with the null hypothesis that there is no difference between the NAM response to split and displaced vortex events, and assessed using a two tailed bootstrap test with the following procedure:

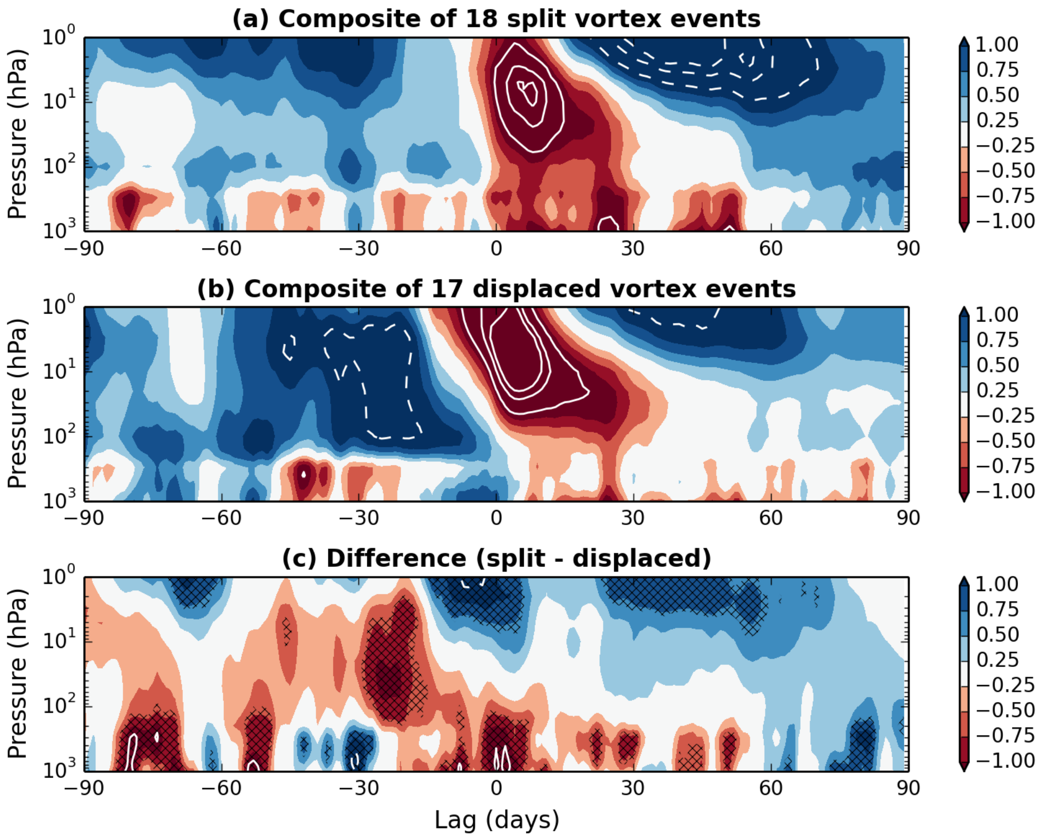

FIGURE 3.12: Composites of the time-height evolution of the NAM during (a) 17 vortex displacement events and (b) 18 splitting events. (c) shows the difference in these composites, and hashed regions represent those that are 95% significant according to a two-tailed bootstrap test. Lag 0 is at the onset of an event as measured at 10 hPa. Contour intervals are 0.25 and the region between -0.25 and 0.25 is unshaded. Data is from the ECMWF Reanalyses 1958–2009.

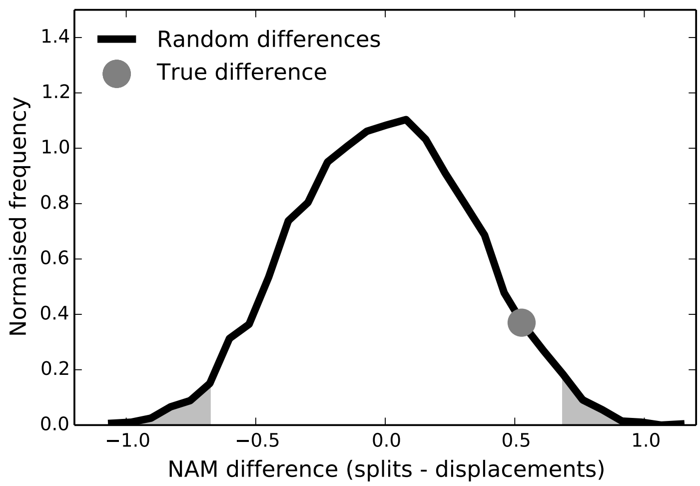

FIGURE 3.13: Distribution of 0-30 day mean surface NAM composite differences between split and displaced vortex events, formed by randomly shuffling the labels 'split' and 'displacement' between events. The 95% significant region (according to a two-tailed test; i.e. $< 2.5\%$ and $> 97.5\%$) is shaded and the true composite difference is at the 94th percentile.

i. The labels 'split' and 'displacement' are randomly re-assigned to the 35 events.

ii. NAM composites and the composite difference of these randomly assigned events are calculated.

iii. The above steps are repeated 10 000 times, to form a distribution of random composite differences. If the true composite difference lies $< 2.5\%$ or $> 97.5\%$ within this distribution, then it can be said to be 95% significant.

Some significant differences are seen between the split and displaced vortex composites. For instance, a more positive stratospheric NAM is seen to precede displaced vortex events, while the dipole in the upper stratospheric and tropospheric NAM near lag 0 represents the difference in baroclinicity of the two types of event. Some regions of significant differences are seen in the tropospheric NAM 0-60 days after the event, but there are also some regions that are not significant. Care must be taken when interpreting the importance of small significant regions these may arise by chance, even if no physical relationship exists.

The difference in the tropospheric anomalies following split and displaced vortex events can be tested more robustly by examining surface anomalies averaged over the 30 days following onset. This difference is again tested using the bootstrap procedure outlined above. The distribution of randomly calculated surface NAM composite differences, and the actual surface NAM composite difference are shown in Figure 3.13. It can be seen that the true NAM difference does not lie in the 95% significant region, so the null hypothesis that there is no difference between surface NAM anomalies following split and displaced vortex events cannot be rejected. It should be noted that the statistical test here is different to that carried out by M13. They tested whether the surface NAM following split and displaced vortex events were different from randomly selected winter dates, finding that anomalies following splits are, but those following displacements are not. They did not, however, test the *difference* between split and displaced vortex events.

The surface NAM does not provide the full description of surface variability, and so in Figure 3.14 composites of MSLP 30 days before and 30 days following the onset dates of displaced and split vortex events are presented. Statistical significance is calculated against the null hypothesis that anomalies before and after split and displaced vortex events are indistinguishable from other winter dates. This is again estimated from a two-tailed bootstrap test, in which 10 000 composites of equal size are formed from randomly selected winter dates, and the percentile of the true composite calculated from this distribution.

The strongest precursor is found for displaced vortex events, with a wave-1 pattern that is similar to the climatological stationary wave pattern [e.g., *Garfinkel and Hartmann*, 2008], suggesting increased wave-1 propagation into the stratosphere. However, the strongest anomalies following events occur after split vortex events, with a pattern resembling the negative phase of the NAM, though with a southern centre of action shifted towards Europe. A further difference between the split and displaced vortex composites is that there is a more negative MSLP anomaly over Scandinavia and Siberia

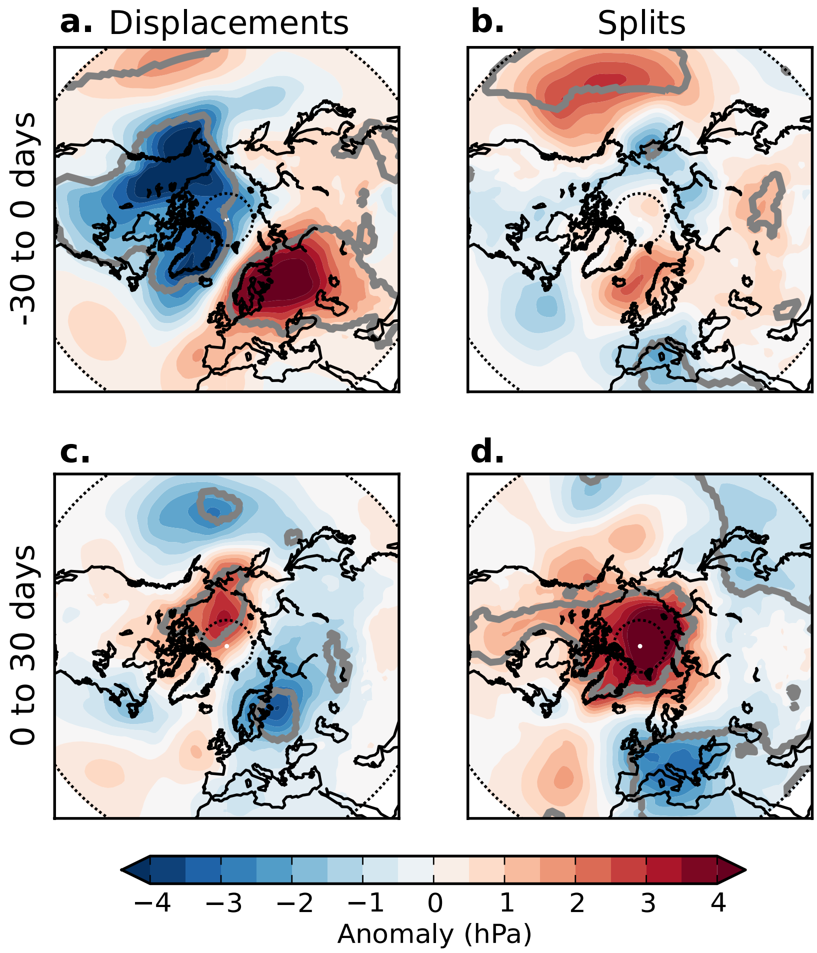

FIGURE 3.14: Composites of mean sea-level pressure anomalies in the 30 days before (a,b) and 30 days after (c,d) the onset dates of displaced (a,c) and split (b,d) vortex events from the Z_{10} method. Data are from the ECMWF renalyses (1958–2009). Anomalies are calculated for each day and gridpoint from the climatology for that day of the year and gridpoint. Grey contours indicate regions of greater than 95% statistical significance according to a bootstrap significance test.

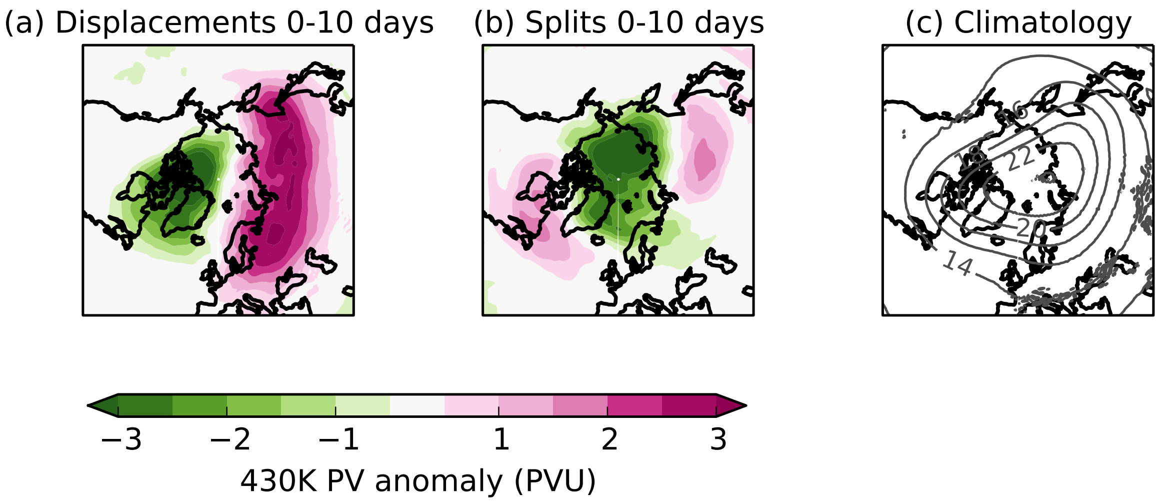

FIGURE 3.15: Composite of PV anomalies on the 430 K isentropic surface averaged over the 10 days following displaced (a) and split (b) vortex events. DJFM average of 430 K PV (c). Units are PVU and the contour interval is 2 PVU. Data are restricted to the ERA-Interim period (1979–2009), meaning a total of 10 displaced and 10 split vortex events enter the composites.

following displaced vortex event. Overall, the main features of Figure 3.14 compare very well with the corresponding diagnostics from M13.

3.4.2 Tropopause response

The mechanism of the stratosphere's influence on the troposphere proposed by *Ambaum and Hoskins* [2002] states that changes in the PV near the tropopause affect the tropopause height and induce tropospheric anomalies below (more details are given in Section 2.4.3). In order to investigate this mechanism, composites of PV anomalies at the 430 K isentropic surface (which lies close to 100 hPa; just above the tropopause), over the 10 days following displaced and split vortex events are shown in Figures 3.15(a,b). Data at this isentropic level were not available for ERA-40 so these composites are limited to the ERA-Interim (1979-2009) period, meaning 10 events of each type enter the composites. The shorter 10-day period was chosen to reflect the typical time scale for the split or displacement of the vortex, rather than the longer time scale taken for the re-formation of the vortex. However, composites taken over 30 days, as in Figure 3.14, show similar structure but with reduced magnitude (similarly, composites taken

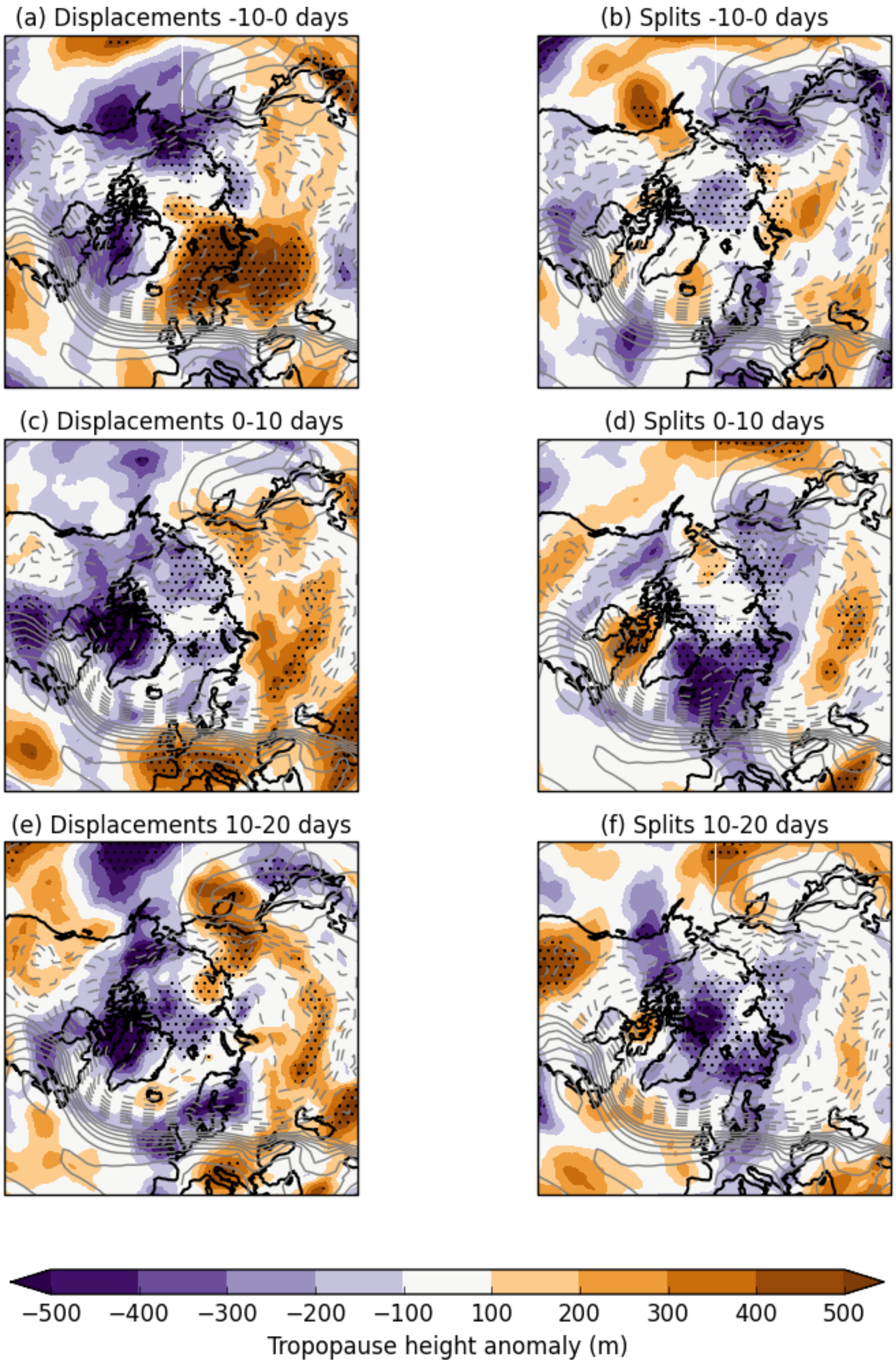

FIGURE 3.16: Composites of tropopause height anomalies averaged 10 days before (a,b), 10 days after (c,d) and 10-20 days after displaced and split vortex events (filled contours). Anomalies are calculated for each day and gridpoint from the climatology for that day of the year and gridpoint. Stippling indicates regions of greater than 95% statistical significance according to a Monte-Carlo significance test. Grey contours indicate the first EOF of NH mean sea-level pressure, which explains 33% of the variance (dashes represent negative values).

over 5 days, as in Figure 3.9, show slightly increased magnitude). In the displaced composite case a region of high PV is seen over Siberia and Scandinavia, consistent with the movement of the vortex over this region. Note that this is shifted further east than the position of the vortex at 850 K ($\sim$ 10 hPa) (Figure 3.9(d)), again indicating the more baroclinic nature of displaced vortex events (this westward tilt with height was also found by *Matthewman et al.* [2009]). The split vortex composite shows two regions of raised PV which are approximately co-located with the two vortices at 850 K.

Again following the reasoning of *Ambaum and Hoskins* [2002], composites of tropopause height averaged over the 10 days before, 0-10 days after, and 10-20 days after split and displaced vortex events are now shown in Figure 3.16 (these composites now use the full ERA (1958-2009) data set). The measure of tropopause height used is that of *Wilcox et al.* [2012], who construct a blended thermal and dynamical tropopause. Significance is again calculated using a two-tailed bootstrap test.

In line with the MSLP anomalies shown in Figure 3.14, tropopause height anomalies are seen to be larger prior to displaced vortex events, with a wave-1-like structure. Following the events, tropopause height anomalies are seen to approximately mirror the stratospheric PV anomalies (Figure 3.15). That is, following displaced vortex events an elevated tropopause is seen over Europe and Scandinavia, with a lowered tropopause over Canada, and following split vortex events two regions of elevated tropopause are present over Canada and Siberia with a depression in between.

It is possible to quantitatively examine (although only approximately) whether these tropopause anomalies are consistent with the changes in stratospheric PV above. Changes in tropopause pressure, Δp_{trop}, are related to changes in stratospheric PV, Δq, through

$$\Delta q \approx -q(1+\text{Bu})\frac{\Delta p_{\text{trop}}}{p_{\text{trop}}}, \tag{3.14}$$

where Bu is the Burger number, which is approximately equal to one for any PV anomaly [*Ambaum and Hoskins*, 2002]. The change in tropopause height, Δh_{trop} can be calculated using the hydrostatic relation

$$\Delta h_{\text{trop}} = -\frac{\Delta p_{\text{trop}}}{p_{\text{trop}}} \frac{RT_{\text{trop}}}{g}, \tag{3.15}$$

where T_{trop} is the tropopause temperature. Hence

$$\Delta h_{\text{trop}} = \frac{\Delta q}{q} \frac{RT_{\text{trop}}}{g(1+\text{Bu})}. \tag{3.16}$$

From Figures 3.15(a,b) a typical 430 K PV anomaly is 2 PVU, and the background climatology is approximately 20 PVU (Figure 3.15(c)), so $\Delta q/q \approx 0.1$. With a typical value of $T_{\text{trop}} = 210$ K, this then gives a change of tropopause height of $\Delta h_{\text{trop}} \approx 300$ m, which is indeed approximately in line with the tropopause height anomalies seen in Figure 3.16. This, along with the fact that the pattern in tropopause height anomalies approximately mirrors that of stratospheric PV anomalies, suggests that these tropopause height anomalies are induced by changes in stratospheric PV above.

Also shown in Figure 3.16 is the surface NAM pattern (the leading EOF of DJFM daily MSLP). It can be seen that following split vortex events more than displaced (especially days 0-10), the negative tropopause height north of Iceland aligns more closely with the minimum in the NAM (this region is also a node of the NAO). This may be significant if it is expected that the fluctuation-dissipation theorem (FDT) [*Nyquist*, 1928] holds in the tropospheric response to stratospheric forcing. For systems in which the FDT holds (which relies on a small applied forcing), the response of a system projected on a mode of variability should linearly scale with the projection of the forcing on that mode [*Ring and Plumb*, 2008]. Under the assumption that the tropopause height perturbation represents the "forcing", appears to project more strongly on the NAM/NAO following split vortex events, consistent with a greater surface response to these events. However, the pattern correlations between the split and displaced vortex tropopause height anomalies and the NAM are not statistically significantly different because the tropopause height field is very noisy. In order to give a more detailed analysis a greater number of events would be needed.

3.5 Conclusions

Recent research has demonstrated the need to distinguish between split and displaced stratospheric polar vortex events because of their different dynamics and impacts on the troposphere. However, previous methods to identify these events are impractical for application to climate model or seasonal prediction simulations because they are highly sensitive to model climatology or rely on non-standard variables. Motivated by this, we have developed a new method to identify displaced and split vortex events which requires only geopotential height at 10 hPa. The method is summarised as follows:

i. To identify the vortex region, a single contour of 10 hPa geopotential height is selected. This is the value of the DJFM mean zonal-mean at 60°N.

ii. Using this contour the centroid latitude and aspect ratio moment diagnostics can be calculated.

iii. Events are identified using a threshold criterion: Displaced events are said to occur if the centroid latitude remains equatorward 66°N for 7 days or more. Split events are said to occur if the aspect ratio remains above 2.4 for 7 days or more. In order to ensure that events are not counted twice, no two events may occur within 30 days.

Results show that vortex moment diagnostics derived from geopotential height in this way are highly correlated with those derived from PV, although fewer high aspect ratio values are seen. The use of geopotential height here is motivated by the fact that it is commonly output by climate models, whereas PV is not. However, in cases where PV is available (such as in reanalyses) its use is preferable because of its quasi-conservative properties and smaller-scale features. The above method can be easily adapted for use with PV-based vortex moments.

Analysis of the stratosphere following events identified by this method demonstrates that it is able to accurately identify split and displaced vortex events. Most of the events identified coincide with those of M13, and about half with events identified by CP07. Composite analysis indicates that the position of the stratospheric polar vortex following these events is at least as extreme as that from the previous methods.

Having identified these events, their impact on the troposphere has been investigated. Composites of the NAM indicate a more negative surface NAM over the month following split vortex events than following displaced vortex events. This supports the finding of M13, using a different event identification method and extended data set. However, using a bootstrap test the composite *difference* of the surface NAM is not found to be statistically significant.

Anomalies of tropopause height following split and displaced vortex events are found to be co-located with lower-stratospheric PV anomalies. They are also of a magnitude consistent with being induced by changes in the stratospheric polar vortex. Surface anomalies induced by changes in tropopause height may therefore explain the different surface anomalies following split and displaced vortex events. However, it is not possible to draw firm conclusions on this because of the relatively small number of events and the noise of the MSLP and tropopause height fields.

Overall, statistically significant results about the difference in the tropospheric response to split and displaced vortex events will require a larger number of events. This is achieved through the analysis of climate model simulations in the next chapter.

CHAPTER 4

Representation of vortex variability in climate models

4.1 Introduction

Over the past decade an increasing number of climate models have included a well-resolved stratosphere, with model lids above the stratopause. This trend is evident in the Coupled Model Intercomparison Projects, CMIP3 and CMIP5 [*Cordero and Forster*, 2006; *Taylor et al.*, 2012], which were evaluated in the Intergovernmental Panel on Climate Change (IPCC) fourth and fifth Assessment Reports respectively [*Solomon et al.*, 2007; *Stocker et al.*, 2013]. CMIP5 includes 15 models with an uppermost level above the stratopause, whereas CMIP3 includes only five. This change in models' representation of the stratosphere has been largely motivated by an increased understanding of the stratosphere's influence on tropospheric climate (discussed in *Gerber et al.* [2012] and Chapters 2 and 3).

The effect of this greater stratospheric resolution was studied by *Charlton-Perez et al.* [2013], who compared stratospheric variability between high-top and low-top models within the CMIP5 ensemble (they defined "high-top" as a model lid above 1 hPa, and "low-top" below). They found that the low-top models have a weaker and less realistic representation of daily to interannual polar stratospheric variability than high-top models, and attributed this to the fact that the low-top models simulate fewer SSW events than high-top. This is combined with a slightly weaker tropospheric NAM response in the two months following SSW events in the low-top compared to high-top models. *Shaw et al.* [2014] also found that high-top CMIP5 models have a greater frequency of extreme stratospheric planetary wave heat flux events than low-top models. They argue that this results in greater biases in the position of the Atlantic jet stream position in low-top models.

These results are supported by similar studies which compared natural variability in high and low-top versions of the same model. *Cagnazzo and Manzini* [2009] found that a high-top model gave a more realistic representation of the influence of ENSO on the NH extratropical stratosphere. Similarly, *Hardiman et al.* [2012] showed the influence of the QBO on the tropical troposphere well as decadal trends in the NAO were more realistically simulated by the high-top than the low-top model. Both *Sassi et al.* [2010] and *Hardiman et al.* [2012] also found a more realistic frequency of SSWs and greater impact on the troposphere in a high-top model.

Other studies have compared simulations of climate change with high- and low-top models. *Huebener et al.* [2007] linked an increased weakening of the stratospheric polar vortex in high-top simulations to a more southward shift of the NH winter storm track, which in turn affects trends in North Atlantic temperatures and precipitation. *Manzini et al.* [2014] investigated climate change simulations of high- and low-top models in the CMIP5 ensemble. They found that the inter-model spread in the simulation of changes of stratospheric polar vortex winds accounts for a significant fraction of the inter-model spread of trends in the surface NAM under climate change.

Despite these findings about the differences between high- and low-top models, it is important to note that a model lid above the stratopause is not a sufficient condition for the accurate representation of stratospheric processes or stratosphere-troposphere coupling. Several other factors are important, such as the parametrisation of gravity waves, resolution of steep PV gradients which impact upon planetary wave propagation, and the generation of planetary waves in the troposphere. Indeed, *Shaw et al.* [2014] found a wide range of biases in stratospheric heat flux extremes among hig-top models and *Charlton-Perez et al.* [2013] found that the frequency of SSWs in high-top CMIP5 models varies widely, from about 2.5 to 8 events per decade.

In this chapter we apply the methods developed in Chapter 3 to evaluate the representation of stratospheric polar vortex variability in the CMIP5 climate models. Motivated by these results which demonstrate a more realistic representation of tropospheric and stratospheric climate in high-top models, we select only models with a lid height above the stratopause. In doing this we extend the work of *Charlton-Perez et al.* [2013] to consider the two-dimensional structure of the polar vortex using moment diagnostics, including the identification of split and displaced vortex events.

The only previous study to apply vortex moment diagnostics to climate model simulations is that of *Mitchell et al.* [2012]. They studied models from the second Chemistry-Climate Model Validation (CCMVal-2) project, although their analysis was limited because only three models of the 18 in CCMVal-2 provided the daily PV which was necessary for the calculation of moment diagnostics. They also did not classify split and displaced vortex events in their analysis, instead focussing on the mean state of the vortex. Using the new methods developed in Chapter 3, we are now able to calculate moment diagnostics and classify split and displaced vortex events using geopotential height from a much larger number of models.

There are three main objectives to this investigation. First, we wish to evaluate the current state of models' representation of the stratospheric polar vortex and stratosphere-troposphere coupling, including whether there are any consistent biases among models. Second, we aim to determine if there is a relationship between model parameters (such as horizontal and vertical resolution) and biases in their representation of vortex variability. This may motivate future model improvements to reduce these biases. Third, we will investigate whether the increased sample size of the CMIP5 ensemble can be used to better understand the mechanism behind the different tropospheric response to split and displaced vortex events, which was described in Chapter 3.

4.1.1 CMIP5 model simulations

For this analysis only climate models with a lid height above the stratopause are selected from the CMIP5 ensemble. In total, 13 such models (listed in Table 4.1) were available from 8 different modelling centres. Although another two (CESM1-WACCM and MIROC-ESM) are listed in the CMIP5 ensemble, appropriate data was not found to be available for these models in the CMIP5 archive (`http://pcmdi3.llnl.gov/esgcet/home.htm`). It can be seen that 12 of the 13 models have an uppermost level which is in the upper mesosphere (70-80 km), but CanESM2 has a significantly lower lid which lies close to the stratopause.

Historical simulations have been used throughout this analysis. These include observed climate forcings, such as from greenhouse gases, ozone depletion, land-use change, tropospheric and stratospheric aerosols and solar variability. The simulation period considered is limited to 1958-2005, so that it coincides with the ERA-40/ERA-Interim reanalysis period (CMIP5 historical simulations end at 2005). Limiting the model simulation analysis to the same period as reanalysis may be important because several studies have suggested that external forcing, such as volcanic eruptions and solar variability, has a significant impact on stratospheric variability [e.g., *Robock*, 2000; *Gray et al.*, 2010]. In order to achieve the largest possible ensemble size, all available ensemble members have been used for each model, which leads to different numbers of years entering the ensemble from different models. This does, however, necessitate that any results appearing in the ensemble mean should also be checked for consistency among the models to ensure that it is not biased by a particular model.

Model	Ensemble size	Lid/ km	Levels	dh/km	dz_1/km	dz_2/km
CanESM2	5	48.1	35	268	1.48	2.30
CMCC-CESM	1	80.6	39	536	1.49	1.89
CMCC-CMS	1	80.6	95	268	0.65	0.68
GFDL-CM3	5	76.3	48	191	1.32	1.75
HadGEM2-CC	3	84.1	60	144	0.82	1.18
IPSL-CM5A-LR	5	70.4	39	254	1.21	1.75
IPSL-CM5A-MR	3	70.4	39	169	1.21	1.75
IPSL-CM5B-LR	1	70.4	39	254	1.21	1.75
MIROC-ESM-CHEM	1	87.8	80	399	0.77	0.73
MPI-ESM-LR	3	80.6	47	268	0.87	1.70
MPI-ESM-MR	3	80.6	95	268	0.65	0.68
MRI-CGCM3	1	80.6	48	107	0.88	1.87
MRI-ESM1	1	80.6	48	107	0.88	1.87

TABLE 4.1: Parameters of the CMIP5 models studied in this chapter. Where the model lid is defined in terms of a pressure, its height was estimated using $z = -H\ln(p/p_0)$ with $H = 7$ km and $p_0 = 1000$ hPa. Following *Anstey et al.* [2013], horizontal resolution, dh, is estimated at 45°N and vertical resolution is shown averaged over two regions; 5-15 km (dz_1) and 15-30 km (dz_2).

4.2 Vortex mean state and variability

4.2.1 Moment diagnostics

The centroid latitude and aspect ratio moment diagnostics are calculated for each of the CMIP5 models over DJFM from the 10 hPa geopotential height field, using the method described in Section 3.3.2. For each model the value of the DJFM mean geopotential height at 60°N and 10 hPa is used to define the appropriate contour for the calculation of the moment diagnostics. This accounts for biases in the mean geopotential height between different models.

The resulting joint distributions of daily centroid latitude and aspect ratio from each of the models are shown in Figure 4.1, along with that from the ERA-40/ERA-Interim reanalysis (hereafter ERA) calculated in Chapter 3. For each model the joint distribution histogram is plotted with a logarithmic colour scale which is normalised according to the number of days entering each box. As discussed in Chapter 3, it can be seen that the joint distribution for ERA has an approximately triangular distribution with high

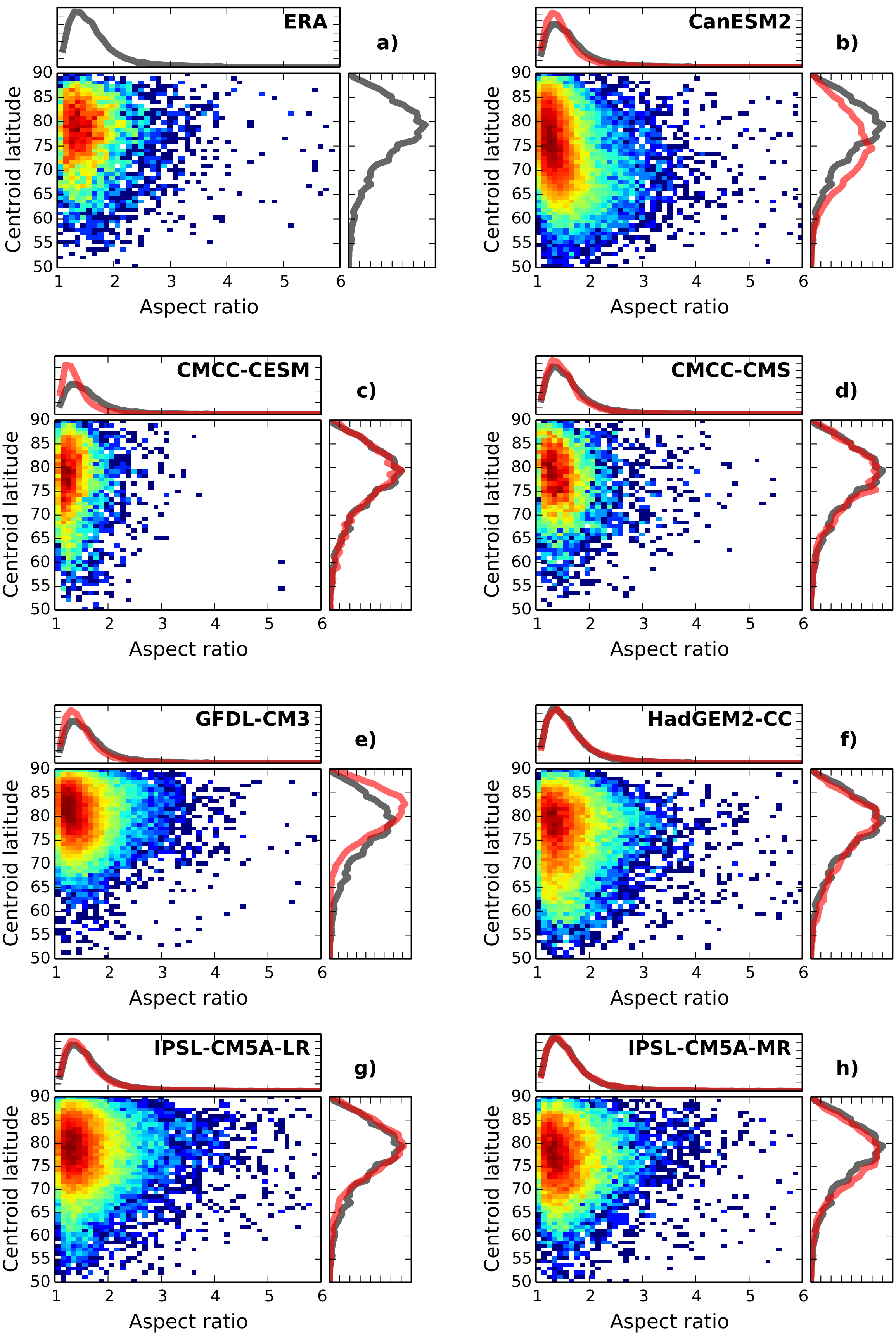

FIGURE 4.1: Distributions of centroid latitude (y-axis) and aspect ratio (x-axis) for the ERA (grey lines; panel (a)) and the CMIP5 models (red lines). Joint distributions are shown with a logarithmic scale such that red squares represent the densest regions.

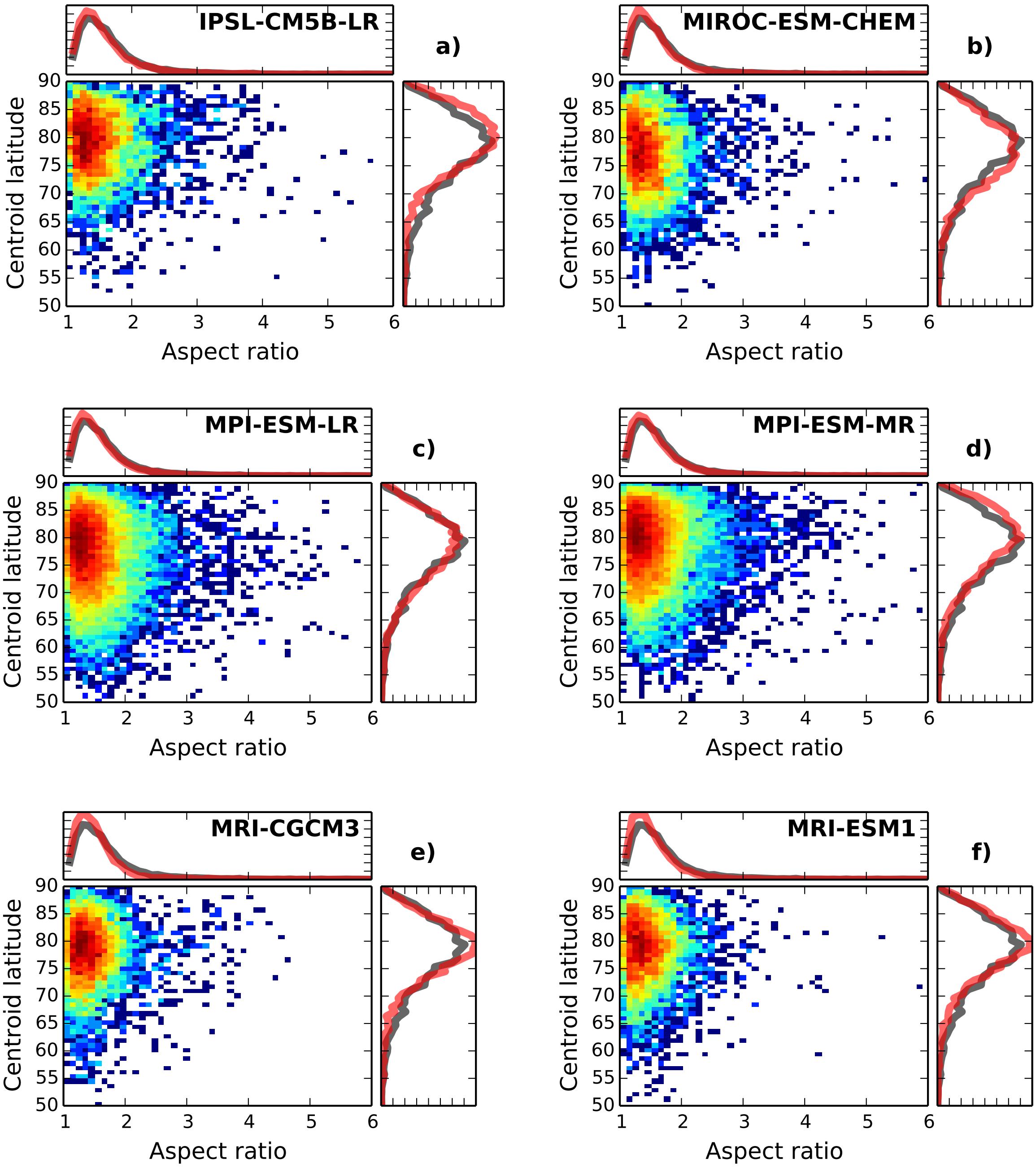

FIGURE 4.1: (Continued)

aspect ratio/poleward centroid latitude, and low aspect ratio/equatorward centroid latitude being relatively more common than high aspect ratio/equatorward centroid latitude. This shape of distribution is well replicated by most of the models, although CanESM2 has a significantly different shape, with the high aspect ratio/equatorward centroid latitude being more common.

It can be seen from this analysis that there are a range of biases among models. CanESM2 has a modal centroid latitude which is about 5° too far equatorward compared to ERA. Contrastingly, GFDL-CM3 has a modal centroid latitude about 2.5° more poleward than observed. CMCC-CESM displays a clear bias in the aspect ratio, with a distribution much less skewed towards high values than in reanalysis.

The winter seasonal cycle of aspect ratio and centroid a latitude in the CMIP5 models is shown in Figure 4.2. For the mean aspect ratio and centroid latitude, the majority of models agree well with reanalysis. CMCC-CESM has a consistently too low mean aspect ratio, while GFDL-CM3 has a consistently too poleward mean centroid latitude, indicating that these biases are not strongly seasonally dependent. On the other hand, the large equatorward bias in the CanESM2 mean centroid latitude is much larger in December and early January than later in winter. The 95th percentile of aspect ratio is lower than reanalysis for the majority of models throughout the season, indicating that models have, on average, too little variability in their aspect ratio.

4.2.2 Displaced and split vortex events

Displaced and split vortex events are identified within the CMIP5 ensemble using the threshold-based method developed in Section 3.3.3. The same thresholds as used for ERA (66°N for centroid latitude and 2.4 for aspect ratio) are used for the models in order to identify, as much as possible, geometrically equivalent events. The same persistence of 7 days was also used. The frequency of displaced and split vortex events for each model is shown in Figure 4.3.

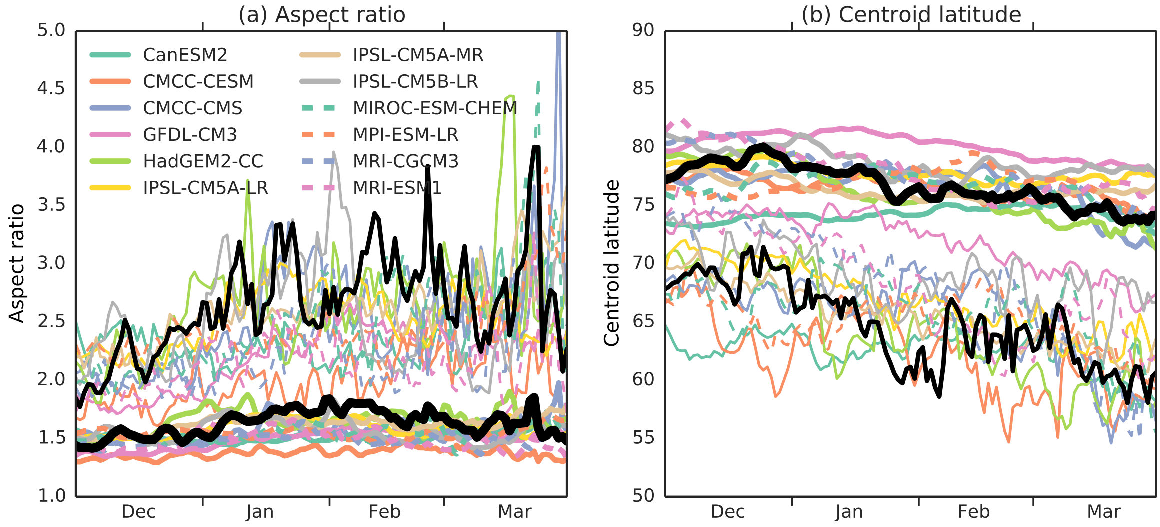

FIGURE 4.2: Seasonal cycle of aspect ratio and centroid latitude in ERA (black) and the CMIP5 models (colours). Thick lines represent the mean and thin lines the 95th or 5th percentile for aspect ratio and centroid latitude respectively.

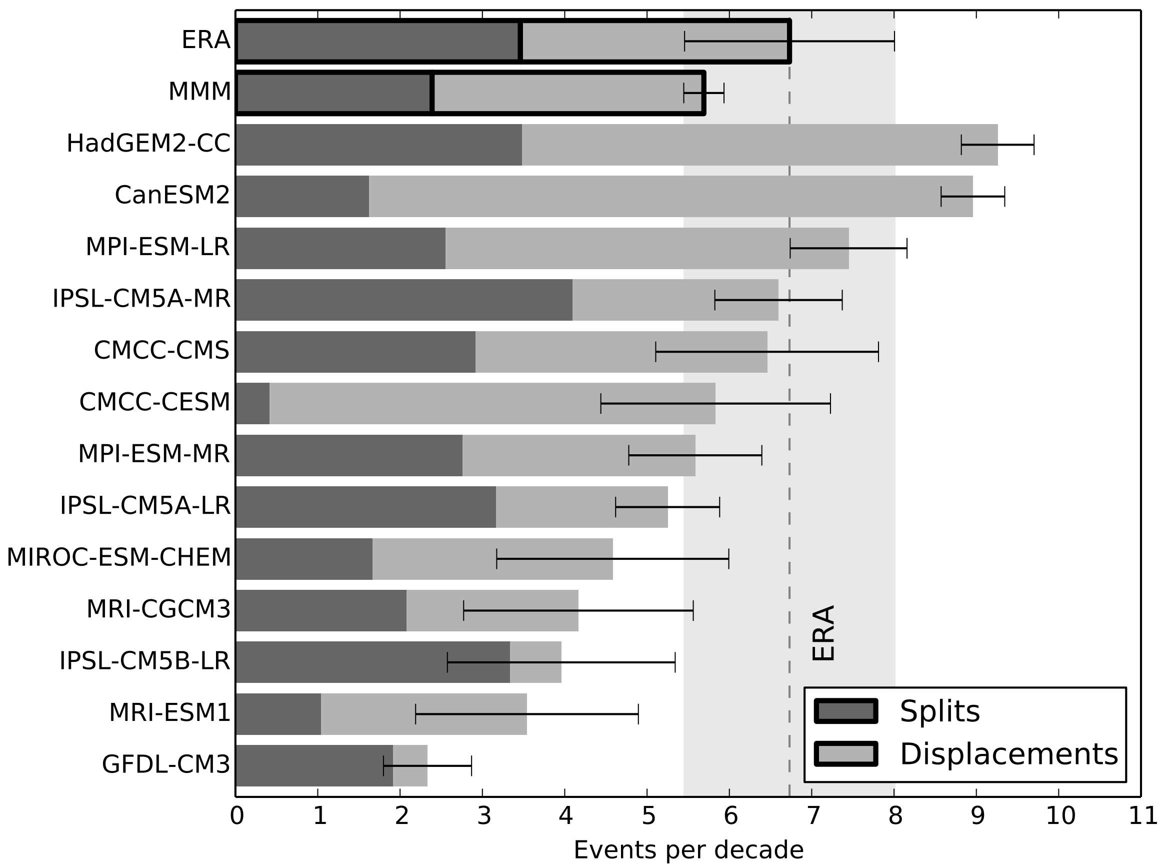

FIGURE 4.3: Frequency of split and displaced vortex events in the CMIP5 models, ERA, and the multi-model mean (MMM). Error bars are for the frequency of all events, and represent one σ range, assuming a binomial distribution of events. The grey shaded region represents the one σ range for ERA, along with the mean (dashed line.)

The total frequency of displaced and split vortex events for each of the CMIP5 models agrees well with the equivalent SSW frequency calculated by *Charlton-Perez et al.* [2013], who identified events based on the reversal of zonal-mean zonal wind at 60°N and 10 hPa. They also found HadGEM2-CC to have the highest frequency of events within the CMIP5 ensemble, while MRI-CGCM3 is the high-top model with the lowest frequency of SSWs in their study (excluding GFDL-CM3 and MRI-ESM1, which *Charlton-Perez et al.* [2013] did not analyse, from the comparison, MRI-CGCM3 becomes the second-lowest frequency in the present study). This similarity between *Charlton-Perez et al.* [2013] and the present study indicates that the close relationship between moment diagnostics-defined events and SSWs defined by zonal-mean zonal wind, as described in Chapter 3, also holds for climate models.

As well as the large differences in the total frequency of displaced and split vortex events, it can be seen in Figure 4.3 that the ratio of frequencies of these events varies significantly between models. For instance CanESM2 and CMCC-CESM simulate almost entirely displaced vortex events, while IPSL-CM5B-LR and GFDL-CM3 simulate almost entirely split vortex events. In the multi-model mean (MMM) these biases largely cancel to give an approximately equal ratio of displaced to split vortex events, which is in agreement with reanalysis.

The seasonal distribution of these displaced and split vortex events is illustrated in Figure 4.4. Some models (CMCC-CMS, HadGEM2-CC and IPSL-CM5A-LR) replicate the observed distribution, with split vortex events being more likely in early winter, and displaced vortex events in late winter. Other models, however, have a very different distribution of events. CanESM2, CMCC-CESM and MPI-ESM-LR all show little seasonal variability in the frequency of events.

It is now considered how model biases in the climatology of the stratospheric polar vortex, discussed in Section 4.2.1, affect the frequency of split and displaced vortex

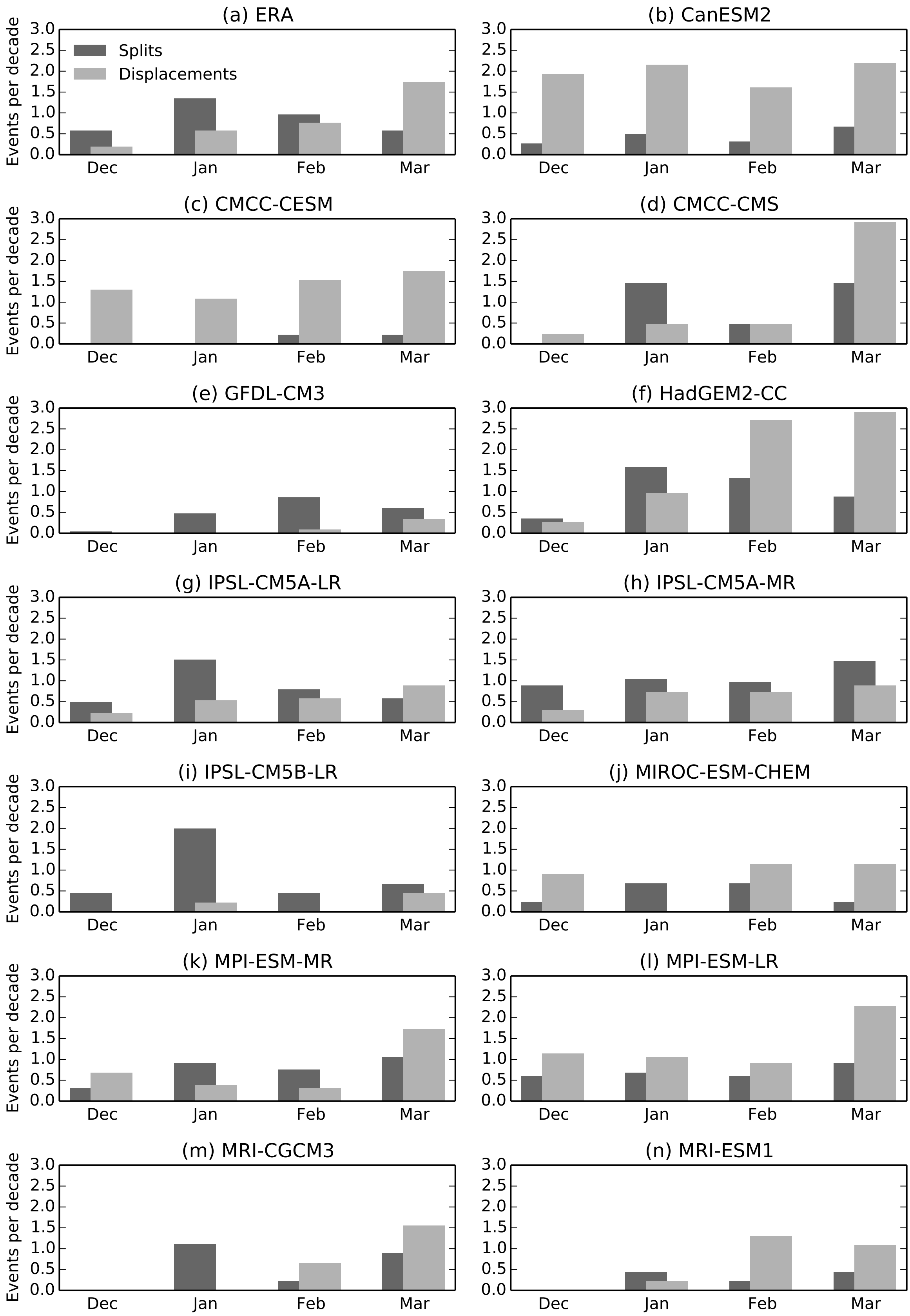

FIGURE 4.4: Seasonal distribution of the occurrence of split and displaced vortex events in ERA (a) and the CMIP5 models.

events. The climatological average state of the vortex is defined by the mode – the peak of the probability distribution function – of the aspect ratio and centroid latitude. Unlike the mean, this quantity is not affected by extreme values and it represents the most likely state of the vortex. The peak can be estimated by the maximum value of a histogram, however this introduces significant random errors and is sensitive to the selection of bin size. A more accurate estimation of the mode can be made by fitting the aspect ratio and centroid latitude with an analytic distribution and then finding the peak of that distribution. Following *Mitchell et al.* [2011], we fit the aspect ratio with a generalised extreme value (GEV) distribution of the form

$$f(x;\mu,\sigma,\xi) = \frac{a^{(-1/\xi)-1}}{\sigma} e^{-a^{-1/\xi}}, \tag{4.1}$$

with

$$a = 1 + \xi \frac{x-\mu}{\sigma}, \tag{4.2}$$

where μ determines the position of the peak along the x-axis, σ determines the variance of the distribution and ξ the skewness. These parameters are determined using the method of maximum-likelihood estimation [*Wilks*, 2006]. This method is also used to fit a Gaussian distribution of the form

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{4.3}$$

where μ determines the position of the peak along the x-axis and σ is the standard deviation, to the cube of the centroid latitude, and then the cube root taken to return the original distribution (this is carried out because an analytic distribution does not fit the unscaled centroid latitude). *Mitchell et al.* [2011] found that these distributions accurately fit the histograms of centroid latitude and aspect ratio in reanalysis data, apart from the extreme tails of the distribution. Qualitative inspection of the distribution for each model confirms that they also provide a similarly good fit to each of the model's histograms.

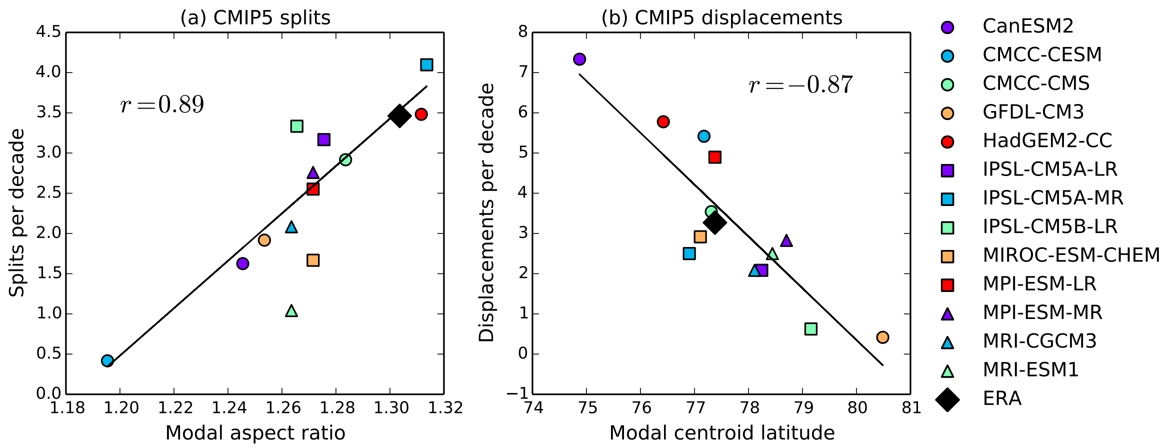

FIGURE 4.5: Comparison of the DJFM mean aspect ratio with frequency of split vortex events (a) and DJFM mean centroid latitude with frequency of displaced vortex events (b) in the CMIP5 ensemble and ERA. Linear best fits and the correlation coefficients for all the models are also shown.

Figure 4.5 shows the relationship between the modal aspect ratio and centroid latitude and frequency of split and displaced vortex events. It can be seen that strong linear relationships exist; the modal aspect ratio accounts for 79% of the variance in the frequency of split vortex events and the modal centroid latitude accounts for 76% of the variance in the frequency of displaced vortex events. This demonstrates that biases in the most likely state of the vortex account for the vast majority of inter-model spread in the representation of extremes. An implication of this is that the models are consistent in their representation of the variability of aspect ratio and centroid latitude, relative to the model climatology.

It can also be seen in Figure 4.5 that the values for ERA lie very close to the best fit lines of the CMIP5 models. This implies that the accuracy of a model's representation of the frequency of displaced and split vortex events can be significantly improved by a more accurate average vortex state. Furthermore, while the ERA value for modal centroid latitude lies approximately in the middle of that for the CMIP5 models, only two models have a larger modal aspect ratio than ERA, indicating that a too circularly-symmetric vortex is a common bias among models. The possible connection between this and planetary wave amplitudes is investigated in the next section.

The structure of the stratospheric polar vortex during split and displaced vortex events in the CMIP5 ensemble is shown in Figure 4.6. This displays composites of 10 hPa geopotential height at the onset date of the events for each model. It can be seen that the majority of models accurately reproduce splitting events as occurring along the 90°W-90°E axis, and displacement events with a vortex shifted towards Scandinavia and Siberia. CanESM2 is an exception to this, with split vortex events which are elliptical but centred quite far from the pole. The IPSL-CM5B-LR model shows a composite for displaced vortex events which actually appears as an uneven split, although this composite only consists of three events so is not statistically significant. There is also significant inter-model spread in the relative strengths of the Aleutian and Azores highs during split vortex events. Several models (GFDL-CM3, IPSL-CM5A-LR, IPSL-CM5B-LR, MRI-ESM1) show an approximately equal strength Aleutian and Azores highs, while others (CMCC-CMS, HadGEM2-CC, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM2) show a weaker Azores high, which is in closer agreement with reanalysis. The more symmetrical Aleutian and Azores highs indicate a greater dominance of wave-2 activity in split vortex events than is found in observations, where not all split vortex events are dominated by wave-2 activity [*Waugh*, 1997; *Mitchell et al.*, 2013].

4.2.3 Planetary wave diagnostics

Given the common bias among the CMIP5 models towards low aspect ratios, it might be suspected that the models have too little wave-2 activity. The relative magnitudes of wave-1 and wave-2 planetary waves are diagnosed using the same method described in Section 3.3.5. In order to determine the climatological wave activity for each model, the resulting distributions of daily wave amplitudes were fitted with a GEV distribution using maximum-likelihood estimation, and the peak of that distribution determined, as in the previous section. A GEV distribution was found to be a good fit for the distributions of each model. As previously, this method has the advantage of determining the most-likely

FIGURE 4.6: Composites of 10 hPa Z at the onset date of split (S) and displaced (D) vortex events in ERA (a,b) and the CMIP5 models. The number of events entering the composite and their type are shown in the title of each plot. The contour interval is 0.3 km.

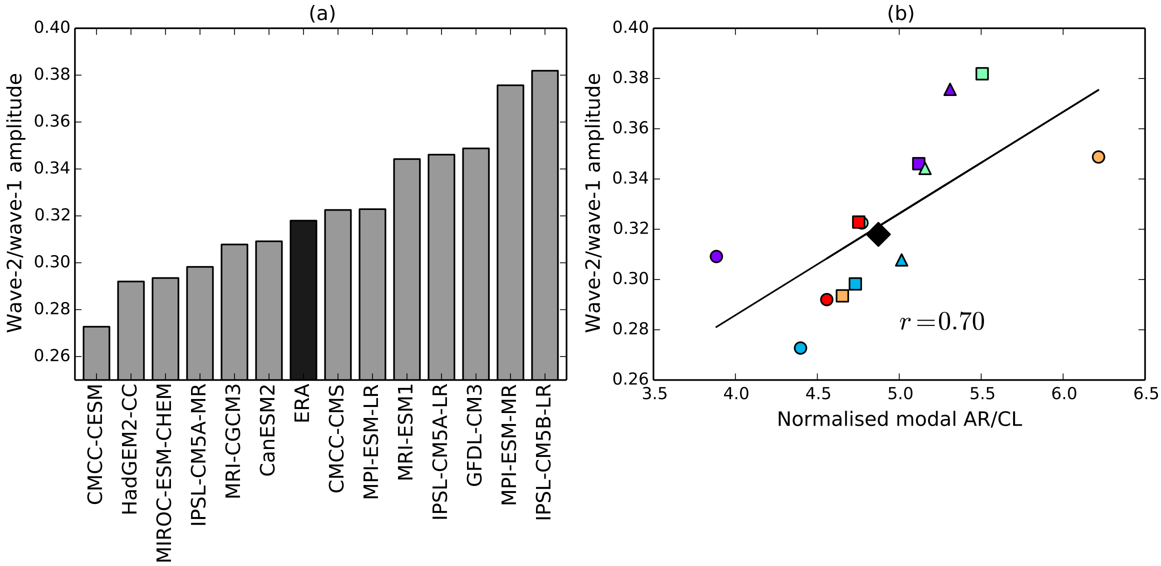

FIGURE 4.7: (a) Ratio of zonal wavenumer-2 to wave-1 amplitude, calculated from the DJFM daily geopotential height field at 60°N, 10 hPa. Amplitudes are calculated from the peak of a GEV distribution fitted to the relevant histogram. (b) Ratio of wave amplitudes against the normalised ratio of modal aspect ratio to modal centroid latitude. Normalisation scales the aspect ratio and centroid latitude by their respective standard deviations. Additionally, the centroid latitude is expressed as the deviation from the pole. Points are coloured as Figure 4.5.

wave amplitude, and unlike the mean it is not affected by large extreme values.

The resulting ratios of wave-2 to wave-1 amplitude for each of the CMIP5 models and ERA are shown in Figure 4.7(a). ERA is seen to lie approximately in the middle of the models; 7 models have a greater ratio, and 6 models lesser. The same ratio of wave amplitude was calculated by *Butchart et al.* [2011] for the CCMVal-2 models. They found a bias towards a relative lack of wave-2 activity, with 13 of 17 models having a lower ratio than reanalysis. The current results show that this bias is not present among the high-top CMIP5 ensemble, and so the bias towards low aspect ratio cannot be attributed to a relative lack of wave-2 activity (this is also true of the absolute value of wave-2 amplitude, in which only 5 of 13 models have a lower value than ERA).

Comparing Figures 4.7(a) and 4.5(a), it can be seen that the model with the lowest ratio of wave amplitudes (CMCC-CESM) also has the lowest aspect ratio. However, across all models the amplitudes of the individual wavenumbers are not significantly correlated with either the modal aspect ratio or the centroid latitude. This suggests that wave-1 activity does not exclusively affect centroid latitude, nor wave-2 exclusively affect aspect ratio, as was found by *Waugh* [1997] and discussed in Appendix 3.3.5. Rather, it is the relative wave amplitudes that affect the relative aspect ratio and centroid latitude climatologies. This is demonstrated in Figure 4.7(b), which shows the ratio of wave amplitudes against the ratio of aspect ratio to centroid latitude. In order that the ratio of moment diagnostics is not biased towards a particular quantity, it is normalised by dividing each of aspect ratio and centroid latitude by their respective standard deviations. The centroid latitude is also expressed as the deviation from the pole (this makes the correlation positive, but does not affect its magnitude). The correlation between the ratio of wave amplitudes to moment diagnostics is 0.70, which is statistically significant at the 95% level. Moreover, if the two outliers (CanESM2 and GFDL-CM3, which were shown to have an unrealistic polar vortex climatology (Figure 4.1)) are removed the correlation becomes much greater ($r = 0.95$). Hence, although there is no one-to-one mapping between individual wave amplitudes and moment diagnostics, the ratio of wave amplitudes strongly determines the relative climatology of the moment diagnostics.

4.3 Stratosphere-troposphere coupling

4.3.1 Zonal-mean response to displaced and split vortex events

The time-height evolution of the atmosphere around split and displaced vortex events in each of the CMIP5 models is displayed in Figure 4.8. This shows composites of polar cap (60°-90°N) geopotential height (Z) anomalies from 90 days before to 90 days following events. The anomalies are calculated from the climatology of each day for each model. The figures extend downwards only to 500 hPa (rather than 1000 hPa). This is because models differ in their representation of geopotential height which is below ground level; some allow negative values, while others set this as an undefined value. This introduces significant errors in the calculation of a climatology and anomalies at levels where the geopotential height is occasionally below ground level. Polar cap Z is highly correlated ($r > 0.95$) with the NAM (calculated from zonal mean Z according to the method of *Baldwin and Thompson* [2009]) over the levels shown in Figure 4.8. *Kushner* [2010] also demonstrated composites of the NAM and polar cap Z following SSWs to be very similar. Indeed, comparing Figures 4.8 (a) and (b) with Figure 3.12 shows that the ERA composites for polar cap Z and the NAM are very similar. In Figure 4.8 the number of events entering each composite is shown in the upper right-hand corner, and it should be noted that composites of a small number of events are likely to be subject to significant statistical uncertainty.

It can be seen that there are large inter-model differences in the evolution of polar cap Z following split and displaced vortex events. For some models (e.g. CanESM2, CMCC-CMS) lower stratospheric anomalies persist for about 45 days, similar to reanalysis, while for others (e.g. IPSL-CM5A-LR, GFDL-CM3) these persist for much longer, beyond 60 days. There are also differences in the stratospheric precursors to events; while some models (e.g. CanESM2, GFDL-CM3, MRI-CGCM3) simulate a stronger negative anomaly prior to displacement events, similar to reanalysis, others (HadGEM2-CC, IPSL-CM5A-MR, MRI-ESM1), show more negative anomalies prior to split vortex events.

Most significantly, there is a large spread in the tropospheric anomalies over the 10-90 days following split and displaced vortex events. Several models (e.g. IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-LR) show only very weak anomalies below approximately 200 hPa, while others (e.g., CMCC-CESM, GFDL-CM3, MRI-ESM1) show stronger anomalies. Among those models which do show stronger tropospheric anomalies, there are also differences in the relative magnitude following split and displaced vortex events. For instance, for GFDL-CM3, and MRI-ESM1 tropospheric anomalies following displaced vortex events are stronger, while for MIROC-ESM-CHEM and MPI-ESM-MR anomalies following split vortex events are stronger, in closer agreement with reanalysis.

As well as these large inter-model differences, there are also some consistent features among models. Almost all models show a barotropic onset to split vortex events, with anomalies occurring at the same time throughout the depth of the atmosphere. In contrast, displaced vortex events appear more baroclinic, with onset occurring first near the uppermost level. This is consistent with the difference found in reanalysis, indicating that this is likely to be a robust difference between the response to split and displaced vortex events. These features are also apparent in the multi-model mean (MMM) (Figures 4.8 C and D). This mean is calculated so as to give each event an equal weight (rather than each model), and so does not give undue weight to models with only a small number of events. On the other hand this does mean that greater weight is given to models with more ensemble members and more events (almost one third of all displaced vortex events come from CanESM2), but the difference in baroclinicity of split and displaced vortex events is observed to be very consistent among models.

4.3.2 Spatial response to displaced and split vortex events

Mean sea-level pressure (MSLP) anomalies averaged over the 30 days following split and displaced vortex events for each of the CMIP5 models are displayed in Figure 4.9. Also shown are the anomalies for ERA (a,b) (the same as Figure 3.14) as well as the multi-model mean (c,d), again calculated so as to give each event an equal weight. The climatology from which anomalies are calculated is determined by a 10-day running mean of the average for each day of the year at each spatial location. The number of events entering each composite is shown, and again, care must be taken interpreting composites of a small number of events due to statistical uncertainty.

Following both split and displaced vortex events, all models show a positive MSLP anomaly near the North Pole, and a negative anomaly centred over Western Europe and the North Atlantic. This pattern is consistent with a negative projection onto the North Atlantic Oscillation. Less consistent among models are anomalies over the North Pacific; many models (e.g. MRI-CGCM3, IPSL-CM5A-LR) show positive anomalies,

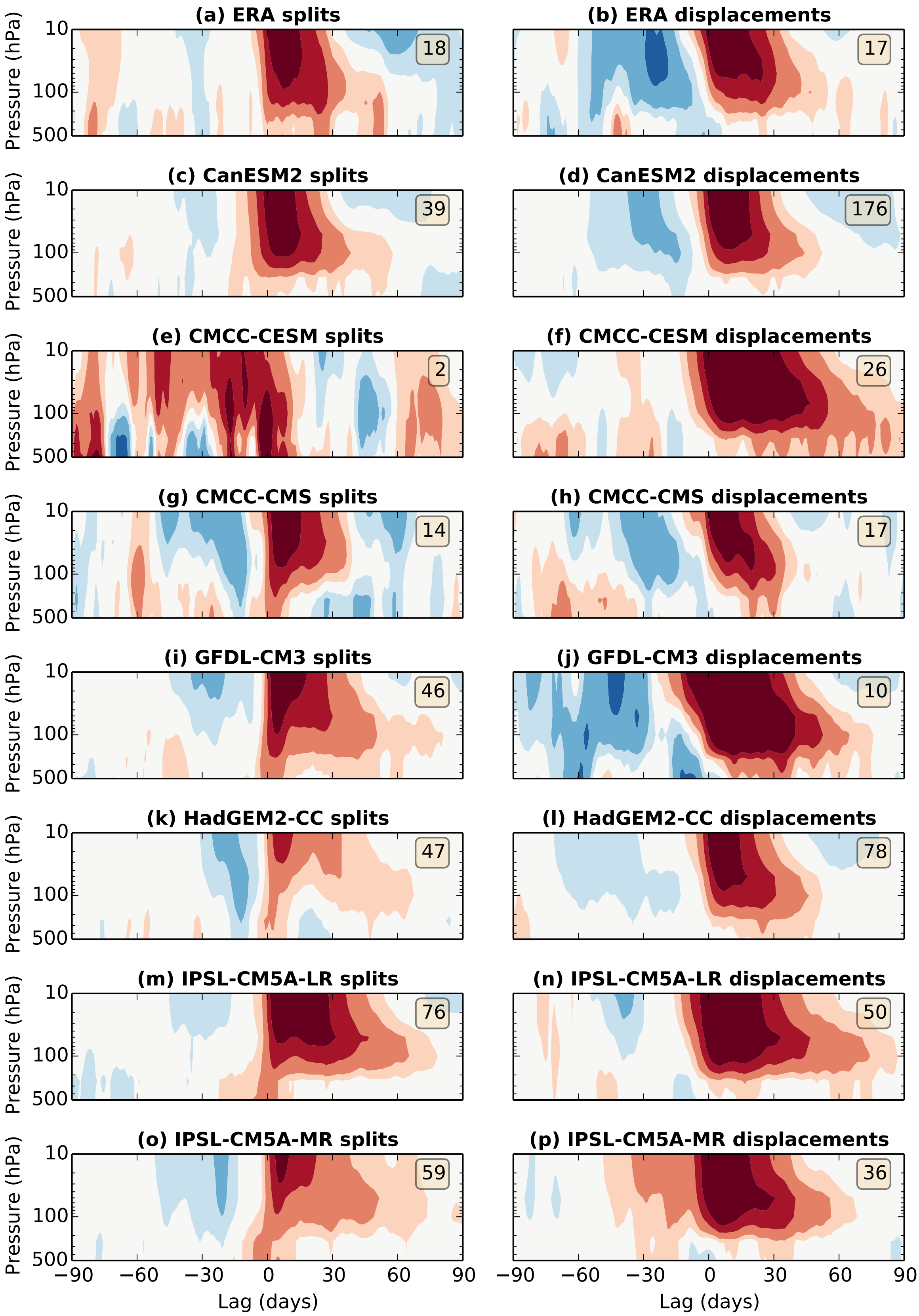

FIGURE 4.8: Composites of normalised polar cap averaged Z anomalies following split and displaced vortex events in ERA (a,b), the CMIP5 models, and the multi-model mean (C,D). Numbers in the upper right of each plot represent the number of events entering the composite.

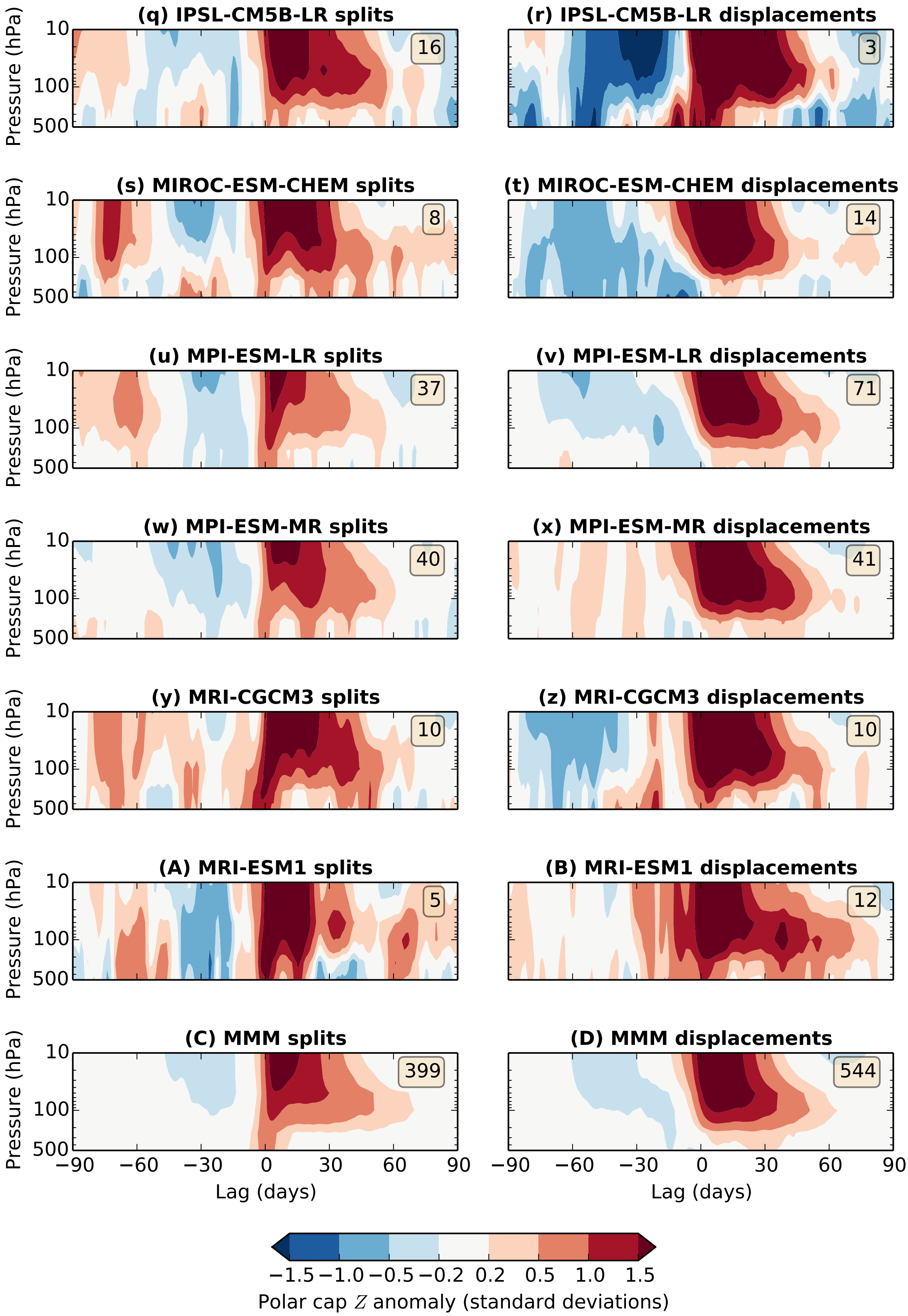

FIGURE 4.8: (Continued)

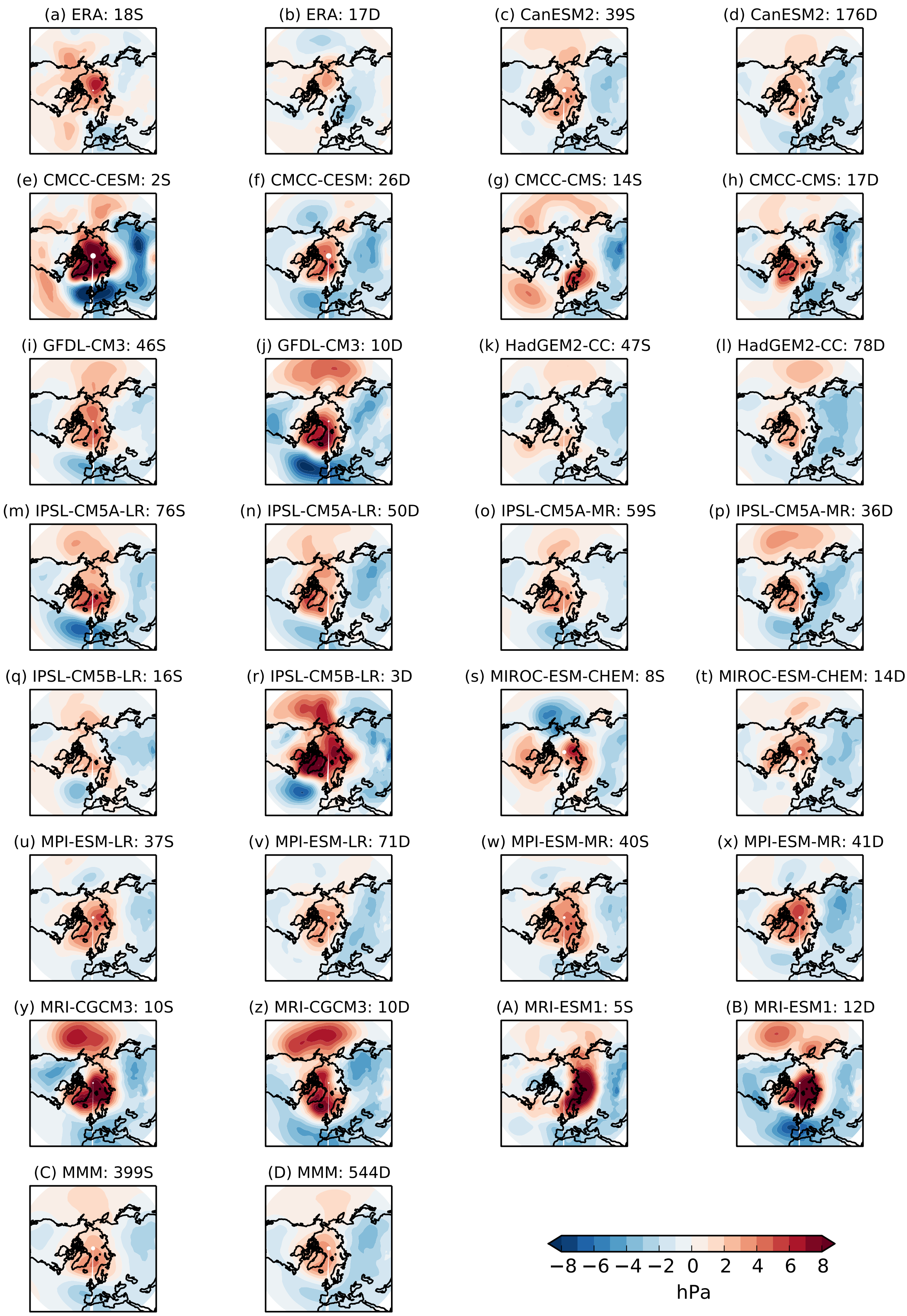

FIGURE 4.9: Composites of mean sea-level pressure anomalies averaged 0-30 days following split (S) and displaced (D) vortex events in the CMIP5 ensemble. Also shown are the ERA composite (a,b) and the multi-model mean (C,D). The multi-model mean is calculated as to give each event an equal weighting. The number of events entering each composite is shown in the title.

while MPI-ESM-LR and MPI-ESM-MR have negative anomalies following both split and displaced vortex events. MIROC-ESM-CHEM has different sign anomalies in the North Pacific following split (negative) and displaced (positive) vortex events. In the multi-model mean, a weakly positive North Pacific anomaly is seen.

This inconsistency in the Pacific anomalies has important consequences for the interpretation of zonal mean anomalies following split and displaced vortex events. For instance, the IPSL-CM5A-LR model shows weak tropospheric anomalies (relative to other models) of polar cap averaged Z following split and displaced vortex events (Figure 4.8 (m,n)), but a relatively strong NAO signal (Figure 4.9 (m,n)), particularly following split vortex events. The reason for this difference is that the model also shows relatively strong positive North Pacific anomalies, that to some extent cancel the North Atlantic anomalies in the polar cap average. Such an effect would also be seen in the NAM, even if calculated from non zonally-averaged Z, since the surface NAM pattern has centres of action of the same sign in the North Atlantic and North Pacific [e.g., *Ambaum et al.*, 2001].

A robust difference found between MSLP anomalies following split and displaced vortex events in almost all models and the multi-model mean is that anomalies over Russia and Eastern Europe are more strongly negative following displaced vortex events. Indeed, the only two models which do not show this difference are CMCC-CESM and MRI-ESM1, but these models simulate only 2 and 5 split vortex events respectively, so the difference is unlikely to be statistically significant. In the multi-model mean, this difference has a magnitude of about 2-3 hPa across most of Russia. In order to understand the possible stratospheric influence on this difference, lower stratospheric anomalies are studied. Figure 4.10 shows composites of 100 hPa Z averaged over the 10 days following the onset of split and displaced vortex events. This shorter time period (rather than the 30 days used for the MSLP composites) is chosen to represent the typical time scale of a split or displacement of the vortex, which is shorter than the time scale taken for the re-formation of the vortex and return towards the climatological mean. However, it should be noted that composites taken over the 30 days show similar structure, but with reduced magnitude (not shown).

As well as the composite for each model, and the split minus displaced vortex difference, a multi-model mean is also shown in Figure 4.10 (Q,R,S). Because this is a mean of model absolute values, and models have different climatologies, the MMMs for split and displaced vortices are scaled to have the same hemispheric mean magnitude. This avoids introducing a bias in the climatology of any particular model into the MMM difference.

For all models with the exception of CanESM2, the 100 hPa Z split vortex composite shows an elliptical vortex with the major axis aligned along the 90°W-90°E line which is similar to the composite at 10 hPa (Figure 4.6). For the displaced vortex composite the 100 hPa vortex is centred over Siberia for almost all models, eastward of the 10 hPa composite which is centred over Scandinavia. This again highlights the more barotropic nature of split vortex events compared to baroclinic displaced vortex events, in which the vortex shows a westward tilt with height (as also demonstrated by *Matthewman et al.* [2009]).

In the split minus displaced vortex difference, the majority of models show a large positive region over Russia and Eastern Europe and a negative region centred over northern Canada. The positive region over Siberia has a relative minimum, located near 90°E (and in some cases it is negative), which is consistent with the position of minimum in the split vortex composite. It can be seen that the multi-model mean difference (Figure 4.10 (S)) is remarkably similar to the reanalysis difference (Figure 4.10), both in terms of the location and magnitude of anomalies. This suggests that the CMIP5 models, on average, realistically represent the evolution of split and displaced vortex events through the depth of the stratosphere. Again, this result is consistent among the majority of models, and so not highly sensitive to the exclusion or inclusion of any particular model.

Figure 4.11 shows the split minus displaced vortex composite difference for MSLP averaged 0-30 days following onset and the 100 hPa Z composite difference averaged 0-10 days following onset, for both ERA and the CMIP5 MMM. Statistical significance in the MSLP difference is calculated by a two-tailed bootstrap test with the null hypothesis that the anomalies following split and displaced vortex events are populations from the same probability distribution. The following procedure is used:

1. All events are grouped together two random subsets are selected from them, with replacement. These are equal in size to the total number of split and displaced vortex events respectively.

2. The difference of the averages of these two subsets is taken.

3. The above is repeated 5000 times to form a distribution of random composite differences.

4. If the actual composite difference lies lower than the 2.5% or higher than the 97.5% levels of this distribution then it can be said there is a less than 95% chance that an anomaly at least this large would arise if anomalies following split and displaced vortex events are populations from the same distribution. Hence the null hypothesis can be rejected.

For the case of ERA, very little statistical significance in the composite difference is seen, while in the CMIP5 MMM there are large statistically significant regions. This is due to the greatly increased sample size in CMIP5; a total of 943 events compared to just 35 in ERA.

In the CMIP5 MMM difference the most significant feature is the large positive anomaly (a result of a more negative anomaly following displaced vortex events) over

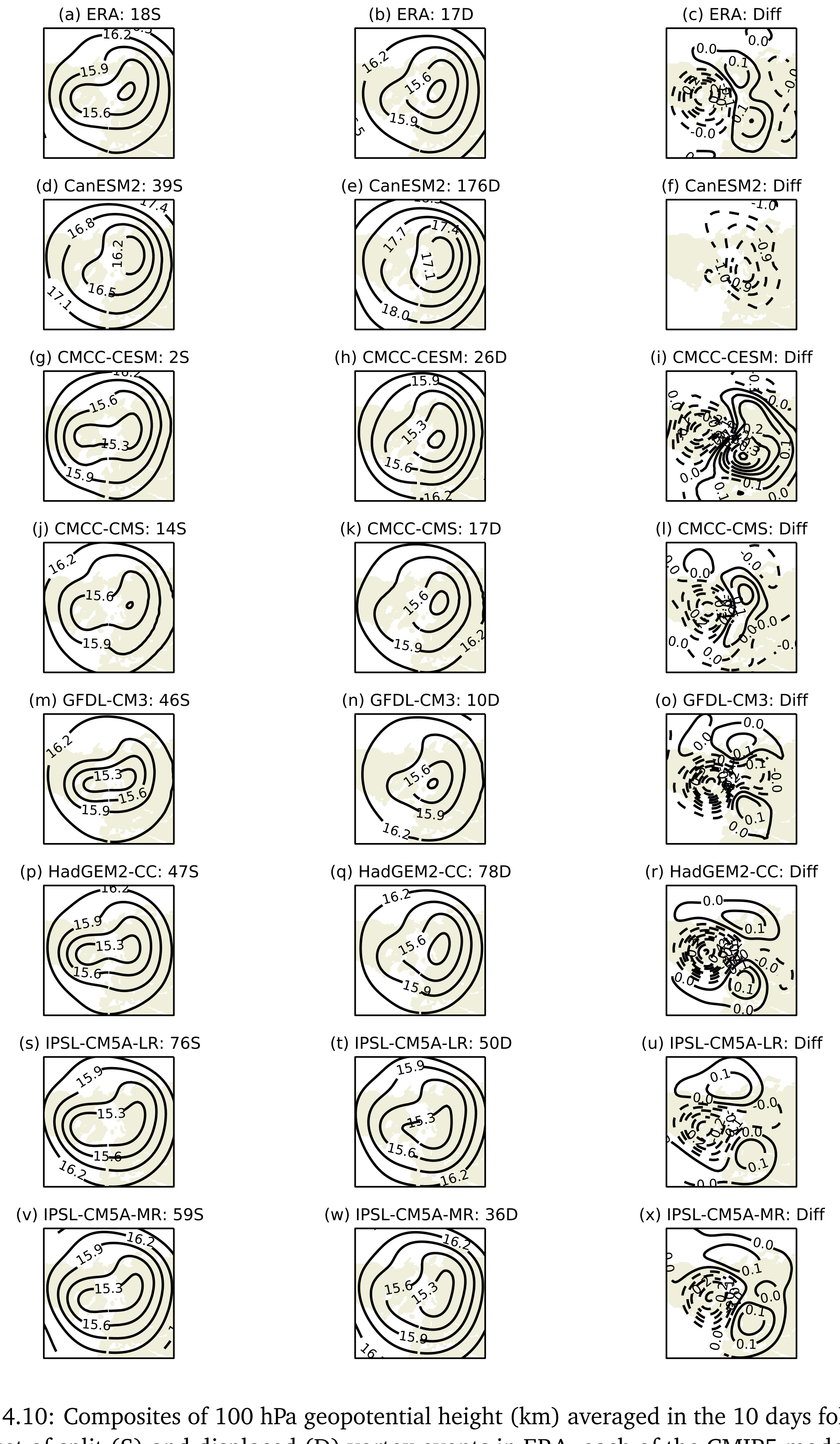

FIGURE 4.10: Composites of 100 hPa geopotential height (km) averaged in the 10 days following the onset of split (S) and displaced (D) vortex events in ERA, each of the CMIP5 models, and the multi-model mean (MMM). The right hand column displays the difference of splits minus displacements. The multi-model mean is calculated so as to give each event an equal weighting.

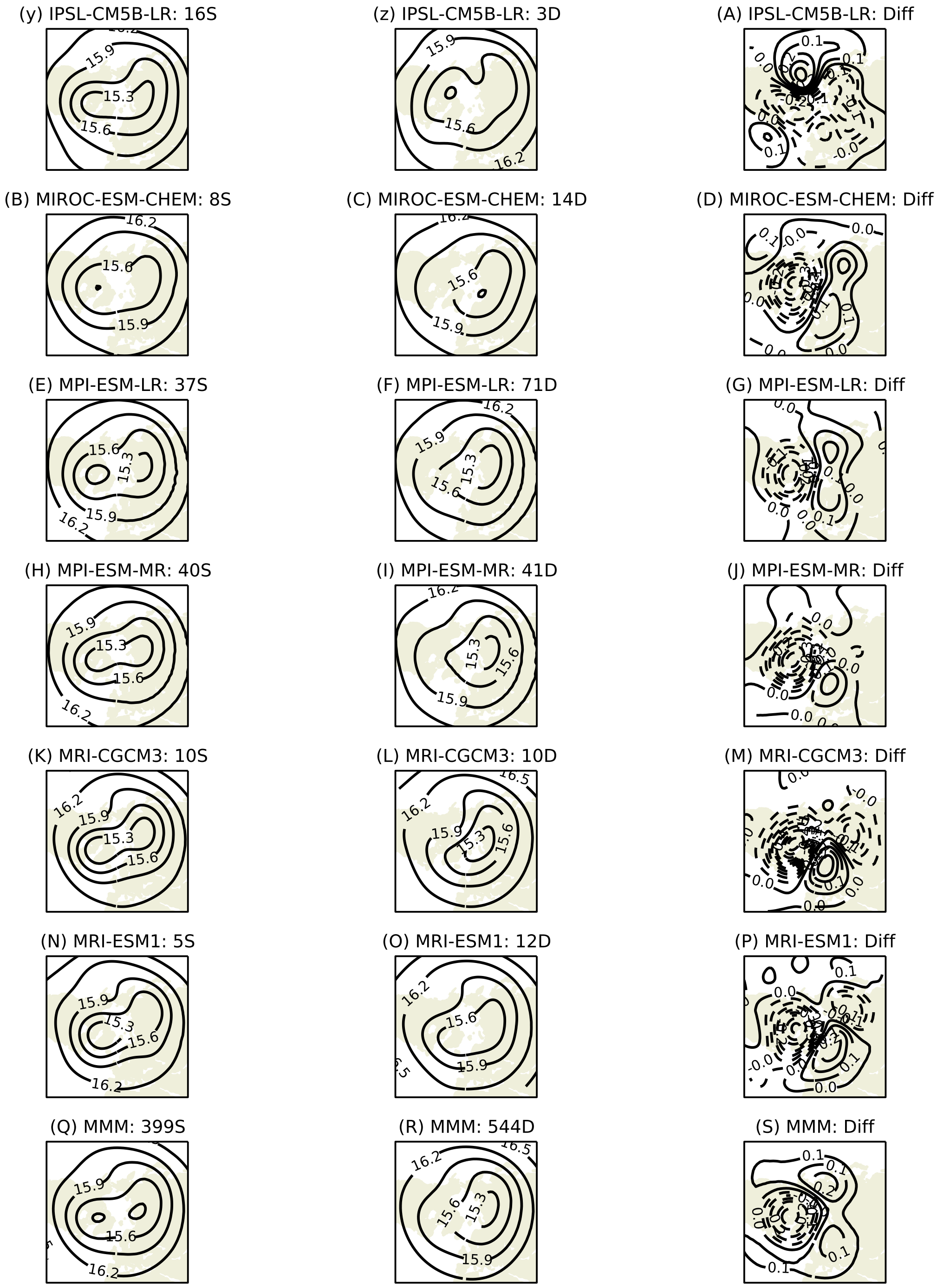

FIGURE 4.10: (Continued)

Scandinavia, Eastern Europe and Russia. There is also a significant negative anomaly over northern Canada and a positive anomaly in the western Atlantic. This pattern is zonally asymmetric and so does not project strongly onto the polar cap average, therefore explaining the small difference in polar cap averaged Z (Figure 4.8 (C,D)). The CMIP5 difference pattern also does not strongly project onto the NAO as there is a similarly negative NAO following both split and displaced vortex events (Figure 4.9 (C,D)). For the ERA MSLP difference there are stronger negative anomalies over Europe (although not statistically significant) and positive anomalies over the North Pole, which does project more strongly onto the NAO, as discussed in Section 3.4.

The CMIP5 difference of Figure 4.11(b) shows the positive 100 hPa Z anomalies over Siberia over-lie the positive MSLP anomalies, while the negative 100 hPa Z over northern Canada over-lies negative MSLP anomalies. A somewhat similar, but not statistically significant pattern is seen in ERA, although the Siberian anomaly is more polewards and the negative anomaly over Canada is much weaker. Importantly, the 100 hPa pressure surface lies close to the tropopause, and 100 hPa Z can therefore give an indication of tropopause height. Indeed, comparing the 100 hPa Z and tropopause height anomalies for ERA (Figures 4.10 (a,b) and 3.16 (c,d)) shows that anomalous negative Z is approximately co-located with an anomalously high tropopause, although the tropopause height field is much noisier. The implications of this result for mechanisms of stratosphere-troposphere coupling are discussed in Section 4.4.2.

4.4 Discussion

4.4.1 Effect of model resolution

We have demonstrated that there are large differences in the representation of the average state of the stratospheric polar vortex among high-top CMIP5 models. This amounts to an inter-model range in the modal centroid latitude of more that 5° and in the aspect ratio of 0.12, which corresponds to (77% and 22% of the ERA standard

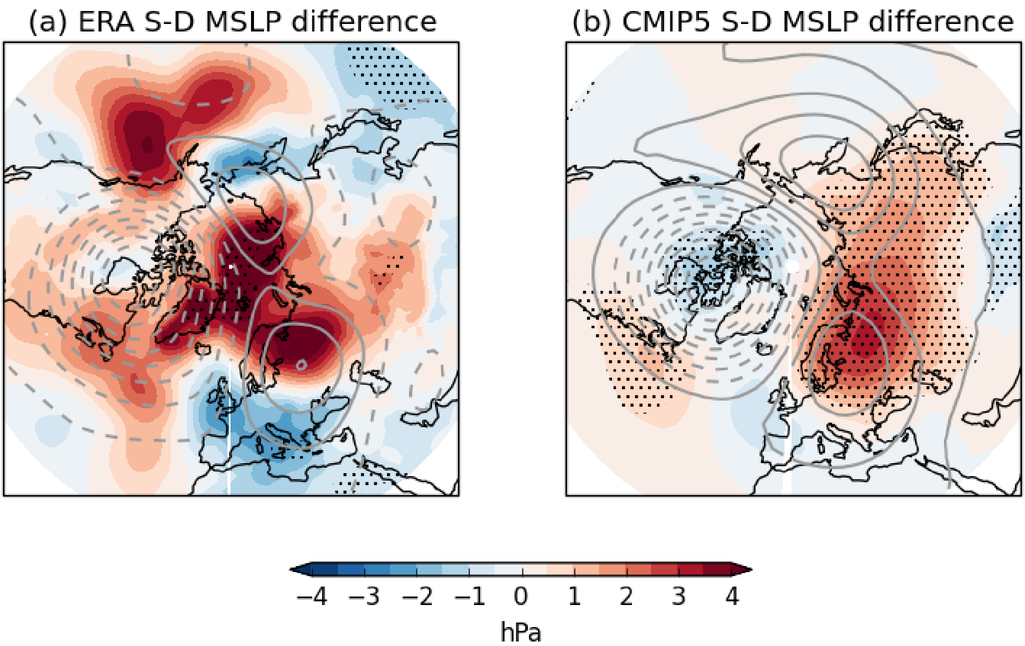

FIGURE 4.11: Difference (S-D) of composites of mean sea-level pressure averaged 0-30 days following split (S) and displaced (D) vortex events in ERA and the CMIP5 ensemble. Stippling indicates regions that are >95% significant according to a two-tailed bootstrap test. Grey contours represent the difference in 100 hPa geopotential height averaged 0-10 days following events the contour interval is 40 m, dashed contours represent negative values and the lowest magnitude contours are ±20 m.

deviation respectively). The multi-model mean centroid latitude is approximately in agreement with ERA because a similar number of models have an equatorward as a poleward bias. In contrast, only two of 14 models have higher modal aspect ratio than reanalysis.

Importantly, we have shown that these biases in the undisturbed state of the vortex are closely related to biases in the frequency of split and displaced vortex events. An implication of this is that these models have a realistic representation of variability relative to the average state. In order to understand the origin of these biases we now consider whether any of the model horizontal and vertical resolution properties listed in Table 4.1 are related to models' polar vortex climatology.

There are no statistically significant correlations between the modal aspect ratio or centroid latitude and horizontal resolution or between the centroid latitude and vertical resolution. However, a stronger relationship is found between vertical resolution and the modal aspect ratio and this is shown in Figure 4.12. Even so, the relatively wide scatter of points as well as the small correlation coefficient values indicate that vertical resolution fails to account for a substantial fraction of inter-model variability in the modal aspect ratio. Indeed, the p-values of these correlations are 0.062 and 0.188 for the 5-15 km and 15-30 km vertical resolutions respectively (calculated using a two-tailed t-test). However, these relationships appear quite nonlinear, with aspect ratio being relatively more sensitive to changes in resolution when the resolution is coarse (high $\mathrm{d}z$) and less sensitive when the resolution is finer (low $\mathrm{d}z$). This can be addressed by instead calculating the Spearman's rank correlations, which test the monotonicity of the relationships. These are -0.60 ($p = 0.031$) and -0.75 ($p = 0.003$) for vertical resolutions over 5-15 km and 15-30 km respectively, therefore indicating more statistically significant relationships. However, because the two measures of vertical resolution are themselves correlated ($r = 0.79$), it is difficult to interpret which of the two regions (if any) has the largest impact on the modal aspect ratio.

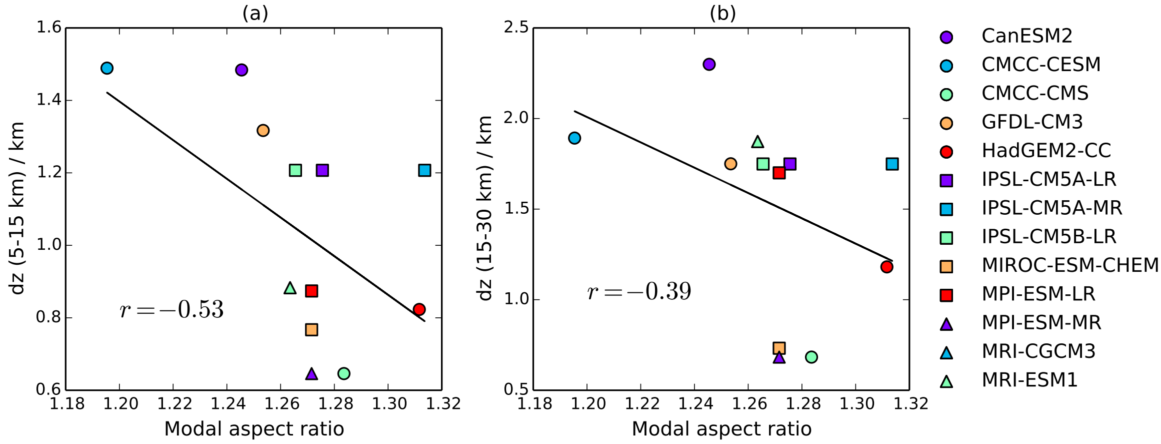

FIGURE 4.12: Correlation between modal aspect ratio and model vertical resolution over two regions (5-15 km and 15-30 km). Correlations, r, are shown along with the p-value calculated by a bootstrap test with null hypothesis $r = 0$.

Two important caveats should be noted in interpreting these correlations. First, comparing pairs of models from the same family but with different resolutions does not show a correlation between vertical resolution and modal aspect ratio. In the present ensemble this comparison is limited to IPSL-CM5A-LR/IPSL-CM5A-MR and MPI-ESM-LR/MPI-ESM-MR. It can be seen from Figure 4.12(a) that the IPSL models have the same vertical resolution but different modal aspect ratios and the MPI models have different vertical resolutions but very similar modal aspect ratios. While this is a very limited comparison of only two pairs of models, it may suggest that it is in fact other model differences which may give rise to the correlations shown in Figure 4.12.

A second caveat is that interpreting this significance, it should be noted that six different relationships between model resolution parameters and moment diagnostics have been tested using the parameters in Table 4.1 (treating the measures of vertical resolution as a single parameter since they are highly correlated). Assuming these parameters to be independent, there is an approximately 26% chance of finding at least one 95% significant correlation among those tested.

It is interesting to note, however, that *Anstey et al.* [2013] found vertical resolution among CMIP5 models to correlate with increased NH winter blocking frequency (although, as with this study, the relationship did not hold in comparing models from the same family). They found the strongest relationship to be with UTLS vertical resolution (i.e. over the 5-15 km region). Since blocking events are known to be closely linked to stratospheric variability (as discussed in Sections 2.4.1 and 2.4.2), the combination of the present study with *Anstey et al.* [2013] may suggest that UTLS vertical resolution is an important factor in the representation of stratosphere-troposphere coupling.

There is also some physical motivation for the importance of UTLS vertical resolution in the representation of stratosphere-troposphere coupling because of the sensitivity of planetary wave propagation to vertical gradients in this region. This sensitivity can be measured by the quasi-geostrophic refractive index, n_s [*Matsuno*, 1970], given by

$$n_s^2 = \frac{1}{a\overline{u}}\frac{\partial \overline{q}}{\partial \phi} - \frac{s^2}{a^2\cos^2\phi} - \frac{f^2}{4N^2H^2}, \tag{4.4}$$

where s is the zonal wavenumber, and

$$\frac{\partial \overline{q}}{\partial \phi} = 2\Omega\cos\phi - \frac{\partial}{\partial \phi}\left[\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}(\overline{u}\cos\phi)\right] - \frac{a}{\rho_0}\frac{\partial}{\partial z}\left(\frac{\rho_0 f^2}{N^2}\frac{\partial \overline{u}}{\partial z}\right). \tag{4.5}$$

The refractive index is significant because planetary waves tending to propagate towards regions of high n_s and becoming evanescent in regions where $n_s < 0$. It can be seen that n_s depends both on the the vertical gradient in static stability ($\partial N^2/\partial z$) and the vertical shear of zonal wind ($\partial \overline{u}/\partial z$). Using a linear primitive equation numerical model, *Chen and Robinson* [1992] found a large vertical gradient in static stability near the tropopause as well as a strong vertical shear of the zonal flow. They found that small changes in these quantities have a large effect on n_s, and therefore that the tropopause acts as a "valve" for the propagation of planetary waves between the troposphere and stratosphere. In a more recent observational study, *Grise et al.* [2010] have confirmed the existence of fine-scale structure in static stability near the tropopause. It therefore may be the case that a higher model vertical resolution near the tropopause is necessary to capture this structure and hence more accurately represent planetary wave propagation from troposphere to stratosphere.

The fact that there is not a significant relationship between the modal centroid latitude and vertical resolution may suggest that wave-2 propagation is more sensitive to vertical resolution than wave-1. Indeed, it might be expected that wave-2 propagation is more sensitive to near-tropopause resolution because the 'window' for permissible wave-2 propagation is narrower (*Charney and Drazin* [1961]; Equation 2.7), meaning smaller differences in vertical gradients are required to exclude wave-2 propagation than wave-1 propagation. However, no significant correlations are found among the CMIP5 models between vertical resolution and wave amplitudes (using the measures shown in Figure 4.7).

Overall, a more systematic study is necessary to understand the importance of UTLS vertical resolution in the representation of stratosphere-troposphere coupling by climate models. This should consist of a 'clean comparison' of models in which only vertical resolution is varied, and which therefore avoids the difficulty of interpreting results when many parameters are varied at once as in the CMIP5 ensemble.

4.4.2 Measures of stratosphere-troposphere coupling

We have shown that zonally-averaged quantities such as polar cap Z (which is highly correlated with the NAM [*Kushner*, 2010]) following split and displaced vortex events are much less consistent among models than the NAO. This inconsistency is dominated by differences in the North Pacific, with some models showing positive MSLP anomalies and others negative. Interestingly, a similar result was found in the recent study of *Davini et al.* [2014], who found the blocking pattern associated with SSWs to be consistent with that associated with the NAM over the North Atlantic, but not over the North Pacific.

As discussed in Section 2.4 many studies of stratosphere-troposphere coupling have focused on the lag-height behaviour the NAM. For instance the comparison of stratosphere-troposphere coupling in high-top and low-top CMIP5 models by *Charlton-*

Perez et al. [2013]. This and several other studies make further approximations as to the zonal nature of the coupling by calculating the NAM based on zonal-mean geopotential height, according to the method of *Baldwin and Thompson* [2009]. Our results suggest that because of the difference in model consistency over the two ocean basins, zonal-mean diagnostics or the NAM alone are not good descriptors of inter-model variability. Therefore, we suggest that the NAO index or the full two-dimensional surface fields should be shown alongside the NAM when making inter-model comparisons.

This difference between the NAM and NAO signals in CMIP5 models may also give some insight into the physical relevance of these two modes of variability. Some studies have suggested that the NAO is in fact a regional manifestation of the planetary-scale annular structure of the NAM [e.g., *Thompson and Wallace*, 1998; *Wallace and Thompson*, 2002]. Furthermore, many observational studies have asserted that tropospheric anomalies following SSWs represent the NAM [e.g., *Baldwin and Dunkerton*, 1999, 2001; *Thompson et al.*, 2000].

On the other hand, *Ambaum et al.* [2001] suggested that the NAO paradigm is a more physically relevant measure of NH variability than the NAM. They found that MSLP anomalies over the North Atlantic and Pacific are not significantly correlated (also shown by *Deser* [2000]) and argued that that the annular NAM pattern is a statistical artifact. *Huth* [2006] also showed that principal component analysis favours the NAO as the more physically relevant mode of variability. Furthermore, *Ambaum and Hoskins* [2002] found that changes in North Pacific tropospheric subtropical and polar jets are much less correlated with the strength of the stratospheric polar vortex than are the North Atlantic jets.

Under the significant assumption that the CMIP5 models can accurately represent the physics underlying these modes of variability (the investigation of which is beyond the scope of this chapter), our results tend to favour the NAO rather than the NAM as the more physically relevant mode, at least in terms of stratosphere-troposphere coupling.

4.4.3 *Difference between split and displaced vortex events*

As well as the consistent NAO signal following split and displaced vortex events, we have found that there are also some consistent differences in anomalies following the two types of event. In particular, MSLP anomalies following displaced vortex events are more negative over Scandinavia and Siberia than following split vortex events. From the fact that these MSLP differences are co-located with 100 hPa Z (Figure 4.11), which is in turn related to tropopause height, it may be possible to gain some understanding of the mechanism behind the difference in the surface response to split and displaced vortex events.

This co-location of surface anomalies and tropopause height is consistent with a localised spinup/spindown caused by stretching/compression of the tropospheric column. Changes in tropopause height are, in turn, caused by the bending of isentropic surfaces towards PV anomalies resulting from the movement of the stratospheric polar vortex. Such a mechanism was discussed by *Ambaum and Hoskins* [2002] and in Section 2.4.3.

Other mechanisms that have been proposed for stratosphere-troposphere coupling fail to account for these regional differences. For instance, the amplification of intrinsic modes of variability [*Robinson*, 1991] can only explain differences which project onto these modes such as the NAO or NAM, unlike observed difference. Stratosphere-troposphere coupling by the reflection of planetary waves [*Perlwitz and Harnik*, 2003; *Shaw et al.*, 2010] also cannot explain the observed difference since it does not project onto the dominant tropospheric planetary wave modes.

This argument relates only to the mechanism underlying the *difference* between the surface responses to split and displaced vortex events, and not to the overall responses. It is important to note that there are many similarities in the responses, especially in the

NAO region. All the mechanisms discussed in Section 2.4.3 can be used to explain a stratospheric influence on the NAO, and since they are not physically inconsistent with one another, it is possible that a number may operate at the same time.

Unfortunately, the small number of observed split and displaced vortex events combined with large tropospheric noise means there is very little statistical significance in the observed difference of MSLP anomalies (see Figure 4.11(a)). Hence it is not possible to compare our model and observational results for this difference, and there remains the possibility that CMIP5 models do not realistically represent the surface responses to split and displaced vortex events. Therefore, it is possible that stratosphere-troposphere coupling mechanisms in models are different to those in the real world.

4.5 Conclusions

Applying the method developed in Chapter 3, the climatology of the stratospheric polar vortex and its coupling with the troposphere has been analysed in stratosphere-resolving CMIP5 simulations. The main conclusions of this investigation are as follows:

Model representation of the stratospheric polar vortex and stratosphere-tropsophere coupling. A wide range of biases among CMIP5 models has been found in the average state of the stratospheric polar vortex. Some models have a vortex which is too equatorward, others too poleward. The majority of models have a vortex which is too circularly symmetric. The relative magnitudes of wave-1 and wave-2 activity to a large extent determine these biases, although there is no one-to-one relationship between wavenumbers and the aspect ratio or centroid latitude of the vortex. The bias towards low aspect ratios may also be related to vertical resolution (models with finer resolution having larger average aspect ratios), although this result is not highly statistically significant.

There is also a wide spread in the frequency of split and displaced vortex events, although in the multi-model mean the frequency of these events is in agreement with observations. Importantly, biases in the average state of the vortex have been shown to relate closely to biases in the frequency of split and displaced vortex events. Hence an improvement in the average state of the vortex is likely to lead to an improvement in the representation of extremes.

Almost all models accurately simulate the more barotropic nature of split vortex events compared to displaced vortex events. MSLP anomalies following these events consistently show a negative NAO in line with observations, but are much less consistent in the North Pacific, leading to a large spread when zonal mean quantities are investigated.

Mechanisms for stratosphere-troposphere coupling. Consistent differences in the MSLP anomalies following split and displaced vortex events in the CMIP5 models have been found to be co-located with the difference in near-tropopause Z anomalies. This is consistent with a localised tropospheric response to stratospheric PV anomalies (as discussed in Section 2.4.3) being the mechanism behind the different responses to the two events. It also excludes mechanisms which rely on projections onto major modes of variability such as the NAO or NAM. This result only applies to the *difference* between responses to split and displaced vortex events, not the overall responses, which also have some similarities.

CHAPTER 5

The role of the stratosphere in seasonal prediction

Most of work in this chapter which relates to the Southern Hemisphere was published in *Journal of Climate* [*Seviour et al.*, 2014].

5.1 Introduction

Accurate prediction of the atmospheric circulation several months in advance relies on the presence of low-frequency predictable signals in the climate system. It has now been demonstrated that the stratosphere is an important pathway for the communication of predictable tropical signals across the globe; in particular, the El Niño-Southern Oscillation (ENSO) [*Bell et al.*, 2009; *Ineson and Scaife*, 2009; *Hurwitz et al.*, 2011], Quasi-Biennial Oscillation (QBO) [*Marshall and Scaife*, 2009; *Garfinkel and Hartmann*, 2011], and 11-year solar cycle [*Haigh*, 2003; *Gray et al.*, 2013]. These teleconnections allow for the possibility of significant predictability in regions remote from the direct effect of the signal. Despite this, many operational seasonal forecast models include only a poor representation of the stratosphere [*Maycock et al.*, 2011], and it has been suggested that this contributes to their lack of seasonal forecast skill in the extratropics [*Smith et al.*, 2012].

Furthermore, because stratospheric anomalies persist for longer than those in the troposphere and can influence surface weather patterns, the initial conditions of the stratosphere itself can act as a source of enhanced predictability [*Baldwin et al.*, 2003; *Charlton et al.*, 2003; *Christiansen*, 2005; *Hardiman et al.*, 2011]. Because the effect of the stratosphere on the troposphere is especially pronounced following SSW events, past work has focused on the influence of these events on forecast skill. For instance, both *Kuroda* [2008] and *Sigmond et al.* [2013] found that enhanced tropospheric predictability can be obtained if forecasts are initialised at the onset of SSW events. However, SSWs are highly nonlinear events which previous studies have not found predictable beyond about two weeks in advance [*Marshall and Scaife*, 2010; *Taguchi*, 2014]. This may therefore limit their usefulness in seasonal prediction. SSWs also occur almost exclusively in the NH, with only one event in the approximately 60 year record having been observed in the SH, in September 2002 [*Roscoe et al.*, 2005].

As discussed in Section 2.1.3, the rarity of SSWs in the SH is a result of less dynamical forcing from vertically propagating planetary waves in the SH relative to the NH stratosphere. This reduced dynamical forcing also means that anomalies in the Antarctic lower stratosphere persist for longer than those in the Arctic [*Simpson et al.*, 2011]. Hence, the SH stratospheric circulation may be predictable on longer time scales, and thus more useful for seasonal forecasts despite the lack of SSWs. Indeed, *Thompson et al.* [2005] and *Son et al.* [2013] have found that smaller-amplitude variations in the Antarctic stratospheric polar vortex are followed by coherent temperature and pressure anomalies at the Earth's surface which resemble the Southern Annular Mode (SAM) pattern. These observations led *Roff et al.* [2011] to find that improved forecasts of the SAM up to 30 days ahead may be achieved with a stratosphere-resolving model. As the dominant mode of variability in the extratropical SH, the SAM affects the position of storm tracks, rainfall, surface air temperature, and ocean temperatures across the extratropics [e.g., *Silvestri and Vera*, 2003; *Reason and Rouault*, 2005; *Hendon et al.*, 2007]. As such, there are considerable societal benefits and interests in its prediction

[*Lim et al.*, 2013].

Another reason for interest in the prediction of the Antarctic stratosphere is the interannual variability in springtime ozone depletion, which can significantly affect the amount of harmful ultraviolet radiation reaching the Earth's surface over the Southern Hemisphere. The magnitude of this interannual variability is a significant fraction of the magnitude of long-term depletion caused by emission of chlorofluorocarbons (CFCs) and other ozone-depleting substances. While ozone-depleted air is confined over the polar region by the stratospheric polar vortex during winter and spring (resulting in the ozone hole), this air is released to mid-latitudes following the ultimate breakdown of the vortex (final warming) in late spring/early summer. The extent of the resulting summertime ozone depletion is largely determined by the total deficit in ozone over the Antarctic during spring [*Bodeker et al.*, 2005].

As discussed in Section 2.2, dynamics play an important role in ozone depletion. Indeed, *Salby et al.* [2012] have shown that interannual variations in Antarctic ozone depletion are highly correlated with changes in planetary wave forcing of the stratosphere. They found that the anomalous vertical EP flux at 70 hPa poleward of 40°S during August-September explains almost all the interannual variance of anomalous ozone depletion during September–November. Using this relationship, they postulate that accurate prediction of planetary wave forcing could allow skillful seasonal forecasts of ozone depletion.

In this chapter, we address directly the predictability of the stratospheric polar vortices using a set of hindcasts (or historical re-forecasts) from a new operational seasonal forecast system with a fully stratosphere-resolving general circulation model. The system accurately simulates the climatology of the NH stratospheric polar vortex including the aspect ratio and centroid latitude. However, we find it does not skilfully predict the winter mean vortex strength, the occurence of SSWs, or split and displaced vortex events. On the other hand, we find significant skill in the prediction of the

Antarctic stratospheric polar vortex up to four months in advance, including for the 2002 SSW. Using the observed relationship between column ozone quantities and the stratospheric circulation, we are then able to infer skillful predictions of springtime ozone depletion, confirming the hypothesis of *Salby et al.* [2012]. This exceeds the lead-time of other contemporary ozone forecasts, which are typically no more than two weeks [*Eskes*, 2005]. The forecast system also shows significant levels of skill in the prediction of the surface SAM at seasonal lead times. By studying the variation of hindcast skill with time and height, we demonstrate that this skill is significantly influenced by the descent of predictable stratospheric circulation anomalies.

5.2 Seasonal forecast system

The analysis in this chapter is based on results from a set of hindcast predictions produced by the Met Office Global Seasonal Forecast System 5 (GloSea5) [*MacLachlan et al.*, 2014]. This system is based upon the HadGEM3 coupled general circulation model [*Hewitt et al.*, 2011], with an atmospheric resolution of 0.83° longitude by 0.56° latitude, 85 quasi-horizontal atmospheric levels and an upper boundary at 85 km. The ocean resolution is 0.25° in longitude and latitude, with 75 quasi-horizontal levels.

Initial conditions for the atmosphere and land surface were taken from the ERA-Interim reanalysis [*Dee et al.*, 2011], and initial ocean and sea-ice concentrations from the GloSea5 Ocean and Sea Ice Analysis, based on the FOAM data assimilation system [*Blockley et al.*, 2013]. The ERA-Interim data are linearly interpolated onto model levels between the surface and 64.56 km (near 0.1 hPa), and the 64.56 km values are then replicated onto the four subsequent levels up to 85 km. FOAM data are on the same grid as the ocean model. Beyond initialisation the model takes no further observational data, and contains no flux corrections or relaxations to climatology. The model lacks interactive chemistry and ozone concentrations are fixed to observed climatological values averaged over 1994–2005, including a seasonal cycle from the Stratosphere-troposphere Processes and their Role in Climate (SPARC) climatology [*Cionni et al.*, 2011]. Climate forcings such as CO_2 and CH_4 concentrations are set to observed values up to 2005 and then follow the IPCC RCP4.5 scenario.

Scaife et al. [2014] have shown that this seasonal forecast system produces highly skillful forecasts of the North Atlantic Oscillation (NAO) during the Northern Hemisphere winter. They found the correlation of the ensemble mean DJF average NAO with observed values to be $r = 0.62$, which is statistically significant from zero at the 99% level. They argue that the combined effects of ENSO, QBO and sea-ice teleconnections, as well as the increased ocean resolution which has improved the representation of Northern Hemisphere blocking events [*Scaife et al.*, 2011b], contribute to this skill.

Hindcast accuracy is verified by comparison to the ERA-Interim reanalysis [*Dee et al.*, 2011]. As discussed in Chapter 3, the ERA-Interim data set has been demonstrated to have realistic representation of the stratospheric meridional circulation [*Seviour et al.*, 2012; *Monge-Sanz et al.*, 2013]. It also assimilates observations of ozone concentrations, and this assimilation has been demonstrated to be in close agreement with independent satellite data [*Dragani*, 2011]

In this chapter, hindcasts are analysed for two seasons; December–February (DJF) for prediction of the Northern Hemisphere winter stratospheric polar vortex, and September–November (SON) for the Southern Hemisphere. The SON season is chosen because it represents the time of maximum SH ozone depletion and stratospheric polar vortex variability. For SON a 15-member ensemble of hindcasts was run for each year in the period 1996–2009, while for the DJF analysis a longer 24-member ensemble is available for the winters 1992/1993–2011/2012. The hindcast length is approximately four months from three separate start dates spaced two weeks apart and centered on 1st August (SON) or 1st November (DJF), with an equal number of members initialised on each start date. Members initialised on the same start date differ only by stochastic parameterisation of model physics, using the Stochastic Kinetic Energy Backscatter v2

Season	Hindcast period	Initialisation dates	Ensemble size
DJF	92/93–11/12 (20 years)	25/10, 01/11, 09/11	24 (8 on each date)
SON	1996–2009 (14 years)	25/07, 01/08, 09/08	15 (5 on each date)

TABLE 5.1: Summary table of the two sets of hindcast simulations analysed.

[SKEB; *Bowler et al.*, 2009] scheme. Details of the hindcast runs for these two seasons are summarised in Table 5.1.

It should be noted that the 20-year, 20-member ensemble hindcasts of DJF were extended from a 14-year, 15-member ensemble, the same as the SON hindcasts. These extended simulations are, however, slightly shorter than the original ensemble; ending at the beginning of March rather than the beginning of April. This therefore limits our analysis of the full ensemble so as not to include March. Furthermore, some variables were not produced by the extended DJF ensemble and so the original shorter ensemble must be used in some cases, which is made clear in the text where necessary.

In order to illustrate these hindcast simulations, timeseries of zonal-mean zonal wind ($\overline{u}$) at 10 hPa are shown for 60°N from November–March (Figure 5.1) and 60°S from August–December (Figure 5.2). ERA-Interim values are also shown in both cases (black lines). These are the approximate positions of the centre of the mean position of the stratospheric polar vortex in the mid-stratosphere, and therefore indicate the strength of the stratospheric polar vortex. It can immediately be seen that there is a much greater ensemble spread in the NH, owing to the much greater dynamical variability of the NH stratospheric polar vortex (this is true even accounting for the increased ensemble size in the NH). In both cases it can also be seen that the ensemble spread is initially small and increases rapidly after approximately 15-30 days. This demonstrates the initial constraint to ERA-Interim and the rapid growth of small differences in initial conditions, because of the chaotic nature of the atmosphere. The predictive skill of for the Northern Hemisphere (DJF) is analysed in the next section, and the Southern Hemisphere (SON) in Section 5.4.

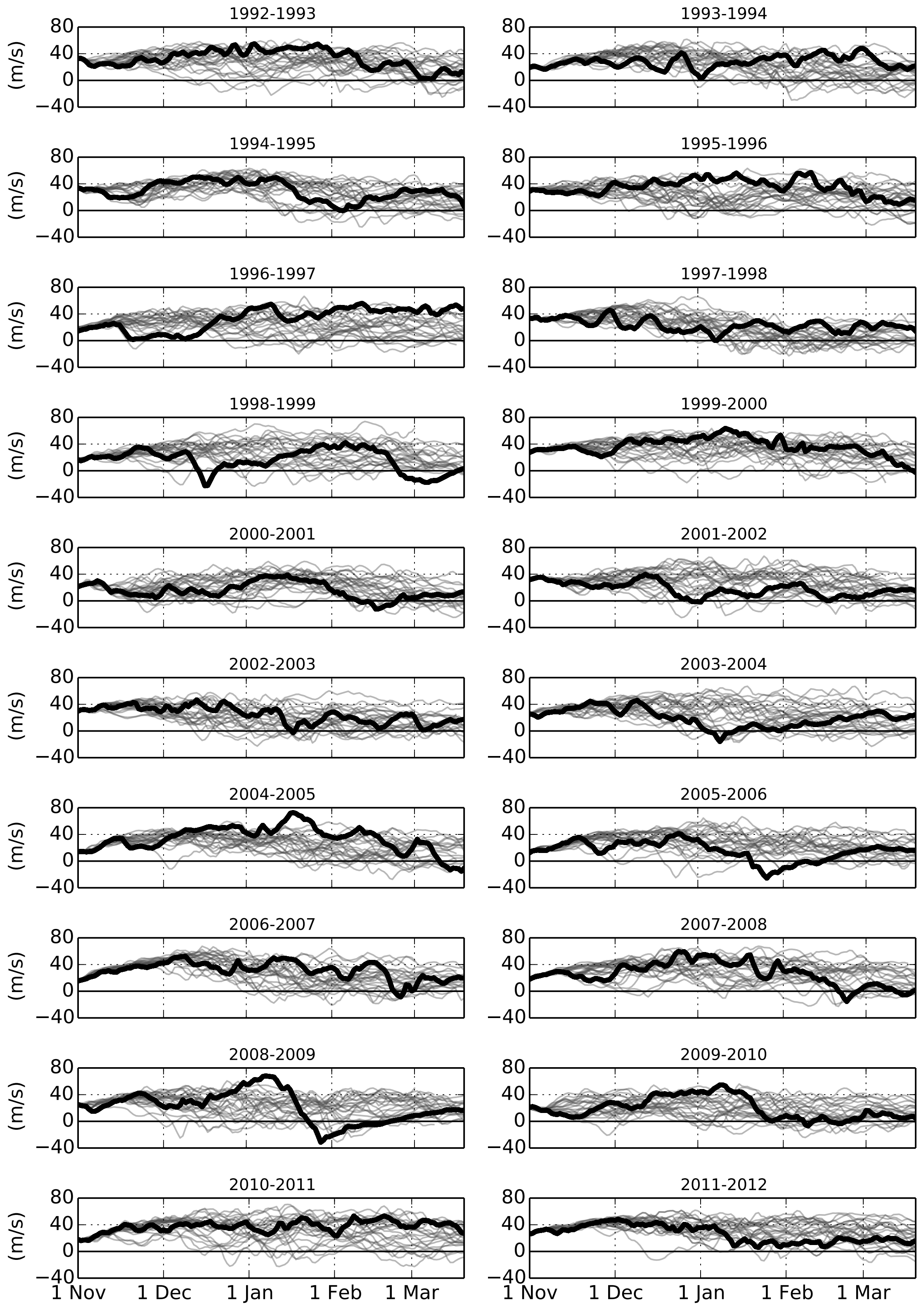

FIGURE 5.1: Timeseries of zonal-mean zonal wind in the Arctic polar vortex (60°N, 10 hPa) in the ERA-Interim reanalysis (thick black lines) and the GloSea5 ensemble hindcasts (thin grey lines). Individual ensemble members are initialised from dates centred on November 1st.

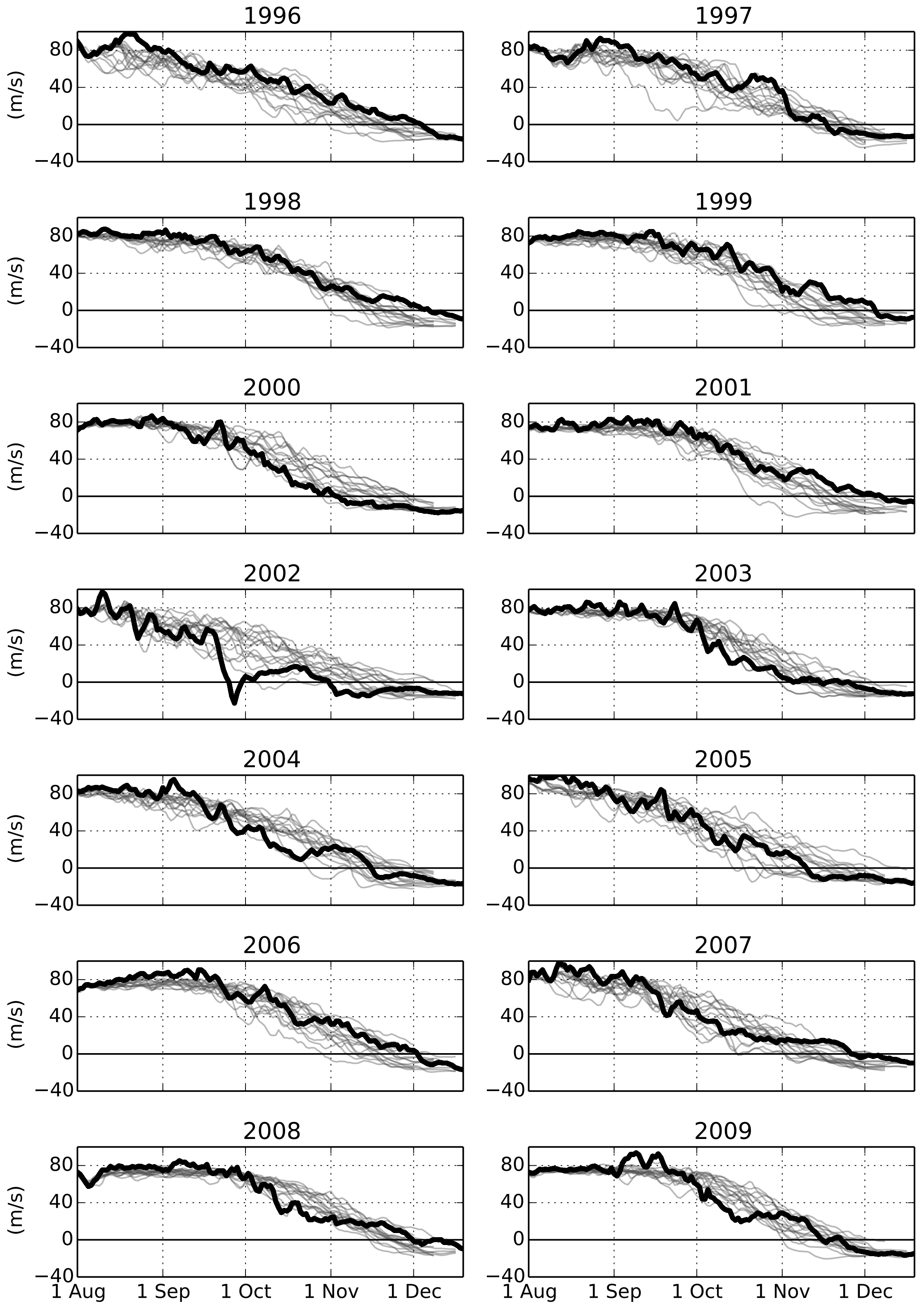

FIGURE 5.2: As figure 5.1 but for the zonal-mean zonal wind in the Antarctic polar vortex (60°S, 10 hPa), and ensemble members initialised from dates centred on August 1st.

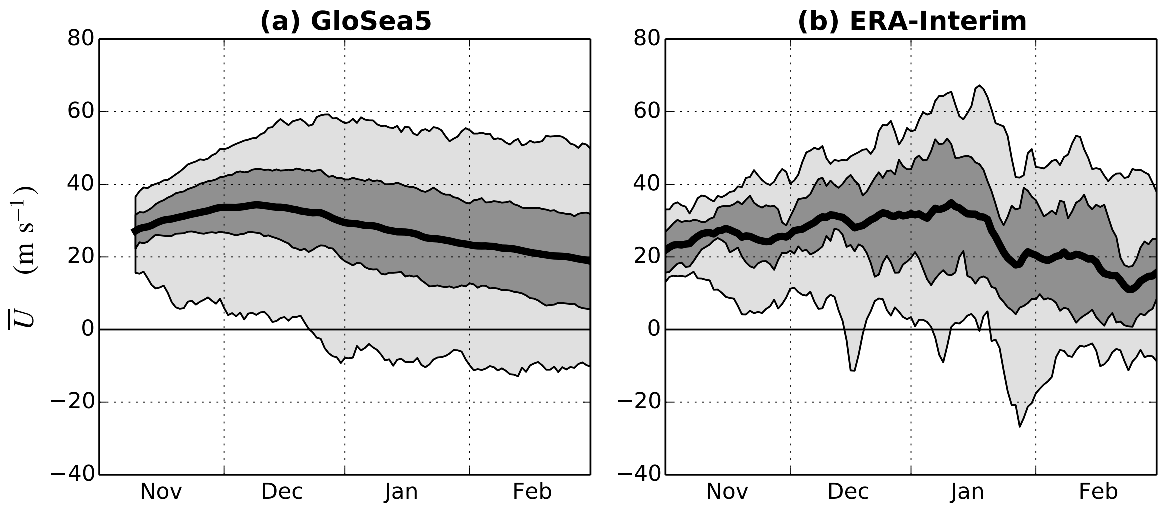

FIGURE 5.3: Time series of daily 10 hPa zonal-mean zonal wind ($\overline{u}$) at 60°N for all GloSea5 ensemble members (a) and ERA-Interim (b) from 1992–2011. The thick black line indicates the mean, dark grey shading the interquartile range and light grey the 95th percentile range.

5.3 Northern Hemisphere results

The climatology of Arctic stratospheric polar vortex winds in the GloSea5 hindcasts is compared to the ERA-Interim reanalysis climatology in Figure 5.3. As in Figure 5.1, the strength of the stratospheric polar vortex is measured by the zonal-mean zonal wind ($\overline{u}$) at 60°N and 10 hPa. The composite for the GloSea5 hindcasts is formed from all the individual ensemble members over the winters 1992/1993–2011/2012 (a total of 480), while that from ERA-Interim is a composite of the same 20 winters. It can be seen that the mean, interquartile range and 95th percentile range of the GloSea5 values agree well with the ERA-Interim values, although the ERA-Interim values are noisier as would be expected from a sample size consisting of fewer years.

Figure 5.4 shows the joint distribution of the apsect ratio and centroid latitude moment diagnostics calculated over DJF from the GloSea5 hindcasts, along with the equivalent values from ERA-Interim. These diagnostics have been calculated from geopotential height following the method described in Chapter 3. Both aspect ratio and centroid latitude distributions closely match those of ERA-Interim, and the joint

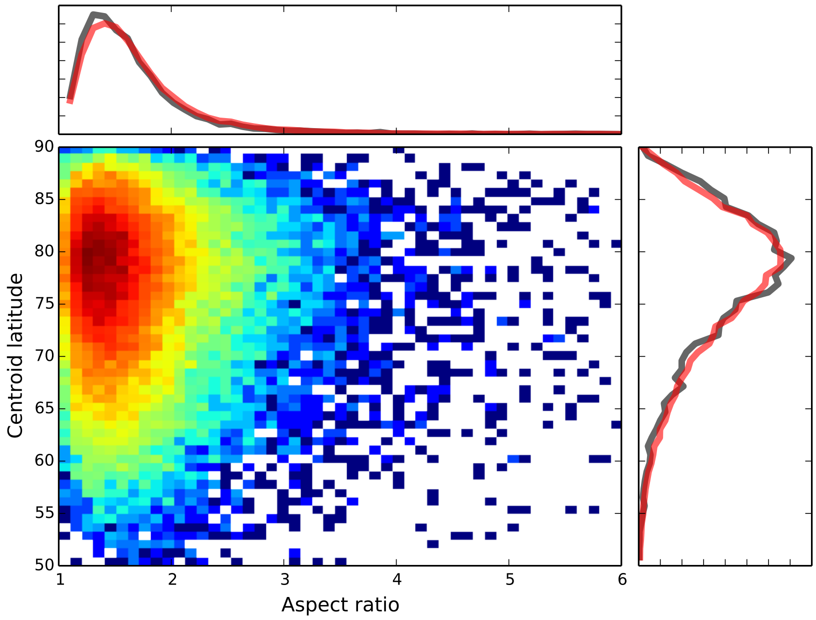

FIGURE 5.4: Distribution (see Figure 4.1) of the centroid latitude and aspect ratio diagnostics derived from geopotential height for all GloSea5 ensemble members (red lines) and ERA-Interim (grey lines) over 1992–2011. The joint distribution is plotted with a logarithmic colour scale such that red represents the densest regions.

distribution shows the characteristic triangular shape which is related to the occurrence of split and displaced vortex events. Together, Figures 5.3 and 5.4 demonstrate that GloSea5 accurately simulates the mean state and variability of the stratospheric polar vortex in the mid-stratosphere.

The GloSea5 hindcast predictions of interannual variability of the NH stratospheric polar vortex winds are shown in Figure 5.5. Anomalies are defined from the relevant climatology of either GloSea5 or ERA-Interim. For GloSea5, this climatology is calculated from the mean of each day across all ensemble members in all years, while for ERA-Interim the climatology is the mean of each day, smoothed with a 30-day running mean (in order to account for its increased noise due to the reduced sample size). Results are shown for DJF averages, corresponding to a 1 month average lead time. The correlation between the GloSea5 ensemble mean and ERA-Interim is $r = 0.24$ which is not statistically significant at the 95% level (under the null hypthesis that the two timeseries are uncorrelated). Significance is calculated using a two-tailed

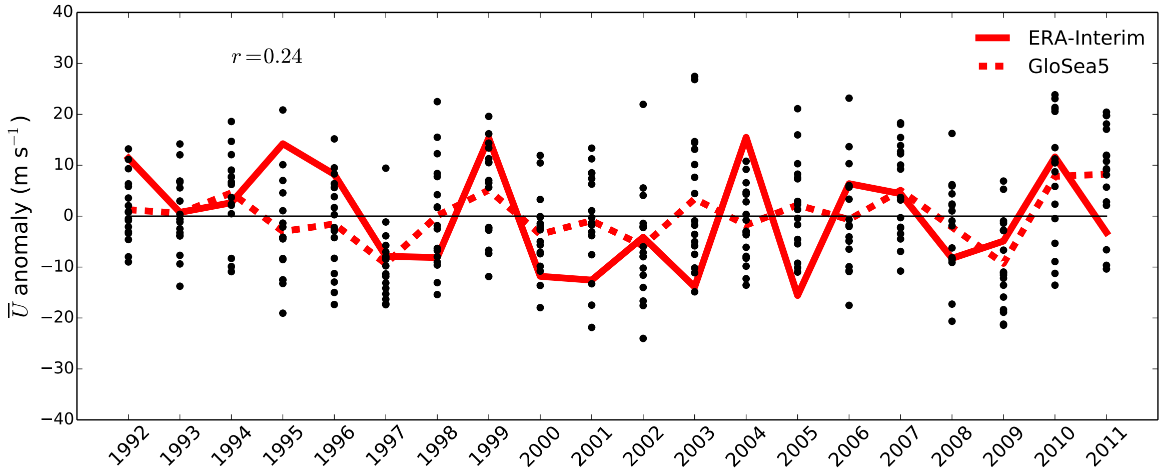

FIGURE 5.5: DJF mean anomalies of $\overline{u}$ at 60°N and 10 hPa for the GloSea5 ensemble mean and ERA-Interim. Black dots represent individual ensemble members. The correlation between the GloSea5 ensemble mean and ERA-Interim is $r = 0.24$, which is not statistically significant from zero. Years refer to the year of the initialisation date.

bootstrap test, whereby the percentile of the observed correlation is calculated from the distribution of correlations of a large number ($\sim$ 10 000) of pairs of time series formed by re-sampling with replacement from the original time series. As elswhere in this thesis, these significance tests are used because they make fewer assumptions about the underlying structure of the data than parametric tests [*Wilks*, 2006].

Although no significant skill is found in the prediction of the seasonal mean strength of the stratospheric polar vortex, it might nonetheless be the case that skilful predictions of SSW events can be made. This is assessed using receiver operating characteristic (ROC) curves, a standard method in forecast evaluation, particularly of binary events [e.g., *Wilks*, 2006]. In order to calculate the ROC curve, the following procedure is followed:

1. For each ensemble member in each year, determine whether an SSW occurs (winters with one SSW and two SSWs are treated the same).

2. Select a threshold for the prediction of an SSW (e.g., 60% of ensemble members forecast an SSW).

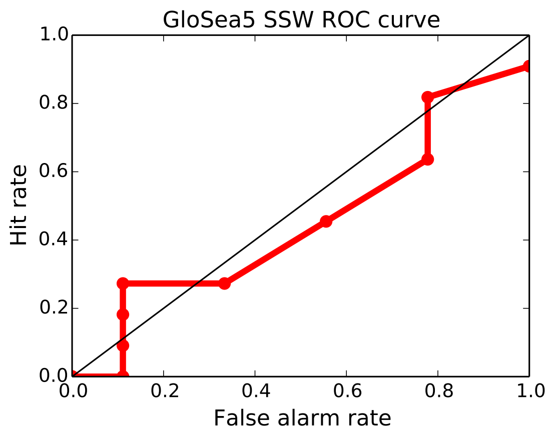

FIGURE 5.6: Receiver operating characteristic (ROC) curve for the prediction of SSW events during DJF for 1992–2011. SSWs are defined as a reversal to easterly $\overline{u}$ at 60°N 10 hPa, and must be separated by at least 30 days.

3. For the given threshold determine the fraction of years for which a SSW was correctly predicted ("hit rate") and the fraction for which a SSW was predicted but none occurred ("false alarm rate").

4. Repeat the steps 2-3 for a range of thresholds from 0-100%.

In a skilful system the ROC curve should indicate a higher hit rate than false alarm rate, bending towards the upper left corner of the graph, while a random forecast will pass along the 1-1 line.

Figure 5.6 shows the ROC curve for SSWs during DJF, determined by the traditional reversal of $\overline{u}$ at 60°N 10 hPa, with the additional criterion that events must be separated by at least 30 days. It can be seen that the calculated ROC curve lies close to the 1-1 line indicating little skill in these predictions. This is despite quite large variations in the fraction of ensemble members predicting SSWS; from 47% (winter 2001–2002) to 100% (winter 1997–1998).

A similar analysis for the prediction of split and displaced vortex events is shown in Figure 5.7. These events have been calculated from the geopotential height-derived moment diagnostics, using the same procedure as decribed in Chapter 3. Again, little

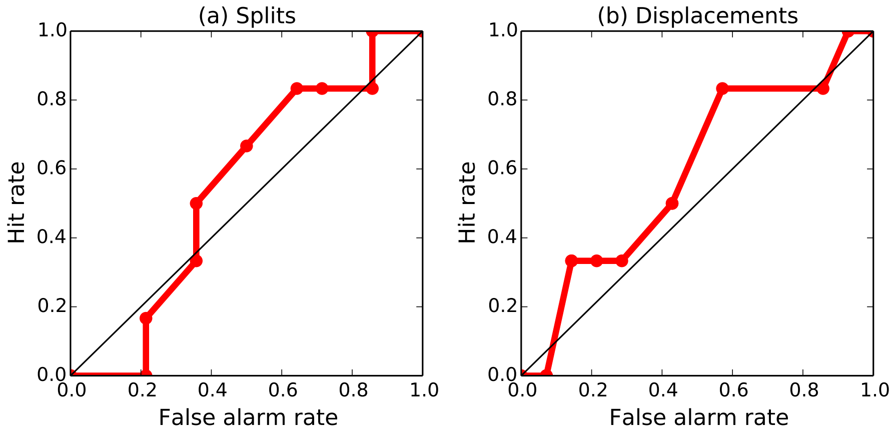

FIGURE 5.7: Receiver operating characteristic (ROC) curves for the prediction of split (a) and displaced (b) vortex events in all GloSea5 ensemble members over 1992–2011. Split and displaced vortex events are identified from geopotential height-derived moment diagnostics using the method described in Chapter 3.

skill can be seen in the predictions of these events, with both ROC curves lying close to the 1-1 line. It should be noted that GloSea5 predicts a high frequency of split and displaced vortex events (8.5 events/decade; 3.5 split, 5.0 displaced), compared to the observed value of 7 events/decade. This is perhaps not surprising given the HadGEM3 model used in GloSea5 is of the same family as HadGEM2-CC, which was found to have the highest frequency of events in Chapter 4.

The evolution of NH hindcast skill as a function of lag and height is evaluated in Figure 5.8. This shows the correlation of ERA-Interm and GloSea5 ensemble mean polar cap (60-90°N) average geopotential height anomalies (Z'; which is highly correlated with the NAM [*Kushner*, 2010]). Values of Z' are smoothed with a 30-day running mean before correlations are calculated, and plotted such that values for the 15th December represent the correlation of the ERA-Interim and GloSea5 ensemble mean December mean values (without this smoothing, a noisier but similar pattern of correlations is seen). Geopotential height data on several levels in the stratosphere is only available in the shorter 14-year, 15-member ensemble, and so the analysis in this figure is limited to these data. Between 1st-9th November the ensemble mean is taken as the average of

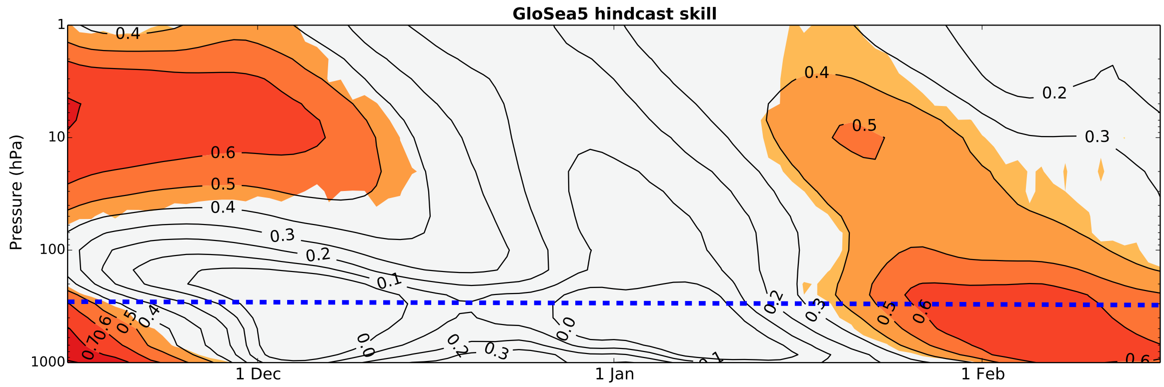

FIGURE 5.8: Correlation of GloSea5 ensemble mean polar cap (60-90°N) geopotential height anomalies (Z') with ERA-Interim values from 1996–2009, as a function of time and height. All values are smoothed with a 30-day running mean before correlations are calculated. The contour interval is 0.1 and all colored regions are greater than zero at the 95% confidence interval, using a bootstrap test at each time at height. The blue dashed line indicates the approximate polar cap mean tropopause level [*Wilcox et al.*, 2012].

the 10 initialised ensemble members, and the average of all 15 ensemble members is used after this date.

In mid-November, significant correlations are seen in Figure 5.8 in both the stratosphere and troposphere, as would be expected from the initialisation of the hindcasts from ERA-Interim data (skill is seen to rapidly decay in the tropopause region, however). This skill persists for longer in the stratosphere than the troposphere, owing to the longer timescales of stratospheric variability [e.g., *Simpson et al.*, 2011], however, by mid-December no significant correlations are seen in the stratosphere or troposphere.

Significant correlations return in the troposphere in late Janurary/February, and to a lesser extent in the stratosphere. This result is similar to that for the NAO prediction in the same system, which has a greater skill in February than January (Adam Scaife, Met Office Hadley Centre, personal communication, 2013). A possible explanation for this behaviour is due to the influence of ENSO, which has been determined to have the greatest effect on the NH extratropics during late Janurary and February [*Bell et al.*, 2009]. If indeed the influence of ENSO is important, it is not clear whether this arises from a tropospheric or stratospheric pathway, although the fact that tropospheric skill is greater than stratospheric may suggest the tropospheric pathway is more important.

Detailed analysis of the mechanisms behind the NH surface skill is beyond the scope of this chapter which investigates the role of the stratosphere. Because skill in the prediction of the NH stratosphere has been demonstrated to be low, attention is now turned to the SH. However, the implications of these results are discussed further in Section 5.5.

5.4 Southern Hemisphere results

5.4.1 Stratospheric polar vortex

The climatology of Antarctic stratospheric polar vortex winds in the GloSea5 hindcasts is compared to the ERA-Interim reanalysis climatology in Figure 5.9. The strength of the stratospheric polar vortex is measured by the zonal-mean zonal wind ($\overline{u}$) at 60°S and 10 hPa, which is approximately the center of the mean position of the vortex in the mid-stratosphere. The composite for the GloSea5 hindcasts is formed from all the individual ensemble members over 1996–2009 (a total of 210), while that from ERA-Interim is a composite of all years from 1979–2010 (a total of 32 years). It can be seen that the mean of the GloSea5 hindcasts agrees very closely with ERA-Interim throughout the spring, with only a slight bias towards weaker winds in August and September. The interquartile and 95th percentile ranges of GloSea5 and ERA-Interim also agree well, although the ERA-Interim values are noisier as would be expected from a sample size consisting of fewer years.

The GloSea5 hindcast predictions of interannual variability of the Antarctic stratospheric polar vortex winds are shown in Figure 5.10(a). Anomalies are defined from the relevant climatology of either GloSea5 or ERA-Interim. For GloSea5, this climatology is calculated from the mean of each day across all ensemble members in all years, while for ERA-Interim the climatology is the mean of each day, smoothed with a 30-day running mean (in order to account for its increased noise due to the reduced sample size). Results are shown for September–November (SON) averages, corresponding to a

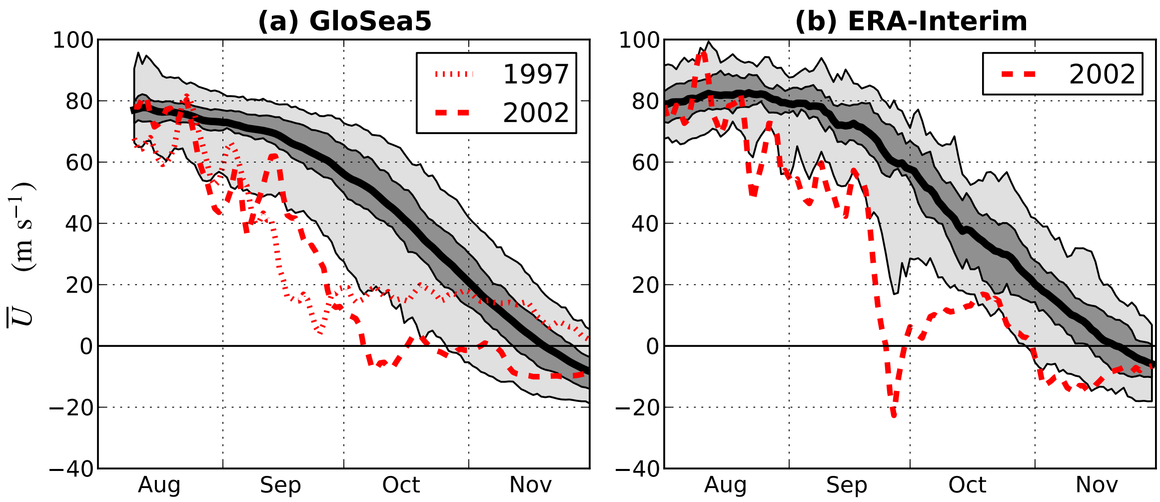

FIGURE 5.9: Time series of daily 10 hPa zonal-mean zonal wind ($\overline{u}$) at 60°S for all GloSea5 ensemble members from 1996–2009 (a) and ERA-Interim from 1979–2010 (b). The thick black line indicates the mean, dark gray shading the interquartile range and light gray the 95th percentile range. Individual time series of the ensemble member of GloSea5 for 2002 which simulated an SSW, and 1997 which simulated a near-SSW, and the year with an observed SSW (2002) are shown in red.

1 month average lead time. The correlation between the GloSea5 ensemble mean and ERA-Interim is 0.73, which is statistically significant from zero at the 99% confidence level, and has a 95% confidence interval of (0.37, 0.90). This correlation does not depend strongly on particular years; the correlation remains significant at the 95% level ($r = 0.57$) if the year 2002 (which has the greatest anomaly) is excluded. Significance is calculated using a two-tailed bootstrap test, whereby the percentile of the observed correlation is calculated from the distribution of correlations of a large number ($\sim 10,000$) of pairs of time series formed by re-sampling with replacement from the original time series. These significance tests make fewer assumptions about the underlying structure of the data than parametric tests [*Wilks*, 2006], and are used throughout this study.

The skill shown in Figure 5.10(a) cannot be accounted for by persistence of initial anomalies. In fact, there is a negative correlation between $\overline{u}$ on 9th August, when the last ensemble member is initialised, and the SON mean ($r = -0.54$). Hence, a persistence forecast would be negatively correlated with observed values. This relationship may

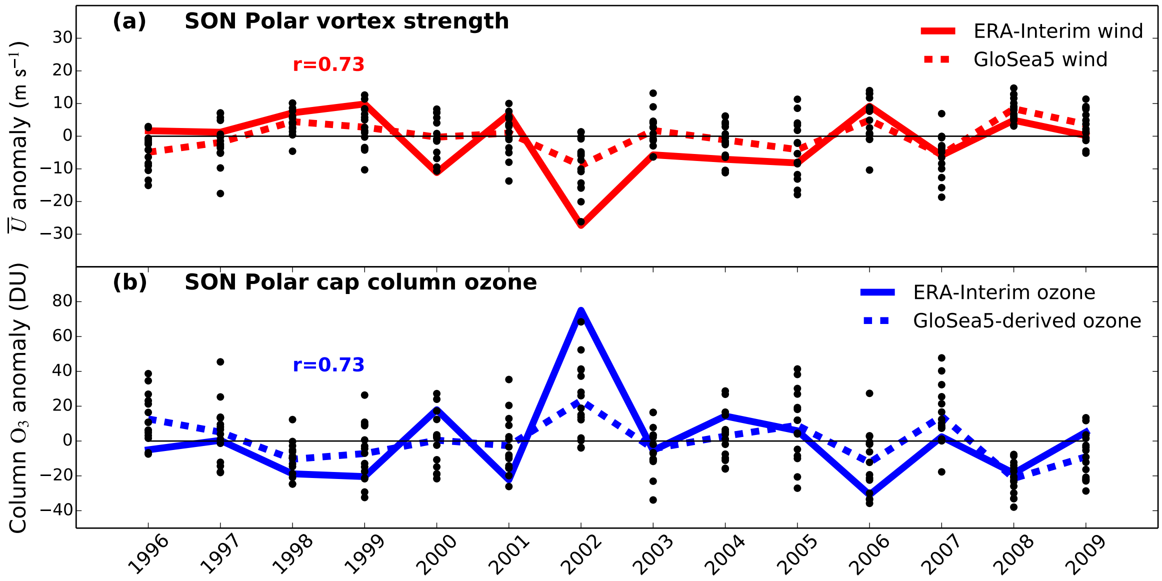

FIGURE 5.10: (a) SON mean anomalies at 10 hPa and 60°S in ERA-Interim and the GloSea5 hindcast ensemble mean. (b) SON mean polar cap averaged (60–90°S) total column ozone anomalies from ERA-Interim and those derived from the GloSea5 anomalies as described in the text. Individual ensemble members are shown as black dots. Hindcasts are initialized near 1st August.

be consistent with ideas of a pre-conditioning of the polar vortex [e.g., *McIntyre and Palmer*, 1983, Section 2.1.3]. The standard deviation of all GloSea5 ensemble members is 7.5 m s^{-1} and that of ERA-Interim is 9.7 m s^{-1} indicating that the GloSea5 ensemble spread may be too small. However, there are large uncertainties in these values due to the short hindcast period and the large 2002 anomaly.

Following *Charlton and Polvani* [2007], SSWs are defined as a temporary reversal of $\overline{u}$ at 60°S and 10 hPa, occurring before the final transition to summer easterlies (final warming). Under this definition, one SSW event was simulated in the GloSea5 hindcasts, in 2002. A similar magnitude event (in terms of departure from climatology) occurred in a 1997 ensemble member, although $\overline{u}$ did not quite become easterly. Time series of stratospheric polar vortex winds for these two events are shown in Figure 5.9(a) along with the observed 2002 SSW in Figure 5.9(b). It can also be seen in Figure 5.9(a) that 2002 has the most anomalous stratospheric polar vortex in the GloSea5 hindcasts,

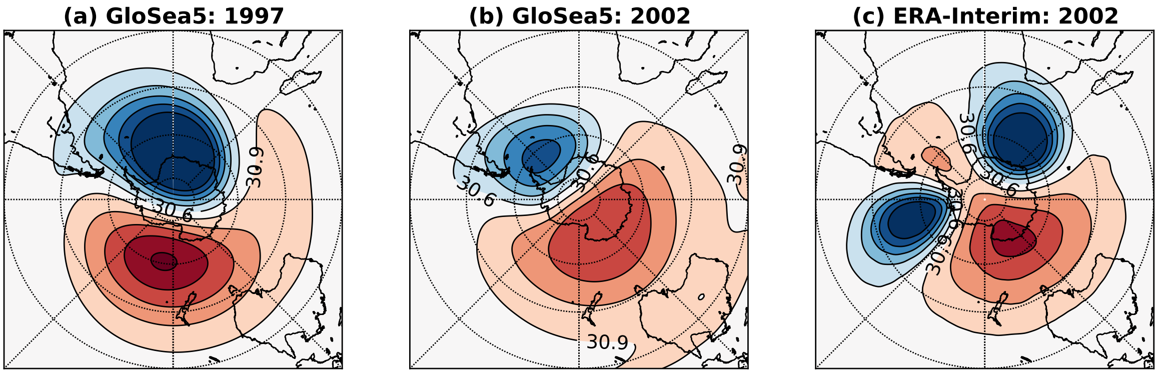

FIGURE 5.11: Geopotential height at 10 hPa on the date at which $\overline{u}$ at 60°S and 10 hPa is at its minimum value, for the two GloSea5 ensemble members which simulate a SSW (a,b), and for ERA-Interim at the central date of the 2002 SSW (c). Units are km and the contour interval is 0.3 km.

with 14 of 15 ensemble members simulating negative anomalies, and the most negative ensemble mean. It is therefore possible that an increased likelihood the 2002 event was to some degree detectable about two months in advance, although it has not been determined whether this predictability comes from a preconditioning of the vortex, as suggested by *Scaife et al.* [2005b], or the result of external forcing.

Both the SSW events simulated by GloSea5 were vortex displacement events, in contrast to the vortex splitting event which occurred in 2002 [*Charlton et al.*, 2005]. This is demonstrated in Figure 5.11, which shows geopotential height in the mid-stratosphere at the date of minimum $\overline{u}$ at 60°S and 10 hPa, for the two simulated events in GloSea5 and the observed event in ERA-Interim. A detailed quantitative anaysis using moment diagnostics was not found necessary in this case because a qualitative inspection is possible with only two events.

The timing of the final warming of the stratospheric polar vortex also has a significant effect on stratospheric temperature and ozone concentrations [*Yamazaki*, 1987], as well as on the coupling of the stratosphere to the troposphere [*Black and McDaniel*, 2007]. The predictability of these events was investigated in GloSea5, but not found to be highly significant. This is probably because the mean timing of the final warming is towards the end of the four month hindcast simulation (around 20th November at 10 hPa), and the final warming does not occur before the end of the hindcast for some ensemble members, thereby introducing a bias in the mean.

5.4.2 *Ozone depletion*

GloSea5 does not include interactive ozone chemistry, so in order to make ozone forecasts concentrations must be inferred from other meteorological variables. Total ozone quantities over the Antarctic polar cap have been found to be highly correlated with vertical EP flux poleward of 40°S [*Weber et al.*, 2011; *Salby et al.*, 2012]. EP flux diagnostics are not routinely produced directly by operational seasonal forecast systems and requires high frequency output at high spatial resolution to calculate. However, vertical EP flux dominates variability of the stratospheric polar vortex, so it may be possible to use the strength of the vortex to infer ozone quantities.

SON mean total column ozone quantities area-weighted averaged over the polar cap (60–90°S) are shown in Figure 5.12(a) for ERA-Interim and the Total Ozone Mapping Spectrometer (TOMS) satellite instrument [*Kroon et al.*, 2008]. ERA-Interim data are highly correlated with TOMS, verifying the accuracy of ERA-Interim against direct satellite measurements (TOMS values are slightly higher than ERA-Interim; this is probably because TOMS cannot make observations during the polar night). The long-term trend in polar cap total column ozone is calculated by fitting a second-order polynomial to the data. This long-term trend is due to changes in concentrations of CFCs and other ozone-depleting substances, and largely unrelated to dynamical variability. On the other hand, shorter-term interannual changes are strongly related to dynamical variability. In Figure 5.12(b) anomalies of polar cap total column ozone from the long-term trend are plotted against anomalies of the SON mean $\overline{u}$ at 60°S and 10 hPa. It can be seen that these two quantities are highly correlated ($r = -0.92$), meaning polar vortex variability explains approximately 85% of the variance of polar cap total column

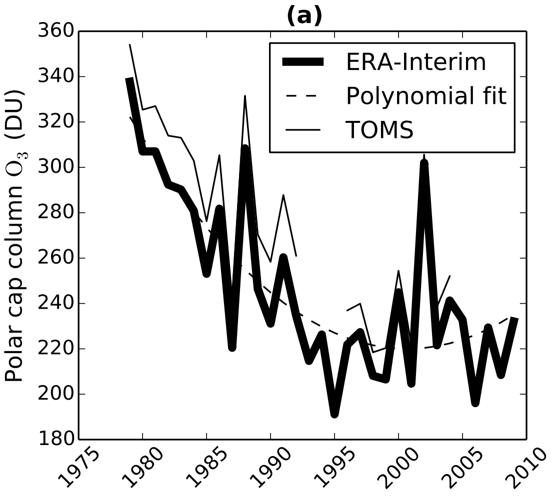

FIGURE 5.12: (a) Time series of SON mean polar cap averaged (60-90°S) total column ozone in ERA-Interim and the TOMS satellite instrument. The ERA-Interim data are fitted with a 2nd-order polynomial. (b) Anomalies of ERA-Interim column ozone from the polynomial fit plotted against SON mean anomalies at 10 hPa and 60°S for each year from 1979–2009.

ozone anomalies.

This strong correlation makes it possible to use GloSea5 forecasts of polar vortex winds to produce inferred predictions of polar cap total column ozone quantities. This is carried out by a leave-one-out cross-validation procedure [*Wilks*, 2006]; the linear regression of ERA-Interim ozone and $\overline{u}$ anomalies for all years 1979–2009 except the hindcast year is used to produce the hindcast for each ensemble member. Thus no information from the hindcast year enters the hindcast itself. Figure 5.10(b) shows the GloSea5 ozone hindcasts along with the assimilated values from ERA-Interim. The correlation between the GloSea5 ensemble mean and ERA-Interim is 0.73, which is statistically significant at the 99% level, and has a 95% confidence interval of $(0.38, 0.91)$. Errors from the regression in Figure 5.10(b) for the inferred ozone quantities for each ensemble member are small compared to the spread between ensemble members, and so are not plotted in this figure.

5.4.3 Southern Annular Mode

The SAM index in both GloSea5 and ERA-Interim is depicted as the difference between the normalized anomalies of zonally averaged mean sea-level pressure at 40°S and 65°S [*Gong and Wang*, 1999]. These anomalies are calculated from the respective climatologies of GloSea5 and ERA-Interim. The ERA-Interim SAM index calculated in this way is also highly correlated with other measures of the SAM, such as the station-based index of *Marshall* [2003]. The GloSea5 hindcast skill for the prediction of the seasonal (SON) mean SAM index is shown in Figure 5.13. The correlation of the GloSea5 ensemble mean and ERA-Interim is 0.64, which is statistically significant at the 95% level, and has a 95% confidence interval of (0.18,0.92) confirming skillful prediction of the SAM at 1 month average lead times. This is similar to the value for the December–February (DJF) NAO correlation skill of 0.62 found by *Scaife et al.* [2014] in the same seasonal forecast system. The 1-year lag autocorrelation of the SON mean SAM is negative ($r = -0.36$), and accounting for this by sampling pairs of consecutive years in the bootstrap test leads to a narrower confidence interval than presented above. The variability of the SAM simulated by GloSea5 is broadly realistic with a standard deviation of all ensemble members of 0.98 compared to 0.90 in ERA-Interim over the same period.

The SAM is strongly related to surface temperatures over much of the SH extratropics. Figure 5.14(a) shows the correlation of the SON mean SAM from ERA-Interim over 1996–2009 with SON mean gridded station-based surface temperature data from the HadCRUT4 data set [*Morice et al.*, 2012]. The HadCRUT4 data set has been chosen to demonstrate the relationship between the SAM and surface temperature because of the scarcity of temperature observations in the Southern Hemisphere, meaning reanalysis data is poorly constrained in many regions. The same relationship between surface temperatures and the SAM is shown for the GloSea5 ensemble mean in Figure 5.14(b). Many of the observed correlations are reproduced in the hindcasts, such as the opposite

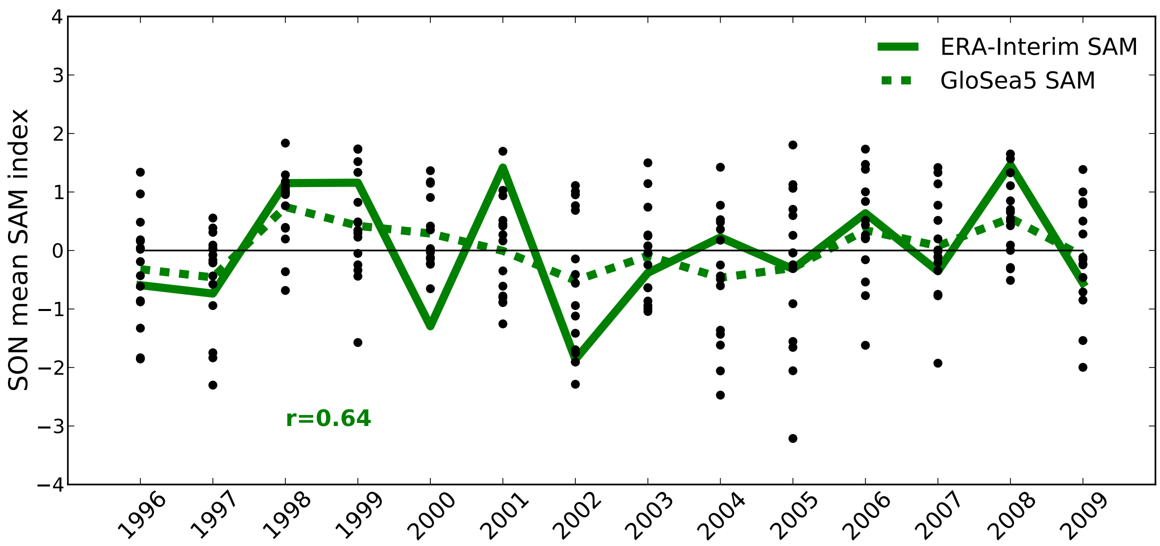

FIGURE 5.13: SON mean Southern Annular Mode (SAM) index in individual GloSea5 hindcast ensemble members (dots), ensemble mean (dashed green curve) and ERA-Interim (solid green curve). The SAM is calculated from mean sea-level pressure data, and hindcasts initialized near 1st August. The correlation of the ensemble mean and ERA-Interim values is 0.64, which is statistically significant at the 95% level.

signed correlations over east Antarctica and the Antarctic Peninsula/Patagonia, as well as between eastern Australia and New Zealand. These results are in agreement with *Gillett et al.* [2006] who analysed the temperature patterns associated with the SAM over the longer observational record of 1957–2005.

The GloSea5 ensemble mean SON surface temperature correlation with HadCRUT4 is shown in Figure 5.14(c). Also highlighted (black circles) are the points with the strongest observed correlations with the SAM ($|r| > 0.5$). Regions of significant positive correlations are found over east Antarctica, Patagonia, New Zealand, and eastern Australia. These are regions which also have a strong correlation with the SAM, indicating that the significant surface temperature skill is related to skill in prediction of the SAM. On the other hand, there are also some significant negative correlations in subtropical regions, which may indicate a model bias in the temperature pattern associated with the SAM in these regions.

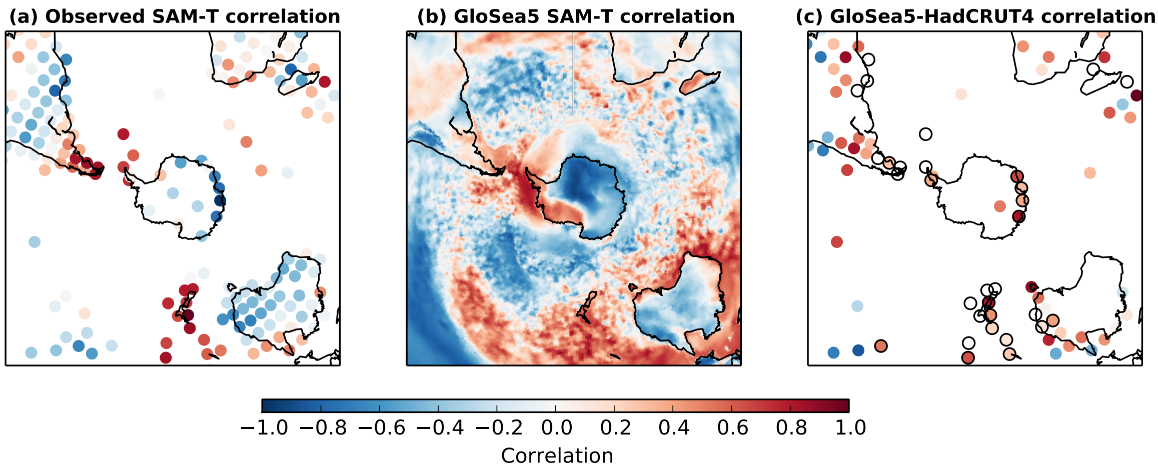

FIGURE 5.14: (a) Correlation of the ERA-Interim SON mean SAM with SON mean HadCRUT4 gridded station-based temperature observations over 1996–2009. (b) Correlation of the SON GloSea5 ensemble mean hindcast SAM with the SON hindcast ensemble mean near-surface temperature. (c) Correlation of observed SON mean HadCRUT4 and hindcast GloSea5 ensemble mean temperature. In (c) only correlations which are significant from zero at the 95% level according to a bootstrap test at each gridpoint are shown. Black circles represent points which have an observed correlation with the SAM with magnitude greater than 0.5.

5.4.4 Stratosphere-troposphere coupling

It is now investigated whether the statistically significant skill in hindcasts of the stratospheric polar vortex affects that of the surface SAM. Forecast skill as a function of lead-time and height is studied for polar cap (60-90°S) mean geopotential height anomalies (Z')[1]. Figure 5.15(a) shows the correlation of Z' in ERA-Interim with the GloSea5 ensemble mean hindcast values. Values are smoothed with a 30-day running mean before correlations are calculated, and plotted such that values for 15th September represent the correlation of the ERA-Interim and GloSea5 ensemble mean September mean values (without this smoothing, there are noisier but still significant correlations in a similar pattern). Between 1st-9th August the ensemble mean is taken as the average of the 10 initialized ensemble members, and the average of all 15 ensemble members is used after this date.

[1]Throughout the troposphere and stratosphere daily Z' is highly correlated ($r > 0.9$) with the SAM index calculated from zonal mean geopotential height [*Baldwin and Thompson*, 2009].

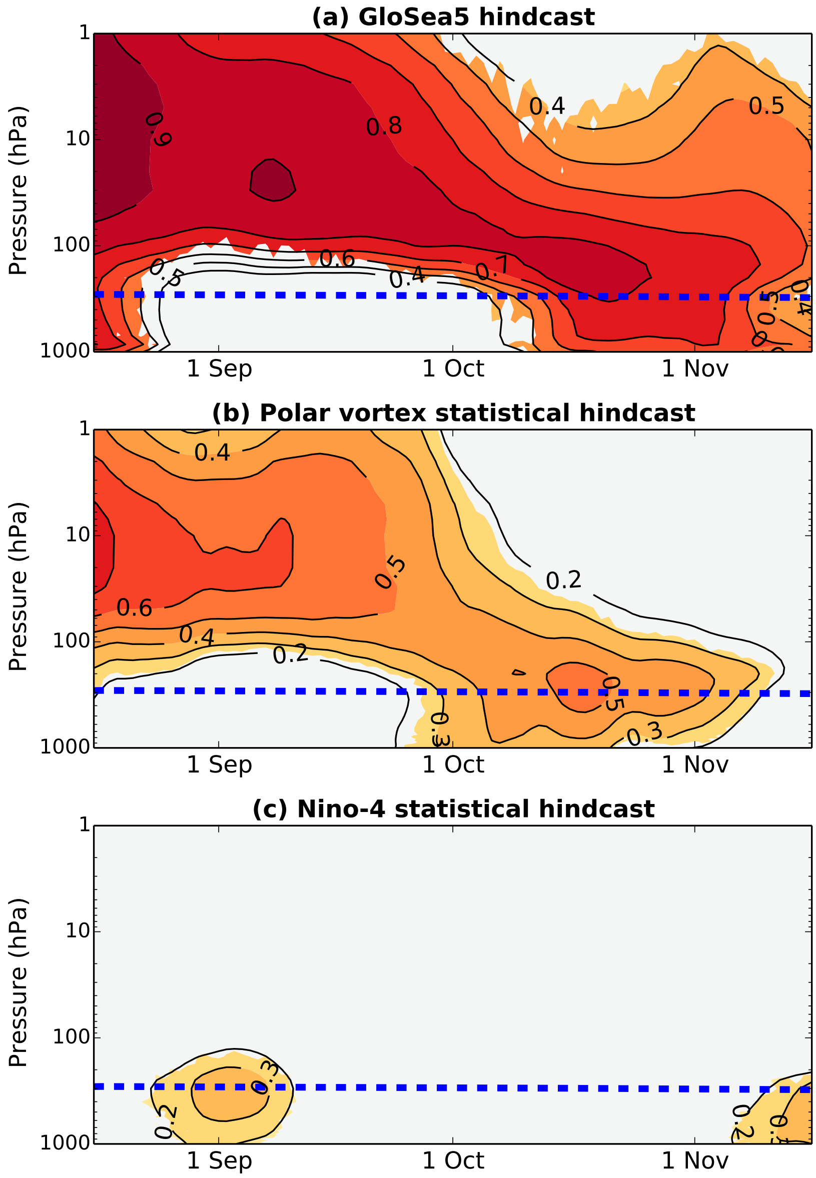

FIGURE 5.15: (a) Correlation of GloSea5 ensemble mean polar cap (60-90°S) geopotential height anomalies (Z') with ERA-Interim values from 1996–2009, as a function of time and height. (b) Correlation of ERA-Interim from 1979–2010 values with those predicted by a linear statistical model based on Z' at 10 hPa on 1st August. (c) As (b) but based on the July-mean Niño-4 index. All values are smoothed with a 30-day running mean before correlations are calculated. The contour interval is 0.1 and all colored regions are greater than zero at the 95% confidence interval, using a bootstrap test at each time at height. The blue dashed line indicates the approximate polar cap mean tropopause level [*Wilcox et al.*, 2012].

As would be expected from the initialization of GloSea5 from ERA-Interim data, correlations are high in both the troposphere and the stratosphere for the August mean, due to predictability on weather timescales. However, tropospheric and lower-stratospheric skill rapidly decays and becomes statistically insignificant throughout September. In contrast, stratospheric correlations remain statistically significant throughout the hindcast simulation, and as high as 0.8 through to mid-October (corresponding to a 2 month lead time).

Importantly, the region of high levels of stratospheric skill descends with time and is present at the tropopause at the same time as a re-emergence of significant tropospheric skill in mid-October. This re-emergence cannot be accounted for by the persistence of tropospheric anomalies, so must be the result of the effect of another predictable signal on the extratropical tropospheric circulation. An obvious candidate for such a signal is the polar stratosphere, since this remains predictable throughout the hindcast period. The re-emergence of tropospheric skill also occurs at the same time as the strongest observed coupling between the stratosphere and troposphere found in other studies [e.g., *Thompson et al.*, 2005; *Simpson et al.*, 2011].

In order to determine the stratospheric influence on tropospheric skill, a simple statistical forecast model is formed, which has as its only input the initial conditions of the Antarctic stratosphere. A leave-one-out cross validation (LOOCV) [*Wilks*, 2006] procedure is employed as follows:

i. Remove the predictand year, i, from the set of all N years, leaving $N-1$ predictor years.

ii. Calculate the linear regressions of Z' at 10 hPa on 1st August with Z' at all other times and heights using the $N-1$ predictor years.

iii. Given the value of Z' at 10 hPa on 1st August for year i (the predictand year), use the linear regressions to produce a forecast for Z' at all other times and heights for this year.

iv. Repeat the above steps for $i = 1, 2, \ldots, N$ to produce N forecasts, each with slightly different regression coefficients.

The method ensures that no information from the predictor year enters the regression, and provides an estimate of the predictability of an unknown year given the available observations. Here, ERA-Interim values are used from 1979–2009; giving $N = 32$ years.

Figure 5.15(b) shows the correlation of 30-day running means of these statistical hindcasts with ERA-Interim values. As might be expected, skill is initially high in the mid-stratosphere but not the troposphere. As with the GloSea5 hindcasts, the region of high skill descends with time, and statistically significant correlations re-emerge in the troposphere throughout October. This demonstrates that skillful forecasts of the Antarctic troposphere during October can be produced based only on knowledge of Z' in the mid-stratosphere on 1st August. It also suggests that the re-emergence of tropospheric skill in the GloSea5 hindcasts in October is likely to be caused by predictable stratospheric anomalies which descend with time.

However, it is also possible that a third factor both influences the 1st August stratosphere and the October and November tropsophere. ENSO may be such a factor, since it has been shown to influence both the surface SAM [*Lim et al.*, 2013] and the polar stratosphere [*Hurwitz et al.*, 2011]. The influence of ENSO is therefore assessed using the same leave-one-out cross-validation procedure, and shown in Figure 5.15(c). The input to the statistical model is the July mean Niño-4 index (sea-surface temperatures averaged over 5°S-5°N, 160°-150°W) from the HadISST1 data set [*Rayner et al.*, 2003]. Similar results are obtained using the July mean Niño-3.4 index or Southern Oscillation Index. The Niño-4 index-based statistical hindcasts show some significant tropospheric correlations around 1st September and in November, but not during October. Hence, ENSO cannot account for the October re-emergence of tropospheric skill in the GloSea5 hindcasts, at least in this statistical model.

Importantly, the longer 32-year (1979–2010) period of the ERA-Interim reanalysis (rather than the 14-year (1996–2009) period of the GloSea5 hindcasts) is used for the statistical analysis presented in Figures 5.15(b) and (c). The correlation between both the 1st August Z' at 10 hPa and the July mean Niño-4 index with the SON SAM is not statistically significantly different during 1996–2009 compared with 1979–2010. This was tested using a bootstrap test, which correlates subsets of 14 years from the (detrended) 32 years. Hence correlations found for the shorter period are deemed to be a marginal distribution of those over the longer period, so a more robust measure of sources of predictability can be obtained by studying the longer observational record. A more detailed justification for this choice of analysis period in the statistical hindcasts is given in Appendix 5.A.

Similar features are seen if the statistical hindcasts are repeated using the shorter period, although tropospheric skill from the polar vortex emerges later (in November), and that from Niño-4 earlier (in October). These statistical hindcasts also show lower skill than the GloSea5 hindcasts at almost all times in both the tropsophere and stratosphere, which may indicate the importance of non-linearities or the influence of other external factors which can be captured by the full dynamical model.

5.5 Discussion

5.5.1 Northern Hemisphere

The fact that hindcasts of the NH stratospheric polar vortex have been shown to be less skilful than those of the SH is not unexpected because of the much greater dynamical variability and chaotic nature of the NH. Indeed, previous studies have not found SSWs to be predictable (in a deterministic sense) beyond about two weeks [*Marshall and Scaife*, 2010; *Taguchi*, 2014]. However, given the fact that the GloSea5 hindcasts have been shown to produce skilful predictions of the DJF NAO [*Scaife et al.*, 2014], it is perhaps surprising that somewhat greater skill was not found in Section 5.3. Even if the vortex was to respond passively to NAO variability, a greater degree of skill might be expected.

A possible explanation for this may be that GloSea5 does not produce skilful forecasts of the North Pacific, so that the North Pacific and North Atlantic 'destructively interfere' in their infuence on the stratosphere. However, *MacLachlan et al.* [2014] found GloSea5 forecasts of DJF surface temperatures to have similar skill in the North Atlantic and North Pacific, as well as the surface NAM to have similar skill as the NAO. Therefore, the reason for the relative lack of skill in hindcasts of the NH stratospheric polar vortex, and whether more skilful forecasts would be possible, remains unknown. This does, however, suggest that the source of skilful DJF NAO hindcasts in GloSea5 is unlikely to be of stratospheric origin, and other model improvements such as the increased ocean resolution may be more important.

5.5.2 Southern Hemisphere

5.5.2.1 Model limitations

We have demonstrated that Antarctic total column ozone amounts are predictable up to four months in advance during the austral spring, even with a model which lacks interactive chemistry. While using such a model has the advantage of being less computationally expensive than a chemistry-climate model, there are also some drawbacks. Primarily, the model will not be able to simulate zonal asymmetries in ozone concentrations and their influence on the stratospheric circulation or the feedback between ozone concentrations and stratospheric temperatures. Both these factors have been shown to be important in driving long-term trends in the SAM as a result of ozone depletion [*Thompson and Solomon*, 2002; *Crook et al.*, 2008; *Waugh et al.*, 2009].

Perhaps more relevant for seasonal forecasts is the fact that we have not been able to determine whether the observed strong correlation between the stratospheric circulation and Antarctic ozone concentrations is dominated by a chemical or dynamical mechanism.

If the relationship is dominated by a chemical mechanism, whereby enhanced descent over the pole inhibits the activation of ozone-depleting substances, we would expect the correlation to weaken as concentrations of these substances return to pre-industrial levels. Accurate forecasts of ozone with models lacking interactive chemistry would then not be possible. On the other hand, if the mechanism is largely dynamical, whereby transport of ozone-rich air from the tropics is the important factor, we would not expect the relationship to change in time. Although a study to distinguish these mechanisms has been carried out for chemistry-climate models [*Garny et al.*, 2011], it has not been possible to do so in observations. In either case, we do not expect the relationship to break down soon, as concentrations of ozone-depleting substances are not projected to return to 1980 levels until the late 21st century [*WMO*, 2011].

5.5.2.2 Statistical significance and ensemble size

The correlation skill of 0.64 (95% confidence interval: [0.18,0.92]) for the SON mean SAM in the GloSea5 hindcasts is greater but not inconsistent with that found by *Lim et al.* [2013]. They report a correlation of 0.40 for the SON mean SAM from 1st August initialized forecasts over 1981–2010 using the Predictive Ocean and Atmosphere Model for Australia, version 2 (POAMA2). Over the comparable period of 1996–2009, they find a correlation of 0.54 (Harry Hendon, Australian Bureau of Meteorology, personal communication, 2014). Significantly, POAMA2 has only two model levels in the stratosphere, and so may be unable to simulate the stratosphere-troposphere coupling described here. *Lim et al.* [2013] attribute their results to the influence of ENSO through a tropospheric teleconnection. This is not inconsistent with our result shown in Figure 5.15(c), since we find significant tropospheric predictability from ENSO during November, the same time that *Lim et al.* [2013] find the strongest correlation between ENSO and the SAM. The lack of discrepancy between these two systems despite their different stratospheric resolutions may be a result of the ENSO/SAM connection being too weak in GloSea5, or simply that the relatively short hindcast period used here prevents a statistically significant difference being detected.

Despite this significant correlation skill in hindcasts of the SAM, it is clear from Figure 5.13 that the standard deviation of the GloSea5 ensemble mean SAM is much less than that of observations. The signal-to-noise ratio (ratio of the standard deviation of the ensemble mean to that of all ensemble members) is just 0.4. For a 'perfect' forecast system (one in which observations are indistinguishable from an ensemble member), the signal-to-noise ratio, s, and correlation, r, are directly related by

$$r = \frac{s^2}{\sqrt{(s^2+1)(s^2+n^{-1})}}, \tag{5.1}$$

where n is the ensemble size [*Sardeshmukh et al.*, 2000; *Kumar*, 2009]. Hence, given the value of $s = 0.4$, the expected correlation would be just 0.3, rather than the 0.64 found. This discrepancy can be explained from the fact that the average correlation between ensemble members and observations (0.27) is much greater than that between pairs of ensemble members (0.13). A similar but smaller difference is also found for the stratospheric polar vortex forecasts, and this is also observed by *Scaife et al.* [2014] for the NAO in the same system. These results mean that individual ensemble members have a smaller predictable signal than observations. This effect was recently discussed by *Eade et al.* [2014], who proposed a rescaling of the ensemble mean to have the same variance as the predictable component of the observed variance (which can be estimated by $\sigma^2_{\text{obs}} r^2$, where σ^2_{obs} is the observed variance). However, this procedure is most applicable to forecasts where the scaling can be determined from hindcasts, so that information from the observations does not enter the forecasts themselves. Furthermore, this rescaling does not affect correlation skill scores, and so it is not applied in the current analysis.

Given the above result, it might be expected that more skilful predictions could be obtained with a larger ensemble size. To illustrate the variation of hindcast skill with ensemble size we systematically sample smaller sets of forecasts from the full 15

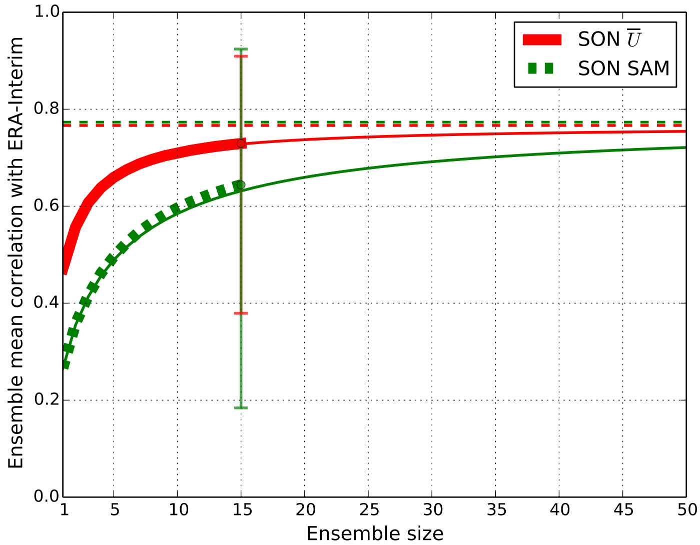

FIGURE 5.16: GloSea5 ensemble mean correlation with ERA-Interim as a function of ensemble size for the SON mean $\bar{u}$ at 10 hPa and 60°S and SON mean SAM (thick lines). A theoretical estimate of the variation of correlation with ensemble size is shown in each case (thin solid lines), along with its asymptote for an infinite sized ensemble (dashed lines). Error bars represent the 95% uncertainty range for the correlation of the full 15-member ensemble, calculated using a bootstrap test.

members for each year, following the method of *Scaife et al.* [2014]. This is repeated many times ($\sim$ 10 000) and an average value for a given sample size calculated. This variation of correlation skill with ensemble size for both the SON mean SAM and stratospheric polar vortex winds is shown in Figure 5.16. These curves closely follow the theoretical relationship of *Murphy* [1990], which relies only on the mean correlation between pairs of ensemble members, $\langle r_{\mathrm{mm}} \rangle$, and the mean correlation between individual ensemble members and observations, $\langle r_{\mathrm{mo}} \rangle$, given by

$$r = \frac{\langle r_{\mathrm{mo}} \rangle \sqrt{n}}{\sqrt{1 + (n-1)\langle r_{\mathrm{mm}} \rangle}}. \tag{5.2}$$

These curves are shown in Figure 5.16, along with their asymptote for an infinite sized ensemble. This shows that the stratospheric forecasts cannot be greatly improved with a larger ensemble size in the current system, but greater correlation scores of the SAM could be achieved with an ensemble size near 30. Although the large uncertainty range does not allow a strong statement about potential predictability, the asymptote near 0.8 is similar to that found by *Scaife et al.* [2014] using a longer hindcast and greater

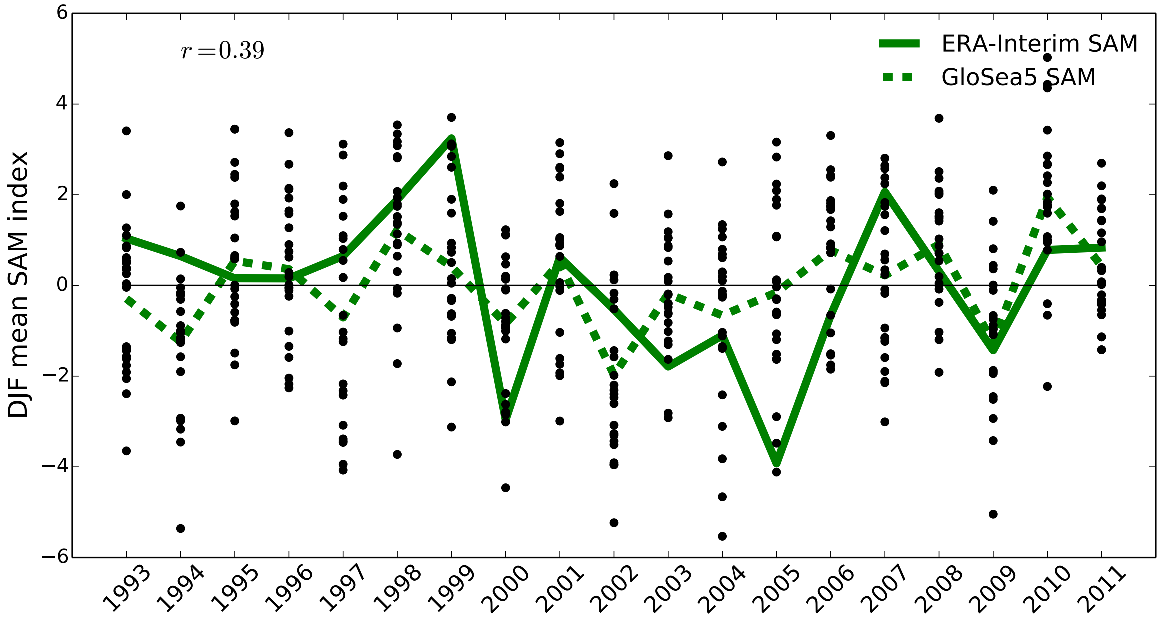

FIGURE 5.17: DJF mean Southern Annular Mode (SAM) index in individual GloSea5 hindcast ensemble members (dots), ensemble mean (dashed green curve) and ERA-Interim (solid green curve). The SAM is calculated from mean sea-level pressure data, and hindcasts initialised near 1st November. The correlation of the ensemble mean and ERA-Interim values is 0.39.

ensemble size for the DJF NAO.

5.5.2.3 Application to other seasons

The dynamics of other seasons are different to those of the austral spring, so results presented here for SON do not imply significant skill in prediction of the SAM at other times. Indeed, the 1-month lead time ensemble mean correlation of the DJF SAM with ERA-Interim is lower than that for SON at $r = 0.39$ (95% confidence interval: [0.15,0.63]), as shown in Figure 5.17. The low signal-to-noise ratio found in Figure 5.13 for SON can also be seen in Figure 5.17.

Shaw et al. [2010] found that the strongest downward wave coupling between the stratosphere and troposphere is present during September to December in the SH. They attribute this to the fact that the lower stratospheric vortex is westerly during this time, but the mid-upper stratospheric vortex is easterly (because the final warming occurs first in the upper stratosphere) and acts as a relecting surface for planetary waves. Following the final warming in the lower stratosphere, *Shaw et al.* [2010] find wave coupling to be much weaker. *Shaw et al.* [2011] extended this analysis to also demonstrate that the dynamical influence of stratospheric ozone depletion on the troposphere through wave coupling is greatest during September-December.

On the other hand, separate studies have found that the largest tropospheric signals associated with stratospheric ozone depletion occur later, in DJF [*WMO*, 2011]. This may seem to contradict the findings of *Shaw et al.* [2010], but the two results can be reconciled if a different mechanism is dominant at this later time. Indeed, as well as an effect on the dynamical coupling between the stratosphere and troposphere, *Grise et al.* [2009] proposed that stratospheric ozone depletion can perturb radiative heating rates in the troposphere which can, in turn, trigger changes in tropospheric dynamics. They used a radiative model to investigate this effect and, importantly, found the largest influence on polar tropospheric temperatures to occur during DJF. A possible physical explanation for this is that it is only after the final warming, when the ozone depleted polar stratospheric air is mixed with lower latitudes, that the radiative effect on the troposphere is significant. While this was only an idealised study which lacked tropospheric dynamics, it may suggest a reconciliation with dynamical coupling being strongest from September-December and radiative from December-February.

The time dependency of SH stratosphere-troposphere coupling is further investigated in Figure 5.18. This shows lag-height correlations of polar cap Z' in the mid-stratosphere (10 hPa), lower-stratosphere (70 hPa) and surface (1000 hPa) at the first of each month from August–January using ERA-Interim data (1979–2010). As in Figure 5.15, values are smoothed with a 30-day running mean before correlations are calculated. Mid-stratosphere-leading significant correlations with the October-November troposphere are seen from 1st August (Figure 5.18(a)), as also shown in Figure 5.15. Furthermore the strongest negative lag correlations of the surface with the stratosphere occur at 1st November (Figure 5.18(l)). This supports the result of *Shaw et al.* [2010]

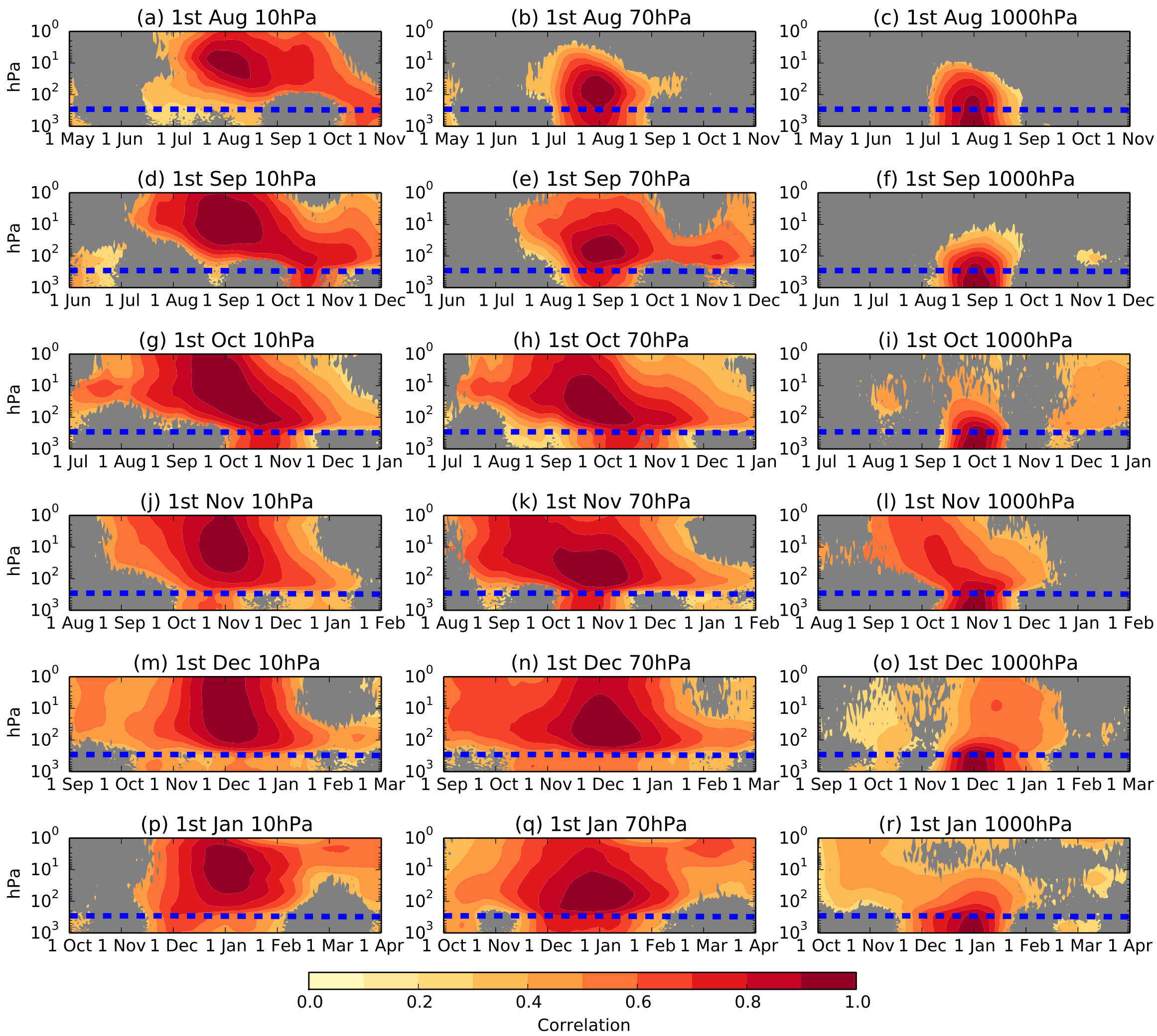

FIGURE 5.18: Correlation of ERA-Interim (1979–2010) Z' at 10 hPa, 70 hPa, and 1000 hPa with Z' at other times and lags. Values are smoothed with a 30-day running mean before correlations are calculated, and colours represent correlations that are 95% significant from zero according to a bootstrap test.

that September-December is the time of strongest stratosphere-troposphere dynamical coupling.

Similar, but weaker lag correlations are seen at 1st Janurary (Figure 5.18(r)). This is unlikely to be due to dynamical coupling since it comes after the stratospheric final warming, and so may be a result of the radiative effect described above. It is important to note that GloSea5 does not contain interactive ozone chemistry, so the radiative effects of ozone variability will not be captured by the model. This may explain, to some extent, the reduced skill in the prediction of DJF SAM compared to the SON SAM, since the predictable effects of the stratosphere on the troposphere are not captured during DJF. Consequently, more skilful forecasts of the DJF SAM may be possible with a model including interactive ozone chemistry.

5.6 Conclusions

Motivated by the results of Chapters 3 and 4, we have analysed the predictability of the polar stratosphere and its influence on the troposphere in a set of hindcasts produced by a stratosphere-resolving seasonal prediction system. Analsysis has focussed on the NH for the boreal winter (DJF) and SH for the austral spring (SON), with forcasts initialised at a 1-month lead time.

No statistically significant skill was found in the prediction of the seasonal mean strength of the NH stratospheric polar vortex, or the occurrence of SSWs, split or displaced vortex events. This result may be surprising given that the same system produces skilful hindcasts of the winter NAO, which is known to influence the polar stratosphere. It does, however, suggest that this NAO skill is unlikely to be influenced by the stratosphere and may be attributable to other model improvements.

On the other hand, skillful prediction of the interannual variability of the spring Antarctic stratospheric polar vortex was found at seasonal lead times. This includes capturing an increased likelihood of the 2002 SSW which is the most extreme year in the GloSea5 ensemble mean and has the only ensemble member in 14 years which simulates a SSW (although another is close to simulating a SSW in 1997). Because this variability is observed to be closely correlated with Antarctic column ozone amounts, we are able to perform skillful predictions of interannual variability in Antarctic ozone depletion.

We also find significant skill in hindcasts of the spring mean SAM index. By studying the variation of this skill with time and height, we suggest that this skill is influenced by stratospheric anomalies which descend with time and are coupled with the troposphere in October and November. In fact, the influence of the stratosphere is such that skillful statistical predictions of the October SAM can be made using only information from 1st August in the mid-stratosphere.

Assuming that the 14 year period studied here is representative of future years, these results suggest that it may now be possible to make skillful seasonal forecasts of interannual variations in springtime ozone depletion and large scale weather patterns across the Southern Hemisphere.

5.A Choice of time period for statistical forecast

The aim of the LOOCV statistical analysis presented in Section 5.4.4 was to estimate the degree of predictability which arises from both the mid-stratosphere at the start of August and the July-mean Niño-4 index. Importantly, this analysis used the longer ERA-Interim period of 1979-2009 rather than the same 1996-2010 period over which the hindcast simulations were run. A choice of the longer period would be justafiable (and, indeed, preferable) if the relationships between these parameters and the forecast perameter (Z') are not physically different over the shorter period. That is, if the shorter period is a marginal distribution of the longer period. This is shown to be the case below.

We use the monthly Southern Oscillation Index (MSLP difference between Darwin and Tahiti, which is highly correlated with the Niño-4 index) data obtained from the Australian Bureau of Meteorology (`http://www.bom.gov.au/climate/current/soihtm1.shtml`), and a station-based SAM index from the British Antarctic Survey [*Marshall*, 2003]. For the 1996-2009 hindcast period we find the correlation of June-July SOI with Oct-Nov SAM to be $r = 0.63$. For 1979-2010 (the ERA-Interim period) $r = 0.32$. In order to justify using the shorter hindcast period (with higher correlation), it would need to be the case that the SAM/SOI correlation is statistically significantly stronger during the hindcast period than the ERA-Interim period, so that these different correlations are not a result of random variability.

To test whether this is the case, we use a bootstrap test which randomly samples (with replacement) 14 years of detrended SAM and SOI from 1979-2013, and calculates the correlation. Figure 5.19 shows a histogram of these correlations along with the 1996-2009 correlation. The 1996-2009 value is not inconsistent with random variability at the 95% level for either a one- or two-tailed test. Therefore we conclude that the SAM/SOI correlation is not statistically significantly greater for 1996-2009. As such the 1996-2009 correlation is a marginal distribution of 1979-2013 so we include the

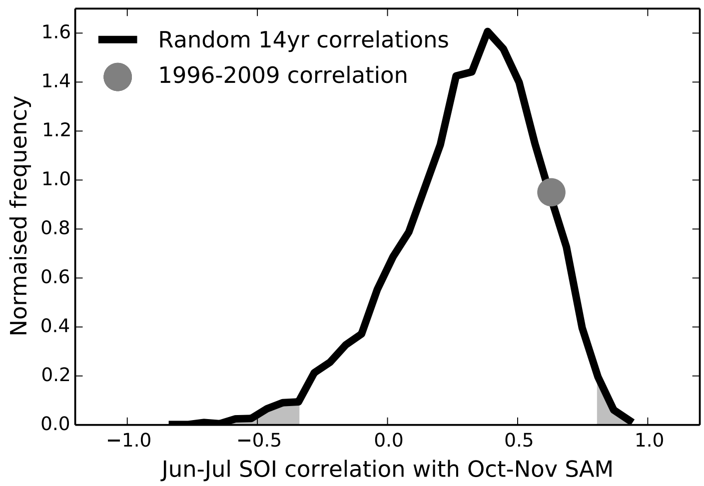

FIGURE 5.19: Histogram of correlations of random 14-year samples of detrended June-July SOI with October-November SAM. Also shown is the correlation over the 1996–2009 hindcast period. Grey shading indicates the $< 2.5\%$ and $> 97.5\%$ ranges.

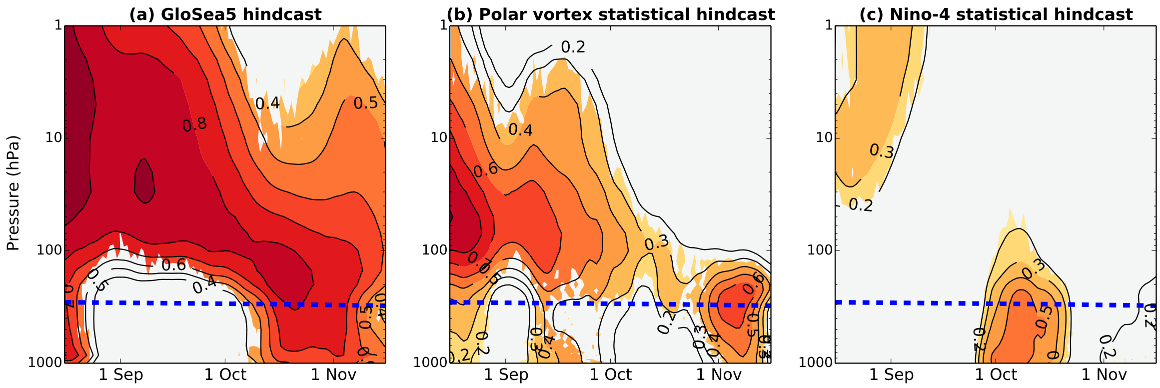

FIGURE 5.20: As Figure 5.15 but all analysis restricted to the 1996–2009 period.

longer ERA-Interim period in our analysis to provide a more robust measure of sources of predictability.

For completeness we include a figure with the same analysis limited to 1996–2009 in this appendix (Figure 5.20). Similar features as Figure 5.15 can be seen although tropospheric skill emerges later in Figure 5.20(b) and earlier in Figure 5.20(c).

CHAPTER 6

Conclusions

6.1 Summary of results

The main findings of this thesis are summarised as follows:

Application of moment diagnostics. It has been demonstrated that vortex moment diagnostics can be successfully applied to the geopotential height field, giving similar results as when applied to conservative fields such as PV. This provides a semi-Lagrangian (or vortex-centric) method which can be readily used to describe the geometry of the stratospheric polar vortex in climate model simulations.

It has been further shown that a simple threshold-based method can be applied to the vortex moment diagnostics in order to identify split and displaced vortex events. The majority of events identified in this way coincide with events defined by other methods, and capture equally extreme vortex states.

The stratospheric polar vortex in climate models. The first multi-model comparison of stratospheric polar vortex geometry and split an displaced vortex events has been carried out using the stratosphere-resolving CMIP5 models. A wide range of biases have been identified in the geometry of the stratospheric polar vortex among models. Some models have a vortex which is on average too equatorward, others too poleward, while the majority of models have a vortex which is too circularly symmetric. Models also vary widely in their frequency of split and displaced vortex events. However, the nature of these events is largely in agreement with observations, in particular the fact that split vortex events appear more barotropic and displaced vortex events baroclinic. The consistency of this difference in baroclinicity among models lends weight to the idea that split vortex events are caused by a resonant excitation of the barotropic mode, as suggested by *Esler and Scott* [2005]. Significantly, the frequency of split and displaced vortex events has been demonstrated to be highly correlated respectively with the aspect ratio and centroid latitude of the average vortex state. It therefore follows that an improvement in the mean state of the vortex is likely to lead to a more accurate representation of these extremes.

Stratosphere-troposphere coupling in climate models and observations. In reanalysis data, using the geopotential height-based vortex moments method, a stronger tropospheric NAM signal is seen following split vortex events than displaced vortex vortex events. This is in agreement with the results of *Mitchell et al.* [2013]. However, a bootstrap significance test of the surface NAM over the month following these events cannot exclude the possibility that this observed difference is due to chance.

In the CMIP5 models, the tropospheric NAM signal following both split and displaced vortex events is weak on average. There is no consistent difference between the two apart from close to the onset of events when there is a negative anomaly for split vortex events which extends barotropically through the depth of the atmosphere. However, looking at two-dimensional tropospheric anomalies in mean sea-level pressure following split and displaced vortex events shows some consistent features. A negative NAO-like signal is seen which is of similar magnitude following both types of event. The Pacific response is much less robust, with some models simulating negative pressure anomalies, and others positive. The discrepancy between the Atlantic and Pacific responses suggests that the annular mode may not be a good metric for stratosphere-troposphere coupling in the NH.

Almost all models show more negative sea-level pressure anomalies over Siberia following displaced vortex events than split vortex events. Overall, the differences in the surface signals following the two types of events are approximately co-located with the difference in lower-stratospheric geopotential height, which in turn follow stratospheric PV anomalies. A similar pattern is also seen in tropopause height in reanalysis data. This suggests the mechanism behind the different surface responses to split and displaced vortex events is one local to lower stratospheric PV anomalies, as proposed by *Ambaum and Hoskins* [2002]. However, it should be stressed that the similarities in the NAO response suggest that other mechanisms more sensitive to zonal-mean anomalies, such as baroclinic instability or planetary wave reflection, also play a role.

Predictability of the polar stratosphere. Using hindcast simulations produced by a stratosphere-resolving seasonal forecast system, no skill has been found in the prediction of NH SSWs or split or displaced vortex events at lead times beyond one month. This suggests that the skillful seasonal prediction of the winter NAO in the same system [*Scaife et al.*, 2014] is not highly influenced by the stratosphere. It may, however, be attributable to other model improvements such as increased atmospheric and oceanic horizontal resolution.

On the other hand, skillful prediction of the SH stratospheric polar vortex during the austral spring at seasonal lead times has been found. This skill is greater than a persistence forecast; indeed, a strong late-summer polar vortex is related to a weak spring vortex, indicating the importance of preconditioning. Using the observed relationship between the strength of the stratospheric polar vortex and polar ozone, it was possible to produce skillful forecasts of interannual variations in polar stratospheric ozone depletion. This prediction is at longer lead times than previous forecasts. Because interannual variability is significant when compared to the long-term ozone depletion trend, and has a significant impact on UV radiation reaching the Earth's surface, such forecasts may be of some interest for populations in the SH.

A further feature of the hindcast simulations is that the year 2002, in which the only observed SH SSW occurred, is also the most extreme of the hindcasts with almost all ensemble members simulating negative stratospheric wind anomalies. It also has one of the two out of 210 ensemble members which simulate SH SSW-like events (although these are displaced vortex events, rather than the split that occurred). This suggests that an increased likelihood of the 2002 event may have been detectable almost two months in advance.

Stratospheric influence on tropospheric predictability. The same seasonal forecast system produces skillful forecasts of the austral spring mean surface SAM at one month lead times. It also accurately simulates the surface temperature pattern associated with the SAM, such that the SAM forecast skill leads directly to skilful surface temperature forecasts over much of Antarctica, New Zealand, and eastern Australia. Interestingly, these forecasts were found to be more skilful during October–November (2 month lead time), than September (1 month lead time). The same pattern is replicated in a statistical hindcast which takes as its only input the polar-cap mean geopotential height at 10 hPa on 1st August. The pattern cannot, however, be replicated by a statistical forecast based on the ENSO index. This suggests, therefore, that the tropospheric skill during October-November is largely attributable to the influence of the predictable stratosphere during this time. The October–November stratospheric SAM is, in turn, highly predictable due to a strong negative correlation with the 1st August stratospheric SAM. The fact that the stratospheric influence is greatest in October-November is also backed-up by observational evidence which shows the largest stratosphere-leading correlations with the surface during this time. These results highlight the importance of including a well-resolved stratosphere and accurate stratospheric initial conditions in seasonal forecast systems.

6.2 Limitations and further investigations

The work presented in this thesis has raised a number of questions, and its limitations have motivated future investigations. Some of these ideas are discussed below:

What is required for a realistic stratosphere? Several studies over the past decade have demonstrated a more realistic climate and improved weather forecasts can be achieved using models which resolve the stratosphere. This has proved persuasive to modelling centres, leading an ever increasing number to include a representation of the stratosphere. Much of the work in this thesis has reaffirmed and provided a more detailed picture of the important role of the stratosphere in surface weather and climate. However, we have also clearly seen that a high-top is not a sufficient condition for a realistic stratosphere. A major challenge for the stratospheric community is to identify where limited computing resources should be best spent in simulating the stratosphere.

It was shown in Figure 4.12 that there appears to be a relationship between the average aspect ratio of the stratospheric polar vortex and vertical resolution among the CMIP5 models. Although this relationship is backed up by the physical understanding of the influence fine-scale vertical structure on planetary wave propagation in this region, it is not highly statistically significant. Furthermore, the relationship does not hold when models of the same family but different resolution are compared. This highlights a general limitation of multi-model, 'ensemble of opportunity', studies such as that in Chapter 4; so many variables are changed between different model simulations it is difficult to attribute model differences to any one factor (also discussed by *Tebaldi and Knutti* [2007]).

These issues could be addressed by performing a series of model integrations in which resolution is systematically varied. This should involve horizontal as well as vertical resolution, since it is likely that horizontal resolution is important for resolving steep PV gradients at the vortex edge which affect wave propagation (although no significant relationships with horizontal resolution were found in Chapter 4). Such a study need not be very computationally expensive, since it was shown in Figure 4.5 that the average state of the vortex is strongly related to the frequency of extreme events. Hence, it is only necessary to simulate enough years to determine the average state, which is far fewer than is necessary to determine a realistic climatology of extremes. If such a study finds any thresholds in resolution, beyond which stratospheric biases are much reduced, then this could act as a recommended resolution for modelling centres.

Synchronisation of the stratosphere and troposphere? A large part of this thesis has focussed on developing an increased understanding of the spatial stucture of stratosphere-troposphere coupling. However, the mechanisms discussed have retained the traditional temporal chain of causation of the form: *A causes B; B casues C* etc.. In the real, chaotic atmosphere, it is unlikely that such a simple mechanism exists. A new approach to understanding stratosphere-troposphere coupling could focus on the synchronisation of modes of variability. Indeed, we have seen here that such modes may be important because of the barotropic nature of split vortex events, suggesting an excitation of the barotropic mode during these events.

The instantaneous phase of an arbitrary signal can be calculated through the Hilbert transform [*Pikovsky et al.*, 2001], and several recent studies have applied this technique to investigate the phase synchronisation of modes of climate variability. For example, *Maraun and Kurths* [2005] found evidence for intermittent synchronization of ENSO and in Indian Monsoon, which they suggested were initiated by volcanic eruptions. *Read and Castrejón-Pita* [2012] also found phase synchronisation between the QBO and the semi-annual oscillation (a oscillation of upper stratosheric equatorial zonal winds with a period of six months), but with a non-stationary ratio of frequencies between the two oscillations.

In principle, a similar technique can be applied to study stratosphere-troposphere coupling; looking, for instance, at whether stratospheric and tropospheric modes are synchronised following particular events, such as SSWs. A difficuly in this case is deciding which are the relevant modes of variability. We could choose the NAM, although, as discussed previously, this has different physical interpretations in the troposphere and stratosphere. *Thompson and Woodworth* [2014] suggested the existence of barotropic and baroclinic annular modes; defined as the leading modes of variability of zonal-mean kinetic energy and eddy kinetic energy respectively. However, they found that this separation is less easy to perform in the NH than the SH.

Modes of variability can also be separated by their temporal structure. This is traditionally carried out through a Fourier spectrum analysis, however, the more modern technique of empirical mode decomposition (EMD) [*Huang et al.*, 1998] may be better suited to studying stratosphere-troposphere coupling. EMD has also been used in atmospheric science by *Coughlin and Tung* [2004] to study the influence of solar variability on the stratosphere. The method decomposes a given time series into a finite number of 'modes', each of which have a characteristic frequency. Unlike Fourier analysis, this frequency is allowed to vary to some degree, so the modes need not be perfectly periodic. As such, it is more applicable to time series of finite length and with a pronounced seasonal variability, such as is seen in the atmosphere. Figure 6.1 shows an example of EMD applied to NH polar-cap average geopotential height. It can be seen that the technique identifies different time scale, quasi-periodic modes of variability, and that these modes occasionally appear coherent through the stratosphere and troposphere. Closer inspection also reveals mode 2 to be more baroclinic than mode 1, consistent with *Thompson and Woodworth* [2014] who found their periodic baroclinic mode to

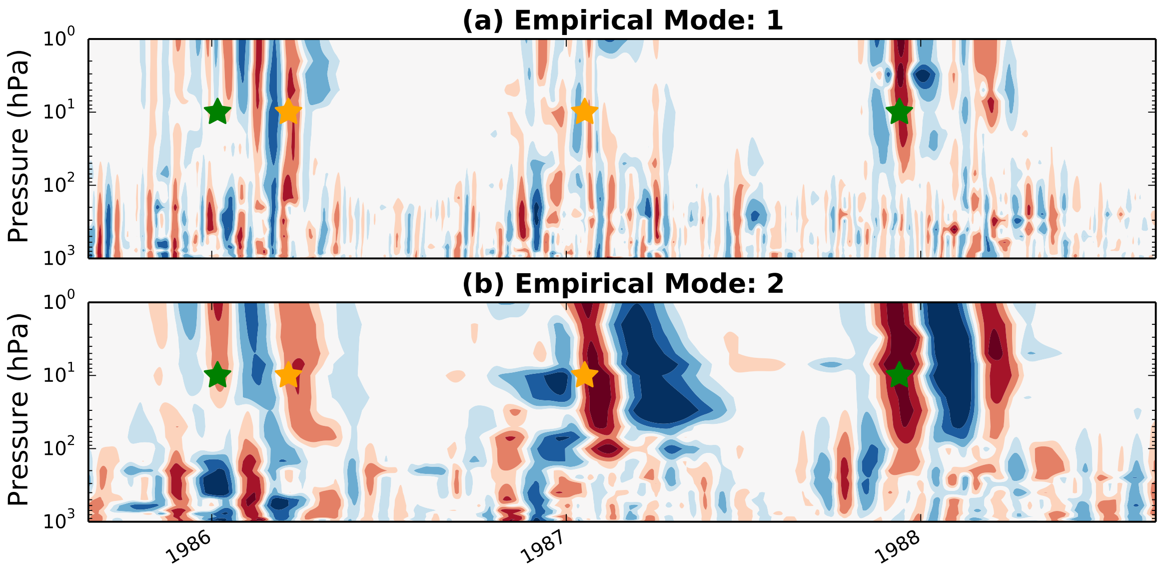

FIGURE 6.1: Time slice of empirical modes 1 and 2 of NH polar-cap averaged (60-90°N) geopotential height in ERA-Interim data. Yellow and green stars represent the onset of displaced and split vortex events respectively.

have a longer time scale than the barotropic mode (although this analysis was for the SH). Further investigations could be carried out to analyse the physical relevance of the modes and to quantify synchronisation between the stratosphere and troposphere. Also, given the results in this thesis which suggest the NAO is a more relevant metric than the NAM in stratosphere-troposphere coupling, it may be more appropriate to study local rather than hemispheric modes of variability in the troposphere.

What factors influence seasonal forecast skill? In Chapter 5 an increase tropospheric seasonal forecast skill was attributed to the influence of the stratosphere through analysis of a statistical forecast. As previously discussed, this method has the disadvantage that it cannot rule out a third factor which separately influences both the stratosphere and troposphere (it was shown that ENSO can be ruled out as such a factor, but it would be impossible to consider all potential influences). A more robust understanding of the factors influencing seasonal forecast skill can be gained by performing a series of hindcasts in which these factors are systematically changed.

An interesting case study for this investigation would the 2002 austral spring, since it was shown that the anomalous nature of this season was, to some degree, captured two months in advance. For instance, the 2002 hindcasts could be re-run with an opposite phase of the QBO, different tropical Pacific or Southern Ocean SSTs, or different polar stratospheric initial conditions. The change in forecasts of the stratospheric polar vortex and the surface SAM could then be analysed, indicating which factors are most important. The main difficulty in this ivestigation would probably come in imposing these different initial conditions in a physically consistent manner (e.g., conserving angular momentum).

Would interactive ozone chemistry improve seasonal forecast skill? The seasonal forecast system analysed in Chapter 5 did not include interactive chemistry, with ozone concentrations set to a climatology. It is therefore unable to capture the feedback between ozone concentrations and the stratospheric circulation, or zonal asymmetries in ozone. *Waugh et al.* [2009] suggested that such asymmetries could have a significant impact on tropospheric climate. This motivates an additional investigation as to whether improved stratospheric or tropospheric forecasts may be achieved by including interactive chemistry in a seasonal forecast system. Such a chemistry scheme is likely to be expensive, so the investigation should determine which reactions have the most impact on forecast skill.

Bibliography

Albers, J. R., and T. Birner (2014), Relative Roles of Planetary and Gravity Waves in Vortex Preconditioning Prior to Sudden Stratospheric Warmings, *J. Atmos. Sci.*, *in press*, doi:10.1175/JAS-D-14-0026.1.

Ambaum, M. H. P., and B. J. Hoskins (2002), The NAO troposphere-stratosphere connection, *J. Climate*, *15*, 1969–1978.

Ambaum, M. H. P., B. J. Hoskins, and D. B. Stephenson (2001), Arctic Oscillation or North Atlantic Oscillation?, *J. Climate*, *14*, 3495–3507.

Andrews, D., and M. McIntyre (1976), Planetary Waves in Horizontal and Vertical Shear: The Generalized Eliassen-Palm Relation and the Mean Zonal Acceleration., *J. Atmos. Sci.*, *33*, 2031–2048.

Andrews, D., J. R. Holton, and C. Leovy (1987), *Middle Atmosphere Dynamics*, 489 pp., Academic Press, San Diego, Calif.

Anstey, J. A., P. Davini, L. J. Gray, T. J. Woollings, N. Butchart, C. Cagnazzo, B. Christiansen, S. C. Hardiman, S. M. Osprey, and S. Yang (2013), Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution, *J. Geophys. Res.*, *118*, 3956–3971, doi:10.1002/jgrd.50231.

Baldwin, M. P., and T. J. Dunkerton (1999), Propagation of the Arctic Oscillation from the stratosphere to the troposphere, *J. Geophys. Res.*, *104*, 30,937–30,946, doi:10.1029/1999JD900445.

Baldwin, M. P., and T. J. Dunkerton (2001), Stratospheric harbingers of anomalous weather regimes., *Science*, *294*, 581–584.

Baldwin, M. P., and D. W. J. Thompson (2009), A critical comparison of stratosphere-troposphere coupling indices, *Q. J. R. Meteorol. Soc.*, *1672*, 1661–1672, doi:10.1002/qj.

Baldwin, M. P., X. Cheng, and T. J. Dunkerton (1994), Observed correlations between winter-mean tropospheric and stratospheric circulation anomalies, *Geophys. Res. Lett.*, *21*, 1141–1144.

Baldwin, M. P., D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O'Neill (2003), Stratospheric memory and skill of extended-range weather forecasts., *Science*, *301*, 636–640, doi:10.1126/science.1087143.

Bancalá, S., K. Krüger, and M. Giorgetta (2012), The preconditioning of major sudden stratospheric warmings, *J. Geophys. Res.*, *117*, D04,101, doi:10.1029/2011JD016769.

Bell, C. J., L. J. Gray, A. J. Charlton-Perez, M. M. Joshi, and A. A. Scaife (2009), Stratospheric Communication of El Niño Teleconnections to European Winter, *J. Climate*, *22*, 4083–4096, doi:10.1175/2009JCLI2717.1.

Black, R. (2002), Stratospheric forcing of surface climate in the Arctic Oscillation, *J. Clim.*, pp. 268–277.

Black, R. X., and B. A. McDaniel (2007), Interannual Variability in the Southern Hemisphere Circulation Organized by Stratospheric Final Warming Events, *J. Atmos. Sci.*, *64*, 2968–2974, doi:10.1175/JAS3979.1.

Blockley, E. W., M. J. Martin, A. J. McLaren, A. G. Ryan, J. Waters, D. J. Lea, I. Mirouze, K. A. Peterson, A. Sellar, and D. Storkey (2013), Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, *Geosci. Model Dev. Discuss.*, *6*, 6219–6278, doi:10.5194/gmdd-6-6219-2013.

Bodeker, G., H. Shiona, and H. Eskes (2005), Indicators of Antarctic ozone depletion, *Atmos. Chem. Phys.*, *5*, 2603–2615.

Boville, B. A. (1984), The influence of the polar night jet on the tropospheric circulation in a GCM, *J. Atmos. Sci.*, *41*, 1132–1142.

Bowler, N., A. Arribas, S. E. Beare, K. R. Mylne, and G. J. Shutts (2009), The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system, *Q. J. R. Meteorol. Soc.*, *776*, 767–776, doi:10.1002/qj.394.

Butchart, N., and E. Remsberg (1986), The Area of the Stratospheric Polar Vortex as a Diagnostic for Tracer Transport on an Isentropic Surface, *J. Atmos. Sci.*, *43*, 1319–1339.

Butchart, N., A. J. Charlton-Perez, I. Cionni, S. C. Hardiman, P. H. Haynes, K. Krüger, P. J. Kushner, P. A. Newman, S. M. Osprey, J. Perlwitz, M. Sigmond, L. Wang, H. Akiyoshi, J. Austin, S. Bekki, A. Baumgaertner, P. Braesicke, C. Brühl, M. Chipperfield, M. Dameris, S. Dhomse, V. Eyring, R. Garcia, H. Garny, P. Jöckel, J.-F. Lamarque, M. Marchand, M. Michou, O. Morgenstern, T. Nakamura, S. Pawson, D. Plummer, J. Pyle, E. Rozanov, J. Scinocca, T. G. Shepherd, K. Shibata, D. Smale, H. Teyssèdre, W. Tian, D. Waugh, and Y. Yamashita (2011), Multimodel climate and variability of the stratosphere, *J. Geophys. Res.*, *116*, D05,102, doi:10.1029/2010JD014995.

Butler, A. H., D. J. Seidel, and S. C. Hardiman (2014), Defining sudden stratospheric warmings, *Bull. Amer. Meteor. Soc.*, *in review*.

Cagnazzo, C., and E. Manzini (2009), Impact of the Stratosphere on the Winter Tropospheric Teleconnections between ENSO and the North Atlantic and European Region, *J. Clim.*, *22*, 1223–1238, doi:10.1175/2008JCLI2549.1.

Castanheira, J. M., and D. Barriopedro (2010), Dynamical connection between tropospheric blockings and stratospheric polar vortex, *Geophys. Res. Lett.*, *37*, L13,809, doi:10.1029/2010GL043819.

Charlton, A. J., and L. M. Polvani (2007), A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks, *J. Climate*, *20*, 449–470.

Charlton, A. J., A. O'Neill, D. B. Stephenson, W. A. Lahoz, and M. P. Baldwin (2003), Can knowledge of the state of the stratosphere be used to improve statistical forecasts of the troposphere?, *Q. J. R. Meteorol. Soc.*, *129*, 3205–3224, doi:10.1256/qj.02.232.

Charlton, A. J., A. O'Neill, W. A. Lahoz, A. C. Massacand, and P. Berrisford (2005), The impact of the stratosphere on the troposphere during the southern hemisphere stratospheric sudden warming, September 2002, *Q. J. R. Meteorol. Soc.*, *131*, 2171–2188, doi:10.1256/qj.04.43.

Charlton-Perez, A. J., M. P. Baldwin, T. Birner, R. X. Black, A. H. Butler, N. Calvo, N. A. Davis, E. P. Gerber, N. Gillett, S. Hardiman, J. Kim, K. Krüger, Y.-Y. Lee, E. Manzini, B. A. McDaniel, L. Polvani, T. Reichler, T. A. Shaw, M. Sigmond, S.-W. Son, M. Toohey, L. Wilcox, S. Yoden, B. Christiansen, F. Lott, D. Shindell, S. Yukimoto, and S. Watanabe (2013), On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models, *J. Geophys. Res.*, *118*, 2494–2505, doi:10.1002/jgrd.50125.

Charney, J. G., and P. G. Drazin (1961), Propagation of Planetary-Scale Disturbances from the Lower into the Upper Atmosphere, *J. Geophys. Res*, *66*, 83–109.

Charney, J. G., and M. E. Stern (1962), On the stability of internal baroclinic jets in a rotating atmosphere, *J. Atmos. Sci.*, *19*, 159–172.

Chen, G., and P. Zurita-Gotor (2008), The Tropospheric Jet Response to Prescribed Zonal Forcing in an Idealized Atmospheric Model, *J. Atmos. Sci.*, *65*, 2254–2271, doi:10.1175/2007JAS2589.1.

Chen, P., and W. Robinson (1992), Propagation of planetary waves between the troposphere and stratosphere, *J. Atmos. Sci.*, *49*, 2533–2545.

Christiansen, B. (2001), Downward propagation of zonal mean zonal wind anomalies from the stratosphere to the troposphere: Model and reanalysis, *J. Geophys. Res.*, *106*, 27,307–27,322, doi:10.1029/2000JD000214.

Christiansen, B. (2005), Downward propagation and statistical forecast of the near-surface weather, *J. Geophys. Res.*, *110*(D14104), doi:10.1029/2004JD005431.

Cionni, I., V. Eyring, J. F. Lamarque, W. J. Randel, D. S. Stevenson, F. Wu, G. E. Bodeker, T. G. Shepherd, D. T. Shindell, and D. W. Waugh (2011), Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, *Atmos. Chem. Phys.*, *11*, 11,267–11,292, doi:10.5194/acp-11-11267-2011.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values.*, 209 pp., Springer, London, UK.

Cordero, E. C., and P. M. d. F. Forster (2006), Stratospheric variability and trends in models used for the IPCC AR4, *Atmos. Chem. Phys.*, *6*, 5369–5380, doi:10.5194/acp-6-5369-2006.

Coughlin, K. T., and L. J. Gray (2009), A continuum of sudden stratospheric warmings, *J. Atmos. Sci.*, *66*, 531–540.

Coughlin, K. T., and K. K. Tung (2004), 11-Year solar cycle in the stratosphere extracted by the empirical mode decomposition method, *Adv. Sp. Res.*, *34*, 323–329, doi:10.1016/j.asr.2003.02.045.

Crook, J. A., N. P. Gillett, and S. P. E. Keeley (2008), Sensitivity of Southern Hemisphere climate to zonal asymmetry in ozone, *Geophys. Res. Lett.*, *35*, L07,806, doi:10.1029/2007GL032698.

Davini, P., C. Cagnazzo, and J. A. Anstey (2014), Blocking view of the stratosphere-troposphere coupling, *J. Geophys. Res.*, *in press*, doi:10.1002/2014JD021703.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kå llberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*, 553–597, doi:10.1002/qj.828.

Deser, C. (2000), On the teleconnectivity of the "Arctic Oscillation", *Geophys. Res. Lett.*, *27*, 779–782, doi:10.1029/1999GL010945.

Dragani, R. (2011), On the quality of the ERA-Interim ozone reanalyses: comparisons with satellite data, *Q. J. R. Meteorol. Soc.*, *137*, 1312–1326, doi:10.1002/qj.821.

Dunkerton, T. (1978), On the Mean Meridional Mass Motions of the Stratosphere and Mesosphere, *J. Atmos. Sci.*, *35*, 2325–2333.

Eade, R., D. Smith, A. Scaife, E. Wallace, N. Dunstone, L. Hermanson, and N. Robinson (2014), Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?, *Geophys. Res. Lett.*, *41*, doi:10.1002/2014GL061146.

Eady, E. T. (1949), Long waves and cyclone waves, *Tellus*, *1*, 33–52.

Eskes, H. (2005), Ozone forecasts of the stratospheric polar vortex-splitting event in September 2002, *J. Atmos. Sci.*, *62*, 812–821.

Esler, J. G., and N. J. Matthewman (2011), Stratospheric Sudden Warmings as Self-Tuning Resonances. Part II: Vortex Displacement Events, *J. Atmos. Sci.*, *68*, 2505–2523, doi:10.1175/JAS-D-11-08.1.

Esler, J. G., and R. K. Scott (2005), Excitation of transient Rossby waves on the stratospheric polar vortex and the barotropic sudden warming, *J. Atmos. Sci.*, *62*, 3661–3682.

Esler, J. G., L. M. Polvani, and R. K. Scott (2006), The Antarctic stratospheric sudden warming of 2002: A self-tuned resonance?, *Geophys. Res. Lett.*, *33*, L12,804, doi: 10.1029/2006GL026034.

Farman, J. C., B. G. Gardiner, and J. D. Shanklin (1985), Large losses of total ozone in Antarctica reveal seasonal ClOx/NOx interaction, *Nature*, *315*, 207–210, doi: 10.1038/315207a0.

Forster, P. M. d. F., and K. P. Shine (1997), Radiative forcing and temperature trends from stratospheric ozone changes, *J. Geophys. Res.*, *102*, 10,841–10,855.

Garfinkel, C. I., and D. L. Hartmann (2008), Different ENSO teleconnections and their effects on the stratospheric polar vortex, *J. Geophys. Res*, *113*, D18,114, doi: 10.1029/2008JD009920.

Garfinkel, C. I., and D. L. Hartmann (2011), The Influence of the Quasi-Biennial Oscillation on the Troposphere in Winter in a Hierarchy of Models. Part I: Simplified Dry GCMs, *J. Atmos. Sci.*, *68*, 1273–1289, doi:10.1175/2011JAS3665.1.

Garny, H., V. Grewe, M. Dameris, G. E. Bodeker, and A. Stenke (2011), Attribution of ozone changes to dynamical and chemical processes in CCMs and CTMs, *Geosci. Model Dev.*, *4*, 271–286, doi:10.5194/gmd-4-271-2011.

Gerber, E. P., S. Voronin, and L. M. Polvani (2008), Testing the Annular Mode Autocorrelation Time Scale in Simple Atmospheric General Circulation Models, *Mon. Weather Rev.*, *136*, 1523–1536, doi:10.1175/2007MWR2211.1.

Gerber, E. P., C. Orbe, and L. M. Polvani (2009), Stratospheric influence on the tropospheric circulation revealed by idealized ensemble forecasts, *Geophys. Res. Lett.*, *36*, L24,801, doi:10.1029/2009GL040913.

Gerber, E. P., A. Butler, N. Calvo, A. J. Charlton-Perez, M. Giorgetta, E. Manzini, and J. Perlwitz (2012), Assessing and Understanding the Impact of Stratospheric Dynamics and Variability on the Earth System, *Bull. Amer. Meteor. Soc.*, *93*, 845–859.

Gillett, N. P., T. D. Kell, and P. D. Jones (2006), Regional climate impacts of the Southern Annular Mode, *Geophys. Res. Lett.*, *33*, L23,704, doi:10.1029/2006GL027721.

Gong, D., and S. Wang (1999), Definition of Antarctic Oscillation index, *Geophys. Res. Lett.*, *26*, 459–462, doi:10.1029/1999GL900003.

Gray, L. J., J. Beer, M. Geller, J. D. Haigh, M. Lockwood, K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, L. Hood, J. Luterbacher, G. A. Meehl, D. Shindell, B. van Geel, and W. White (2010), Solar influences on climate, *Rev. Geophys.*, *48*, RG4001, doi:10.1029/2009RG000282.

Gray, L. J., A. A. Scaife, D. M. Mitchell, S. Osprey, S. Ineson, S. Hardiman, N. Butchart, J. Knight, R. Sutton, and K. Kodera (2013), A lagged response to the 11 year solar cycle in observed winter Atlantic/European weather patterns, *J. Geophys. Res.*, *118*, 13,405–13,420, doi:10.1002/2013JD020062.

Grise, K. M., D. W. J. Thompson, and P. M. Forster (2009), On the Role of Radiative Processes in Stratosphere-Troposphere Coupling, *J. Clim.*, *22*, 4154–4161, doi:10.1175/2009JCLI2756.1.

Grise, K. M., D. W. J. Thompson, and T. Birner (2010), A Global Survey of Static Stability in the Stratosphere and Upper Troposphere, *J. Clim.*, *23*(9), 2275–2292, doi:10.1175/2009JCLI3369.1.

Haigh, J. D. (2003), The effects of solar variability on the Earth's climate, *Philos. Trans. R. Soc. A*, *361*, 95–111.

Hall, A. R. (2005), *Generalized Method of Moments (Advanced Texts in Econometrics)*, Oxford University Press, Oxford, UK.

Hall, T. M., and R. A. Plumb (1994), Age as a diagnostic of stratospheric transport, *J. Geophys. Res.*, *99*, 1059–1070, doi:10.1029/93JD03192.

Hannachi, A., D. Mitchell, L. Gray, and A. Charlton-Perez (2010), On the use of geometric moments to examine the continuum of sudden stratospheric warmings, *J. Atmos. Sci.*, *68*, 657–674.

Hardiman, S. C., N. Butchart, A. J. Charlton-Perez, T. A. Shaw, H. Akiyoshi, A. Baumgaertner, S. Bekki, P. Braesicke, M. Chipperfield, M. Dameris, R. R. Garcia, M. Michou, S. Pawson, E. Rozanov, and K. Shibata (2011), Improved predictability of the troposphere using stratospheric final warmings, *J. Geophys. Res*, *116*, D18,113, doi:10.1029/2011JD015914.

Hardiman, S. C., N. Butchart, T. J. Hinton, S. M. Osprey, and L. J. Gray (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*, 7083–7099, doi:10.1175/JCLI-D-11-00579.1.

Hartley, D. E., J. T. Villarin, R. X. Black, and C. A. Davis (1998), A new perspective on the dynamical link between the stratosphere and troposphere, *Nature*, *391*, 1996–1999, doi:10.1038/35112.

Hartmann, D. L., J. M. Wallace, V. Limpasuvan, D. W. Thompson, and J. R. Holton (2000), Can ozone depletion and global warming interact to produce rapid climate change?, *Proc. Natl. Acad. Sci. U. S. A.*, *97*(4), 1412–7.

Haynes, P. (2005), Stratospheric Dynamics, *Annu. Rev. Fluid Mech.*, *37*, 263–293, doi:10.1146/annurev.fluid.37.061903.175710.

Haynes, P., M. McIntyre, T. G. Shepherd, and K. P. Shine (1991), On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces, *J. Atmos. Sci.*, *48*, 651–678.

Hendon, H. H., D. W. J. Thompson, and M. C. Wheeler (2007), Australian Rainfall and Surface Temperature Variations Associated with the Southern Hemisphere Annular Mode, *J. Clim.*, *20*, 2452–2467, doi:10.1175/JCLI4134.1.

Hewitt, H. T., D. Copsey, I. D. Culverwell, C. M. Harris, R. S. R. Hill, a. B. Keen, A. J. McLaren, and E. C. Hunke (2011), Design and implementation of the infrastructure of HadGEM3: the next-generation Met Office climate modelling system, *Geosci. Model Dev.*, *4*, 223–253, doi:10.5194/gmd-4-223-2011.

Holton, J. (1990), On the global exchange of mass between the stratosphere and troposphere, *J. Atmos. Sci.*, *47*, 392–395.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson (1985), On the use and significance of isentropic potential vorticity maps, *Q. J. R. Meteorol. Soc.*, *111*, 877–946.

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. Lond. A*, *454*, 903–995.

Huebener, H., U. Cubasch, U. Langematz, T. Spangehl, F. Niehörster, I. Fast, and M. Kunze (2007), Ensemble climate simulations using a fully coupled ocean-troposphere-stratosphere general circulation model., *Philos. Trans. A. Math. Phys. Eng. Sci.*, *365*, 2089–2101, doi:10.1098/rsta.2007.2078.

Hurwitz, M. M., P. A. Newman, L. D. Oman, and A. M. Molod (2011), Response of the Antarctic Stratosphere to Two Types of El Niño Events, *J. Atmos. Sci*, *68*, 812–822, doi:10.1175/2011JAS3606.1.

Huth, R. (2006), Arctic or North Atlantic Oscillation? Arguments based on the principal component analysis methodology, *Theor. Appl. Climatol.*, *89*, 1–8, doi:10.1007/s00704-006-0257-1.

Ineson, S., and A. A. Scaife (2009), The role of the stratosphere in the European climate response to El Niño, *Nat. Geosci.*, *2*, 32–36, doi:10.1038/NGEO381.

Jacqmin, D., and R. Lindzen (1985), The causation and sensitivity of the northern winter planetary waves, *J. Atmos. Sci.*, *42*, 724–746.

Jung, T., and J. Barkmeijer (2006), Sensitivity of the Tropospheric Circulation to Changes in the Strength of the Stratospheric Polar Vortex, *Mon. Weather Rev.*, pp. 2191–2207.

Karpetchko, A., E. Kyrö, and B. M. Knudsen (2005), Arctic and Antarctic polar vortices 1957-2002 as seen from the ERA-40 reanalyses, *J. Geophys. Res.*, *110*, D21,109, doi:10.1029/2005JD006113.

Kodera, K., and M. Chiba (1995), Tropospheric circulation changes associated with stratospheric sudden warmings: A case study, *J. Geophys. Res.*, *100*, 11,055–11,068, doi:10.1029/95JD00771.

Kodera, K., K. Yamazaki, M. Chiba, and K. Shibata (1990), Downward propagation of upper stratospheric mean zonal wind perturbation to the troposphere, *Geophys. Res. Lett.*, *17*, 1263–1266.

Kolstad, E. W., T. Breiteig, and A. A. Scaife (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, doi:10.1002/qj.620.

Kroon, M., J. P. Veefkind, M. Sneep, R. D. McPeters, P. K. Bhartia, and P. F. Levelt (2008), Comparing OMI-TOMS and OMI-DOAS total ozone column data, *J. Geophys. Res.*, *113*, D16S28, doi:10.1029/2007JD008798.

Kumar, A. (2009), Finite Samples and Uncertainty Estimates for Skill Measures for Seasonal Prediction, *Mon. Weather Rev.*, *137*, 2622–2631, doi:10.1175/2009MWR2814.1.

Kuroda, Y. (2004), Role of the Polar-night Jet Oscillation on the formation of the Arctic Oscillation in the Northern Hemisphere winter, *J. Geophys. Res.*, *109*, D11,112, doi:10.1029/2003JD004123.

Kuroda, Y. (2008), Role of the stratosphere on the predictability of medium-range weather forecast: A case study of winter 2003-2004, *Geophys. Res. Lett.*, *35*, L19,701, doi:10.1029/2008GL034902.

Kushner, P. J. (2010), Annular Modes of the Troposphere and Stratosphere, The Stratosphere: Dynamics Transport and Chemistry, *Geophys. Monogr. Ser.*, *190*, 59–91.

Kushner, P. J., and L. M. Polvani (2004), Stratosphere-troposphere coupling in a relatively simple AGCM: The role of eddies, *J. Climate*, *17*, 629–639, doi:10.1175/JCLI4007.1.

Labitzke, K., and B. Naujokat (2000), The Lower Arctic Stratosphere since 1952, *SPARC Newsl.*, (15).

Li, B., and S. Ma (1994), Efficient computation of 3D moments, *Pattern Recognition, 1994. Vol. 1 - Conf. A Comput. Vis. Image Process. Proc. 12th IAPR Int. Conf.*, *1*, 22–26, doi:10.1109/ICPR.1994.576218.

Lim, E.-P., H. H. Hendon, and H. Rashid (2013), Seasonal Predictability of the Southern Annular Mode due to its Association with ENSO, *J. Climate*, *26*, 8037–8045, doi: 10.1175/JCLI-D-13-00006.1.

Limpasuvan, V., and D. Hartmann (2000), Wave-Maintained Annular Modes of Climate Variability, *J. Climate*, *13*, 4414–4429.

Limpasuvan, V., and D. L. Hartmann (1999), Eddies and the annular modes of climate variability, *Geophys. Res. Lett.*, *26*, 3133–3136, doi:10.1029/1999GL010478.

Limpasuvan, V., D. W. J. Thompson, and D. L. Hartmann (2004), The life cycle of the Northern Hemisphere sudden stratospheric warmings, *J. Climate*, *17*, 2584–2596.

Love, A. (1893), On the stability of certain vortex motions, *Proc. London Math. Soc.*, pp. 18–43.

MacLachlan, C., A. Arribas, K. A. Peterson, A. Maidens, D. Fereday, A. Scaife, M. Gordon, M. Vellinga, A. Williams, R. E. Comer, J. Camp, P. Xavier, and G. Madec (2014), Global Seasonal Forecast System version 5 (GloSea5): a high resolution seasonal forecast system, *Q. J. R. Meteorol. Soc.*, *in press*, doi:10.1002/qj.2396.

Manney, G. L., L. Froidevaux, M. L. Santee, R. W. Zurek, and J. W. Waters (1997), MLS observations of Arctic ozone loss in 1996-97, *Geophys. Res. Lett.*, *24*, 2697–2700.

Manzini, E., a. Y. Karpechko, J. Anstey, M. P. Baldwin, R. X. Black, C. Cagnazzo, N. Calvo, A. Charlton-Perez, B. Christiansen, P. Davini, E. Gerber, M. Giorgetta, L. Gray, S. C. Hardiman, Y.-Y. Lee, D. R. Marsh, B. a. McDaniel, A. Purich, a. a. Scaife, D. Shindell, S.-W. Son, S. Watanabe, and G. Zappa (2014), Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere - troposphere coupling, *J. Geophys. Res.*, *119*, 7979–7998, doi:10.1002/2013JD021403.

Maraun, D., and J. Kurths (2005), Epochs of phase coherence between El Niño/Southern Oscillation and Indian monsoon, *Geophys. Res. Lett.*, *32*, L15,709, doi:10.1029/2005GL023225.

Marshall, A. G., and A. A. Scaife (2009), Impact of the QBO on surface winter climate, *J. Geophys. Res.*, *114*, D18,110, doi:10.1029/2009JD011737.

Marshall, A. G., and A. A. Scaife (2010), Improved predictability of stratospheric sudden warming events in an atmospheric general circulation model with enhanced stratospheric resolution, *J. Geophys. Res.*, *115*, D16,114, doi:10.1029/2009JD012643.

Marshall, G. J. (2003), Trends in the southern annular mode from observations and reanalyses., *J. Clim.*, *16*, 4134–4143.

Martius, O., L. M. Polvani, and H. C. Davies (2009), Blocking precursors to stratospheric sudden warming events, *Geophys. Res. Lett.*, *36*, L14,806, doi:10.1029/2009GL038776.

Matsuno, T. (1970), Vertical Propagation of Stationary Planetary Waves in the Winter Northern Hemisphere., *J. Atmos. Sci.*, *27*(6), 871–883.

Matsuno, T. (1971), A dynamical model of the stratospheric sudden warming., *J. Atmos. Sci.*, *28*(8), 1479–1494.

Matthewman, N. J., and J. G. Esler (2011), Stratospheric Sudden Warmings as Self-Tuning Resonances. Part I: Vortex Splitting Events, *J. Atmos. Sci.*, *68*, 2481–2504, doi:10.1175/JAS-D-11-07.1.

Matthewman, N. J., J. G. Esler, a. J. Charlton-Perez, and L. M. Polvani (2009), A New Look at Stratospheric Sudden Warmings. Part III: Polar Vortex Evolution and Vertical Structure, *J. Climate*, *22*, 1566–1585, doi:10.1175/2008JCLI2365.1.

Maycock, A. C., S. P. E. Keeley, A. J. Charlton-Perez, and F. J. Doblas-Reyes (2011), Stratospheric circulation in seasonal forecasting models: implications for seasonal prediction, *Clim. Dyn.*, *36*, 309–321, doi:10.1007/s00382-009-0665-x.

McElroy, M. B., R. J. Salawitch, S. C. Wofsy, and J. A. Logan (1986), Reductions of Antarctic ozone due to synergistic interactions of chlorine and bromine, *Nature*, *321*, 759–762, doi:10.1038/321759a0.

McInturff, R. (1978), Stratospheric warmings: Synoptic, dynamic and general circulation aspects, *NASA Ref. Publ.*, *1017*.

McIntyre, M. E. (1982), How well do we understand the dynamics of stratospheric warmings, *J. Meteor. Soc. Japan*, *60*, 37–65.

McIntyre, M. E., and T. N. Palmer (1983), Breaking planetary waves in the stratosphere, *Nature*, *305*, 593–600, doi:10.1038/305593a0.

Mitchell, D. M., A. J. Charlton-Perez, and L. J. Gray (2011), Characterizing the Variability and Extremes of the Stratospheric Polar Vortices Using 2D Moment Analysis, *J. Atmos. Sci.*, *68*, 1194–1213.

Mitchell, D. M., A. J. Charlton-Perez, L. J. Gray, H. Akiyoshi, N. Butchart, S. C. Hardiman, O. Morgenstern, T. Nakamura, E. Rozanov, K. Shibata, D. Smale, and Y. Yamashita (2012), The nature of Arctic polar vortices in chemistry-climate models, *Q. J. R. Meteorol. Soc.*, *138*, 1681–1691, doi:10.1002/qj.1909.

Mitchell, D. M., L. J. Gray, J. Anstey, M. P. Baldwin, and A. J. Charlton-Perez (2013), The Influence of Stratospheric Vortex Displacements and Splits on Surface Climate, *J. Climate*, *26*, 2668–2682, doi:10.1175/JCLI-D-12-00030.1.

Molina, L., and M. Molina (1987), Production of chlorine oxide (Cl2O2) from the self-reaction of the chlorine oxide (ClO) radical, *J. Phys. Chem.*, *91*, 433–436, doi: 10.1021/j100286a035.

Monge-Sanz, B. M., M. P. Chipperfield, D. P. Dee, A. J. Simmons, and S. M. Uppala (2013), Improvements in the stratospheric transport achieved by a chemistry transport model with ECMWF (re)analyses: identifying effects and remaining challenges, *Q. J. R. Meteorol. Soc.*, *139*, 654–673, doi:10.1002/qj.1996.

Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones (2012), Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, *J. Geophys. Res.*, *117*, D08,101, doi:10.1029/2011JD017187.

Mukougawa, H., T. Hirooka, and Y. Kuroda (2009), Influence of stratospheric circulation on the predictability of the tropospheric Northern Annular Mode, *Geophys. Res. Lett.*, *36*(8), 1–5, doi:10.1029/2008GL037127.

Murphy, J. (1990), Assessment of the practical utility of extended range ensemble forecasts, *Q. J. R. Meteorol. Soc.*, *116*, 89–125.

Murray, F. W. (1960), Dynamic stability in the stratosphere, *J. Geophys. Res.*, *55*, 3273–3305.

Nakagawa, K. I., and K. Yamazaki (2006), What kind of stratospheric sudden warming propagates to the troposphere?, *Geophys. Res. Lett.*, *33*, L04,801, doi: 10.1029/2005GL024784.

Nash, E. R., P. A. Newman, J. E. Rosenfield, and M. R. Schoeberl (1996), An objective determination of the polar vortex using Ertel's potential vorticity, *J. Geophys. Res.*, *101*, 9471–9478.

Neef, L., S. Walther, K. Matthes, and K. Kodera (2014), Observations of stratospheric sudden warmings in Earth rotation variations, *J. Geophys. Res.*, *119*, doi:10.1002/2014JD021621.

Newman, P. A. (2010), Chemistry and Dynamics of the Antarctic Ozone Hole, The Stratosphere: Dynamics Transport and Chemistry, *Geophys. Monogr. Ser.*, *190*, 157–171.

Newman, P. A., and J. E. Rosenfield (1997), Stratospheric thermal damping times, *Geophys. Res. Lett.*, *24*, 433–436.

Nigam, S. (1990), On the structure of variability of the observed tropospheric and stratospheric zonal-mean zonal wind, *J. Atmos. Sci.*, *47*, 1799–1813.

Norton, W. A. (2003), Sensitivity of northern hemisphere surface climate to simulation of the stratospheric polar vortex, *Geophys. Res. Lett.*, *30*(12), 1627.

Nyquist, H. (1928), Thermal agitation of electric charge in conductors, *Phys. Rev.*, *32*, 110–113.

O'Neill, A., and V. D. Pope (1988), Simulations of linear and nonlinear disturbances in the stratosphere, *Q. J. R. Meteorol. Soc.*, *114*, 1063–1110, doi:10.1002/qj.49711448210.

O'Neill, A., W. L. Grose, V. D. Pope, H. Maclean, and R. Swinbank (1994), Evolution of the stratosphere during northern winter 1991/92 as diagnosed from UK Meteorological Office analyses, *J. Atmos. Sci.*, *51*, 2800–2817.

Perlwitz, J., and H.-F. Graf (1995), The statistical connection between tropospheric and stratospheric circulation of the Northern Hemisphere in winter, *J. Climate*, *8*, 2281–2295.

Perlwitz, J., and N. Harnik (2003), Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection, *J. Climate*, *16*(18), 3011–3026.

Pikovsky, A., M. Rosenblum, and J. Kurths (2001), *Synchronization: A universal concept in nonlinear sciences*, 411 pp., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Plumb, R. A. (1981), Instability of the distorted polar night vortex: A theory of stratospheric warmings, *J. Atmos. Sci.*, *38*, 2514–2531.

Plumb, R. A., and K. Semeniuk (2003), Downward migration of extratropical zonal wind anomalies, *J. Geophys. Res.*, *108*, 4223, doi:10.1029/2002JD002773.

Polvani, L. M., and P. J. Kushner (2002), Tropospheric response to stratospheric perturbations in a relatively simple general circulation model, *Geophys. Res. Lett.*, *29*, 40–43.

Quiroz, R. S. (1986), The Association of Stratospheric Warmings With Tropospheric Blocking, *J. Geophys. Res.*, *91*, 5277–5285.

Ramanathan, V. (1977), Troposphere-stratosphere feedback mechanism: stratospheric warming and its effect on the polar energy budget and the tropospheric circulation, *J. Atmos. Sci.*, *34*, 439–447.

Randel, W., P. Udelhofen, E. Fleming, M. Geller, M. Gelman, K. Hamilton, D. Karoly, D. Ortland, S. Pawson, R. Swinbank, F. Wu, M. Baldwin, M.-L. Chanin, P. Keckhut, K. Labitzke, E. Remsberg, A. Simmons, and D. Wu (2004), The SPARC intercomparison of middle-atmosphere climatologies, *J. Climate*, *17*, 986–1003.

Randel, W. J. (1988), The seasonal evolution of planetary waves in the Southern Hemisphere stratosphere and troposphere, *Q. J. R. Meteorol. Soc.*, *587*, 1385–1409.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*, 4407, doi:10.1029/2002JD002670.

Read, P. L., and A. A. Castrejón-Pita (2012), Phase synchronization between stratospheric and tropospheric quasi-biennial and semi-annual oscillations, *Q. J. R. Meteorol. Soc.*, *138*, 1338–1349, doi:10.1002/qj.1872.

Reason, C. J. C., and M. Rouault (2005), Links between the Antarctic Oscillation and winter rainfall over western South Africa, *Geophys. Res. Lett.*, *32*, L07,705, doi: 10.1029/2005GL022419.

Reichler, T., J. Kim, E. Manzini, and J. Kröger (2012), A stratospheric connection to Atlantic climate variability, *Nat. Geosci.*, *5*(10), 1–5, doi:10.1038/ngeo1586.

Rhines, P., and W. Holland (1979), A theoretical discussion of eddy-driven mean flows, *Dyn. Atmos. Ocean.*, *3*, 289–325.

Ring, M. J., and R. A. Plumb (2008), The Response of a Simplified GCM to Axisymmetric Forcings: Applicability of the Fluctuation-Dissipation Theorem, *J. Atmos. Sci.*, *65*, 3880–3898, doi:10.1175/2008JAS2773.1.

Robinson, W. (1991), The dynamics of the zonal index in a simple model of the atmosphere, *Tellus A*.

Robock, A. (2000), Volcanic eruptions and climate, *Rev. Geophys.*, *38*, 191–219.

Roff, G., D. W. J. Thompson, and H. Hendon (2011), Does increasing model stratospheric resolution improve extended-range forecast skill?, *Geophys. Res. Lett.*, *38*, L05,809, doi:10.1029/2010GL046515.

Roscoe, H. K., J. D. Shanklin, and S. R. Colwell (2005), Has the Antarctic vortex split before 2002?, *J. Atmos. Sci.*, *62*, 581–588, doi:10.1175/JAS-3331.1.

Rousseeuw, P. J. (1987), Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *J. Comput. Appl. Math.*, *20*, 53–65, doi:10.1016/0377-0427(87)90125-7.

Salby, M. L., E. A. Titova, and L. Deschamps (2012), Changes of the Antarctic ozone hole: Controlling mechanisms, seasonal predictability, and evolution, *J. Geophys. Res.*, *117*, D10,111, doi:10.1029/2011JD016285.

Sardeshmukh, P., G. Compo, and C. Penland (2000), Changes of probability associated with El Niño, *J. Clim.*, *5*, 4268–4286.

Sassi, F., R. R. Garcia, D. Marsh, and K. W. Hoppel (2010), The Role of the Middle Atmosphere in Simulations of the Troposphere during Northern Hemisphere Winter: Differences between High- and Low-Top Models, *J. Atmos. Sci.*, *67*, 3048–3064, doi:10.1175/2010JAS3255.1.

Scaife, A. A., J. R. Knight, G. K. Vallis, and C. K. Folland (2005a), A stratospheric influence on the winter NAO and North Atlantic surface climate, *Geophys. Res. Lett.*, *32*, 1–5.

Scaife, A. A., D. R. Jackson, R. Swinbank, N. Butchart, H. E. Thornton, M. Keil, and L. Henderson (2005b), Stratospheric vacillations and the major warming over Antarctica in 2002., *J. Atmos. Sci.*, *62*, 629–639, doi:10.1175/JAS-3334.1.

Scaife, A. A., C. K. Folland, L. V. Alexander, A. Moberg, and J. R. Knight (2008), European Climate Extremes and the North Atlantic Oscillation, *J. Climate*, *21*, 72–83, doi:10.1175/2007JCLI1631.1.

Scaife, A. A., T. Spangehl, D. R. Fereday, U. Cubasch, U. Langematz, H. Akiyoshi, S. Bekki, P. Braesicke, N. Butchart, M. P. Chipperfield, A. Gettelman, S. C. Hardiman, M. Michou, E. Rozanov, and T. G. Shepherd (2011a), Climate change projections and stratosphere-troposphere interaction, *Clim. Dyn.*, *38*, 2089–2097, doi:10.1007/s00382-011-1080-7.

Scaife, A. A., D. Copsey, C. Gordon, C. Harris, T. Hinton, S. Keeley, A. O'Neill, M. Roberts, and K. Williams (2011b), Improved Atlantic winter blocking in a climate model, *Geophys. Res. Lett.*, *38*, L23,703, doi:10.1029/2011GL049573.

Scaife, A. A., A. Arribas, E. Blockley, A. Brookshaw, R. Clark, N. Dunstone, R. Eade, D. Fereday, C. Folland, M. Gordon, L. Hermanson, J. Knight, D. Lea, C. MacLachlan, A. Maidens, M. Martin, A. Peterson, D. Smith, M. Vellinga, E. Wallace, J. Waters, and A. Williams. (2014), Skillful Long Range Prediction of European and North American Winters, *Geophys. Res. Lett.*, *in press*.

Scherhag, R. (1952), Die explosionsartigen Stratosphärenerwärmungen des Spätwinters 1951/52, *Berichte des Dtsch. Wetterdienstes der US-Zone*, *38*, 51–63.

Schoeberl, M., and D. Hartmann (1991), The dynamics of the stratospheric polar vortex and its relation to springtime ozone depletions, *Science*, *251*, 46–52, doi:10.1126/science.251.4989.46.

Scinocca, J., and P. Haynes (1998), Dynamical forcing of stratospheric planetary waves by tropospheric baroclinic eddies, *J. Atmos. Sci.*, *55*, 2361–2392.

Scott, R. K., and D. G. Dritschel (2006), Vortex-Vortex Interactions in the Winter Stratosphere, *J. Atmos. Sci.*, *63*, 726–740, doi:10.1175/JAS3632.1.

Scott, R. K., D. G. Dritschel, L. M. Polvani, and D. W. Waugh (2004), Enhancement of Rossby wave breaking by steep potential vorticity gradients in the winter stratosphere, *J. Atmos. Sci.*, *61*, 904–918.

Seviour, W. J. M., N. Butchart, and S. C. Hardiman (2012), The Brewer-Dobson circulation inferred from ERA-Interim, *Q. J. R. Meteorol. Soc.*, *138*, 878–888, doi:10.1002/qj.966.

Seviour, W. J. M., D. M. Mitchell, and L. J. Gray (2013), A practical method to identify displaced and split stratospheric polar vortex events, *Geophys. Res. Lett.*, *40*, 5268–5273, doi:doi:10.1002/grl.50927.

Seviour, W. J. M., S. C. Hardiman, L. J. Gray, N. Butchart, C. MacLachlan, and A. A. Scaife (2014), Skillful seasonal prediction of the Southern Annular Mode and Antarctic ozone, *J. Climate*, *27*, 7462–7474, doi:10.1175/JCLI-D-14-00264.1.

Shaw, T. A., J. Perlwitz, and N. Harnik (2010), Downward Wave Coupling between the Stratosphere and Troposphere: The Importance of Meridional Wave Guiding and Comparison with Zonal-Mean Coupling, *J. Clim.*, *23*, 6365–6381, doi:10.1175/2010JCLI3804.1.

Shaw, T. a., J. Perlwitz, N. Harnik, P. A. Newman, and S. Pawson (2011), The Impact of Stratospheric Ozone Changes on Downward Wave Coupling in the Southern Hemisphere *, *J. Clim.*, *24*(16), 4210–4229, doi:10.1175/2011JCLI4170.1.

Shaw, T. A., J. Perlwitz, and O. Weiner (2014), Troposphere-stratosphere coupling: Links to North Atlantic weather and climate, including their representation in CMIP5 models, *J. Geophys. Res.*, *119*, 5864–5880, doi:10.1002/2013JD021191.

Shepherd, T., R. A. Plumb, and S. C. Wofsy (2005), Preface, *J. Atmos. Sci.*, *62*, 565–566, doi:10.1175/JAS-9999.1.

Shepherd, T. G. (2014), Atmospheric circulation as a source of uncertainty in climate change projections, *Nat. Geosci.*, *7*, 703–708, doi:10.1038/ngeo2253.

Sigmond, M., J. F. Scinocca, and P. J. Kushner (2008), Impact of the stratosphere on tropospheric climate change, *Geophys. Res. Lett.*, *35*, L12,706, doi:10.1029/2008GL033573.

Sigmond, M., J. F. Scinocca, V. V. Kharin, and T. G. Shepherd (2013), Enhanced seasonal forecast skill following stratospheric sudden warmings, *Nat. Geosci.*, *6*, 98–102, doi:10.1038/NGEO1698.

Silvestri, G. E., and C. Vera (2003), Antarctic Oscillation signal on precipitation anomalies over southeastern South America, *Geophys. Res. Lett.*, *30*, 2115, doi:10.1029/2003GL018277.

Simmons, A., S. Uppala, D. Dee, and S. Kobayashi (2007), ERA-Interim: New ECMWF reanalysis products from 1989 onwards, *ECMWF Newsl.*, (110), 25–35.

Simpson, I. R., P. Hitchcock, T. G. Shepherd, and J. F. Scinocca (2011), Stratospheric variability and tropospheric annular-mode timescales, *Geophys. Res. Lett.*, *38*, L20,806, doi:10.1029/2011GL049304.

Smith, A. K. (1989), An investigation of resonant waves in a numerical model of an observed sudden stratospheric warming, *J. Atmos. Sci.*, *46*, 3038–3054.

Smith, D. M., A. A. Scaife, and B. P. Kirtman (2012), What is the current state of scientific knowledge with regard to seasonal and decadal forecasting?, *Environ. Res. Lett.*, *7*, 015,602, doi:10.1088/1748-9326/7/1/015602.

Solomon, S., R. R. Garcia, F. S. Rowland, and D. J. Wuebbles (1986), On the depletion of Antarctic ozone, *Nature*, *321*, 755–788, doi:10.1038/321755a0.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (Eds.) (2007), *IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, 996 pp., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Son, S.-W., A. Purich, H. H. Hendon, B.-M. Kim, and L. M. Polvani (2013), Improved seasonal forecast using ozone hole variability?, *Geophys. Res. Lett.*, *40*, 6231–6235, doi:10.1002/2013GL057731.

Song, Y., and W. A. Robinson (2004), Dynamical Mechanisms for Stratospheric Influences on the Troposphere, *J. Atmos. Sci.*, *61*, 1711–1725, doi:10.1175/1520-0469(2004)061.

Stewartson, K. (1978), The evolution of the critical layer of a Rossby wave, *Geophys. Astrophys. Fluid Dyn*, *9*, 185–200, doi:10.1080/03091927708242326.

Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Eds.) (2013), *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group 1 to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1535 pp., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Taguchi, M. (2003), Tropospheric Response to Stratospheric Degradation in a Simple Global Circulation Model, *J. Atmos. Sci.*, *60*, 1835–1846.

Taguchi, M. (2008), Is There a Statistical Connection between Stratospheric Sudden Warming and Tropospheric Blocking Events?, *J. Atmos. Sci.*, *65*, 1442–1454, doi:10.1175/2007JAS2363.1.

Taguchi, M. (2014), Predictability of Major Stratospheric Sudden Warmings of the Vortex Split Type: Case Study of the 2002 Southern Event and the 2009 and 1989 Northern Events, *J. Atmos. Sci.*, *71*, 2886–2904, doi:10.1175/JAS-D-13-078.1.

Taylor, K. E., R. J. Stouffer, and G. a. Meehl (2012), An Overview of CMIP5 and the Experiment Design, *Bull. Amer. Meteor. Soc.*, *93*, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tebaldi, C., and R. Knutti (2007), The use of the multi-model ensemble in probabilistic climate projections., *Philos. Trans. R. Soc. A*, *365*, 2053–2075, doi:10.1098/rsta.2007.2076.

Thompson, D. W. J., and S. Solomon (2002), Interpretation of recent Southern Hemisphere climate change., *Science*, *296*, 895–899, doi:10.1126/science.1069270.

Thompson, D. W. J., and J. M. Wallace (1998), The Arctic Oscillation signature in the wintertime geopotential height and temperature fields, *Geophys. Res. Lett.*, *25*, 1297–1300.

Thompson, D. W. J., and J. M. Wallace (2000), Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability, *J. Climate*, *13*(5), 1000–1016.

Thompson, D. W. J., and J. D. Woodworth (2014), Barotropic and Baroclinic Annular Variability in the Southern Hemisphere, *J. Atmos. Sci.*, *71*, 1480–1493, doi:10.1175/JAS-D-13-0185.1.

Thompson, D. W. J., J. M. Wallace, and G. C. Hegerl (2000), Annular modes in the extratropical circulation. Part II: Trends, *J. Climate*, *13*, 1018–1036.

Thompson, D. W. J., M. P. Baldwin, and J. M. Wallace (2002), Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction, *J. Climate*, *15*, 1421–1428.

Thompson, D. W. J., M. P. Baldwin, and S. Solomon (2005), Stratosphere-troposphere coupling in the Southern Hemisphere, *J. Atmos. Sci.*, *62*, 708–715, doi:10.1175/JAS-3321.1.

Thompson, D. W. J., J. C. Furtado, and T. G. Shepherd (2006), On the tropospheric response to anomalous stratospheric wave drag and radiative heating, *J. Atmos. Sci.*, *63*, 2616–2629.

Tibaldi, S., and F. Molteni (1990), On the operational predictability of blocking, *Tellus A*, *42A*, 343–365.

Tibshirani, R., G. Walther, and T. Hastie (2001), Estimating the number of clusters in a data set via the gap statistic, *J. R. Stat. Soc.*, *63*(Part 2), 411–423.

Tomassini, L., E. P. Gerber, M. P. Baldwin, F. Bunzel, and M. Giorgetta (2012), The role of stratosphere-troposphere coupling in the occurrence of extreme winter cold spells over northern Europe, *J. Adv. Model. Earth. Syst.*, *4*, M00A03, doi:10.1029/2012MS000177.

Toon, O. B., P. Hamill, R. P. Turco, and J. Pinto (1986), Condensation of HNO3 and HCl in the winter polar stratospheres, *Geophys. Res. Lett.*, *13*, 1284–1287.

Tung, K. K., and R. S. Lindzen (1979), A Theory of Stationary Long Waves. Part II: Resonant Rossby Waves in the Presence of Realistic Vertical Shears, *Mon. Weather Rev.*, *107*, 735–750.

Uppala, S. M., P. W. Kallberg, A. J. Simmons, U. Andrae, V. D. C. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. V. D. Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. Mcnally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.

Wallace, J., and D. Thompson (2002), The Pacific center of action of the Northern Hemisphere annular mode: Real or artifact?, *J. Clim.*, *15*, 1987–1991.

Warn, T., and H. Warn (1978), The evolution of a nonlinear critical level, *Stud. Appl. Math.*, *59*, 37–71.

Waugh, D. W. (1997), Elliptical diagnostics of stratospheric polar vortices, *Q. J. R. Meteorol. Soc.*, *123*, 1725–1748.

Waugh, D. W., and W. J. Randel (1999), Climatology of Arctic and Antarctic polar vortices using elliptical diagnostics, *J. Atmos. Sci.*, *56*, 1594–1613.

Waugh, D. W., R. A. Plumb, R. J. Atkinson, M. R. Schoeberl, L. R. Lait, P. A. Newman, M. Loewenstein, D. W. Toohey, L. M. Avallone, C. R. Webster, and R. D. May (1994), Transport out of the lower stratospheric Arctic vortex by Rossby wave breaking, *J. Geophys. Res.*, *99*, 1071–1088.

Waugh, D. W., L. Oman, P. A. Newman, R. S. Stolarski, S. Pawson, J. E. Nielsen, and J. Perlwitz (2009), Effect of zonal asymmetries in stratospheric ozone on simulated Southern Hemisphere climate trends, *Geophys. Res. Lett.*, *36*, L18,701, doi:10.1029/2009GL040419.

Weber, M., S. Dikty, J. P. Burrows, H. Garny, M. Dameris, A. Kubin, J. Abalichin, and U. Langematz (2011), The Brewer-Dobson circulation and total ozone from seasonal to decadal time scales, *Atmos. Chem. Phys.*, *11*, 11,221–11,235, doi:10.5194/acp-11-11221-2011.

Wilcox, L. J., B. J. Hoskins, and K. P. Shine (2012), A global blended tropopause based on ERA data. Part I: Climatology, *Q. J. R. Meteorol. Soc.*, *138*, 561–575, doi:10.1002/qj.951.

Wilks, D. S. (2006), *Statistical Methods in the Atmospheric Sciences*, 2 ed., 627 pp., Academic Press.

Wittman, M. A. H., A. J. Charlton, and L. M. Polvani (2007), The Effect of Lower Stratospheric Shear on Baroclinic Instability, *J. Atmos. Sci.*, *64*(2), 479–496, doi:10.1175/JAS3828.1.

WMO (2011), Scientific Assessment of Ozone Depletion: 2010, *Tech. rep.*, World Meteorological Organization, Global Ozone Reasearch and Monitoring Project, Report 52., Geneva.

Woollings, T., A. Charlton-Perez, S. Ineson, A. G. Marshall, and G. Masato (2010), Associations between stratospheric variability and tropospheric blocking, *J. Geophys. Res.*, *115*, D06,108, doi:10.1029/2009JD012742.

Yamazaki, K. (1987), Observations of the Stratospheric Final Warmings in the Two Hemispheres, *J. Meteor. Soc. Japan*, *65*, 51–66.

Yoden, S., T. Yamaga, P. S, and U. Langematz (1999), A composite analysis of the stratospheric sudden warmings simulated in a perpetual January integration of the Berlin TSM GCM, *J. Meteor. Soc. Japan*, *77*, 431–445.