

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

William Seviour

Atmospheric, Oceanic and Planetary Physics  
Department of Physics  
University of Oxford

Draft compiled on:

July 29, 2014



---

## Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Dynamics of the polar stratosphere . . . . .	1
1.2	Stratosphere-troposphere coupling . . . . .	1
<b>2</b>	<b>A geometrical description of vortex variability</b>	<b>3</b>
2.1	Introduction . . . . .	3
2.2	The $Z_{10}$ method . . . . .	5
2.2.1	Vortex geometry calculation . . . . .	5
2.2.2	Event definition . . . . .	6
2.3	Analysis . . . . .	7
2.4	Summary . . . . .	9
<b>3</b>	<b>Representation of vortex variability in climate models</b>	<b>19</b>
3.1	Introduction . . . . .	19
3.1.1	CMIP5 model simulations . . . . .	21
3.2	Vortex mean state and variability . . . . .	24
3.2.1	Moment diagnostics . . . . .	24
3.2.2	Displaced and split vortex events . . . . .	25
3.3	Stratosphere-troposphere coupling . . . . .	30
3.3.1	Zonal-mean response to displaced and split vortex events . .	30
3.3.2	Spatial response to displaced and split vortex events . . .	33
3.4	Discussion . . . . .	41
3.4.1	Effect of model resolution . . . . .	41
3.4.2	Measures of stratosphere-troposphere coupling . . . . .	45

3.4.3	Difference between split and displaced vortex events . . . . .	47
3.5	Conclusions . . . . .	47
<b>4</b>	<b>The role of the stratosphere in seasonal prediction</b>	<b>49</b>
4.1	Introduction . . . . .	49
4.2	Seasonal forecast system . . . . .	52
4.3	Seasonal forecast results . . . . .	53
4.3.1	Stratospheric polar vortex . . . . .	53
4.3.2	Ozone depletion . . . . .	56
4.3.3	Southern Annular Mode . . . . .	58
4.3.4	Stratosphere-troposphere coupling . . . . .	60
4.4	Discussion . . . . .	63
4.5	Conclusions . . . . .	66
	<b>Bibliography</b>	<b>73</b>

---

## List of Figures

---

2.1	Distributions of $Z_{10}$ and PV-based moment diagnostics. . . . .	6
2.2	Seasonal distribution of displaced and split vortex events. . . . .	7
2.3	PV composites for split and displaced vortex events. . . . .	10
2.4	Clustering algorithms applied to the moment diagnostics. . . . .	12
2.5	NAM composites for split and displaced vortex events. . . . .	13
2.6	Significance of surface NAM difference following split and displaced vortex events. . . . .	14
2.7	Mean sea-level pressure composites for split and displaced vortex events. . . . .	15
2.8	Tropopause height composites for split and displaced vortex events. .	16
3.1	Distributions of moment diagnostics for the CMIP5 models. . . . .	22
3.2	Seasonal cycle of moment diagnostics in the CMIP5 models . . . . .	25
3.3	Frequency of split and displaced vortex events in the CMIP5 models	26
3.4	Seasonal distribution of splits and displacements in the CMIP5 models . . . . .	28
3.5	Comparison of moment diagnostics and frequency of split and displaced vortex events. . . . .	30
3.6	Composites of 10 hPa Z at the onset date split and displaced vortex events in the CMIP5 models. . . . .	31
3.7	NAM composites for splits and displacements in the CMIP5 models	34
3.8	MSLP composites following splits and displacements in the CMIP5 models . . . . .	36
3.9	NAM composites for splits and displacements in the CMIP5 models	39

3.10 Difference of MSLP following split and displaced vortex events. . . . .	42
3.11 Relationship between model resolution and moment diagnostics in the CMIP5 models. . . . .	46
3.12 Vertical resolution and modal aspect ratio. . . . .	47
4.1 Comparison of GloSea5 and ERA-Interim zonal-mean zonal wind climatologies. . . . .	53
4.7 Comparison of GloSea5 and ERA-Interim zonal-mean zonal wind climatologies. . . . .	58
4.8 GloSea5 forecast skill for the stratospheric polar vortex strength and column ozone. . . . .	59
4.9 Comparison of GloSea5 and observed SSWs. . . . .	60
4.10 Relation between stratospheric polar vortex strength and column ozone. . . . .	61
4.11 GloSea5 predictions of the SAM. . . . .	62
4.12 Correlation of GloSea5 forecasts of sea-level pressure and tempera- ture. . . . .	63
4.13 Lag-height correlation of GloSea5 polar cap geopotential height . . .	64
4.14 Variation of GloSea5 forecast skill with ensemble size . . . . .	66
4.15 Timeseries of $\bar{U}$ at $60^{\circ}\text{N}$ , 10 hPa, for all GloSea5 ensemble members.	68
4.16 Timeseries of $\bar{U}$ at $60^{\circ}\text{S}$ , 10 hPa, for all GloSea5 ensemble members.	69
4.17 Lag-height correlation of $Z'$ at 10 hPa and 1000 hPa. . . . .	70
4.18 GloSea5 predictions of the SAM. . . . .	71
4.19 Correlation of the SOI with the SAM. . . . .	71
4.20 Lag-height correlation of GloSea5 polar cap geopotential height . .	72

# CHAPTER 1

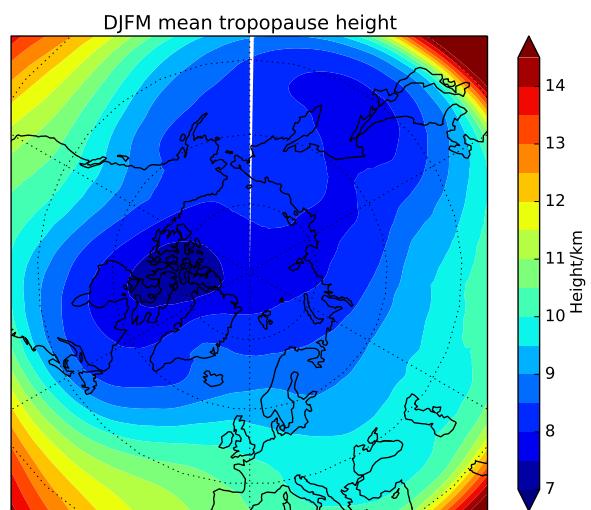
---

## Introduction

---

### **1.1 Dynamics of the polar stratosphere**

### **1.2 Stratosphere-troposphere coupling**



**Figure 1.1**

# CHAPTER 2

---

## A geometrical description of vortex variability

---

Much of the work contained in this chapter is based upon *Seviour et al.* [2013], although the analysis here has been significantly extended and re-written.

### 2.1 Introduction

Stratospheric sudden warmings (SSWs) are extreme events in which the strong westerly winds that usually dominate the winter polar stratosphere become highly disturbed (here, for reasons outlined below, we use the term SSW to encompass a wider range of variability than its traditional definition). These events lead to the mixing of mid-latitude air into the polar vortex region, causing an increase in temperatures by several tens of kelvin over the course of a few days. Traditional methods to identify stratospheric sudden warmings (SSWs) have relied on either zonal-mean [*Andrews et al.*, 1987] or annular mode [*Baldwin and Dunkerton*, 2001] diagnostics. Neither method explicitly deals with the inherent zonal asymmetry in vortex variability. In particular, SSWs are observed to occur in one of two manners: displaced vortex events, where the vortex moves far from the pole, and split vortex events, where the vortex separates into two ‘child’ vortices. These two types have a very different spatial structure and evolution timescale [*Matthewman et al.*, 2009]. Displaced and split vortex events are predominantly associated with vertically propagating Rossby waves of wavenumber 1 and 2 respectively,

and many previous studies have classified SSWs based on wavenumber [e.g. *Nakagawa and Yamazaki, 2006*]. However, this method does not provide a description of the location of the polar vortex itself, which theoretical arguments suggest may be important for understanding stratosphere-troposphere coupling [*Ambaum and Hoskins, 2002*]. In an improvement to these traditional SSW definitions, *Charlton and Polvani [2007]* (hereafter CP07) introduced a classification in which a split vortex event is identified when two vortices with a circulation ratio of 2:1 or higher are present, and all other SSWs are automatically classed as displaced vortex events. However, they maintained the traditional SSW identification which requires there to be a reversal of the zonal-mean zonal wind at 10 hPa and 60°N.

An increased understanding of stratospheric variability can be gained by using vortex-centric diagnostics, such as two-dimensional (2D) vortex moments [*Waugh, 1997; Waugh and Randel, 1999; Mitchell et al., 2011a,b*], which provide a geometrical description of the vortex and have no reliance on zonal-mean properties. Using a classification based on these diagnostics, *Mitchell et al. [2013]* (hereafter M13) identified a greater number of SSWs than CP07. This is primarily because they did not use a zonal mean threshold criterion. Importantly, M13 also 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. Their result supported a similar conclusion by *Nakagawa and Yamazaki [2006]*, who found that the impact of events associated with an enhanced upward flux of wavenumber-2 planetary waves was more likely to reach the surface. These results underline the need to correctly identify the precise type of SSW, in order to understand stratosphere-troposphere coupling within climate models.

Distinguishing between displaced and split vortex events using the method of M13 requires the use of potential vorticity (PV), which is not commonly output by climate models. For this reason, previous attempts to apply PV-based techniques in a multi-model study have led to the majority of models being excluded [*Mitchell et al., 2012*]. Furthermore, their method used a hierarchical clustering technique [*Hannachi et al., 2010*], which is very sensitive to the exact shape of the distribution of vortex variability, so is unsuitable for application to a range of models with different climatologies. In this chapter, we develop an improved method which; (a), is based on the geometry of the vortex, but requires only the 10 hPa geopotential height; and (b), identifies events using a simple threshold instead of a clustering technique. We apply this new method to the ERA-40 and ERA-Interim reanalysis

## 2.2. THE $Z_{10}$ METHOD

datasets and demonstrate that the method captures a similar number of events which are in good agreement with, and at least as extreme as, those of M13.

## 2.2 The $Z_{10}$ method

### 2.2.1 Vortex geometry calculation

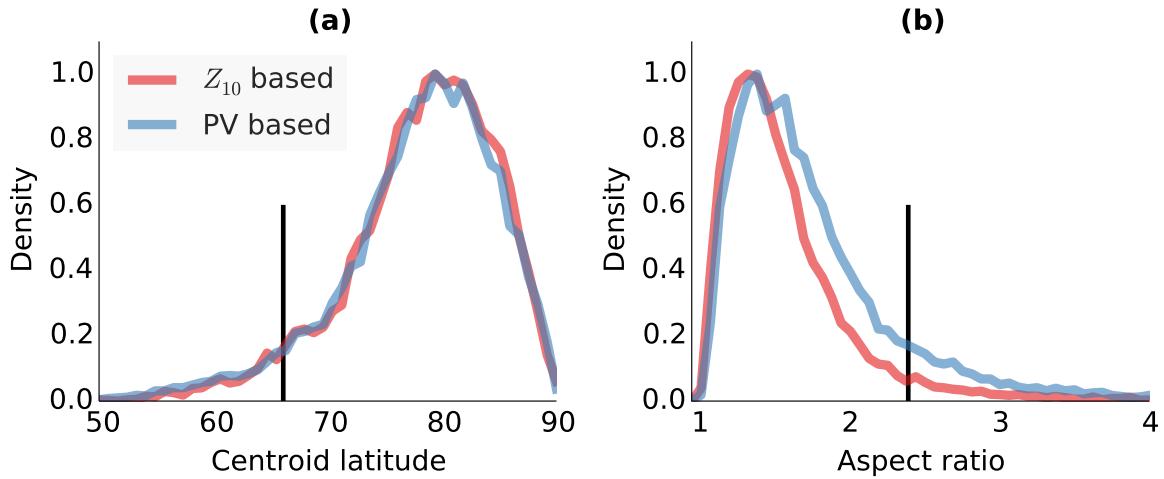
Northern Hemisphere winter daily-mean data for December-March (DJFM) were employed. The European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-40 dataset [*Uppala et al., 2005*] is used from 1958-1979 and ERA-Interim [*Dee et al., 2011*] from 1979-2009. The geometry of the vortex is described in terms of the latitude of the centre of the vortex (centroid latitude) and aspect ratio (a measure of how stretched the vortex is).

The method calculates the 2D vortex moment of order  $a + b$ , defined in Cartesian coordinates as

$$\mu_{ab} = \iint_S [X(x, y) - \bar{X}] x^a y^b dx dy, \quad (2.1)$$

where  $\bar{X}$  is a single contour of PV or geopotential height, representing the vortex edge. The value of this contour is chosen to match that of the 10 hPa geopotential height ( $Z_{10}$ ) or 850 K PV ( $PV_{850}$ ) DJFM zonal-mean at 60°N (however, the results are not sensitive to the exact choice of contour between about 50°-70°N). Following the method of *Matthewman et al. [2009]* (to which readers are referred for technical details), Equation 2.1 allows for the explicit calculation of the centroid latitude and aspect ratio on a particular pressure level or isentropic surface. Other moment diagnostics can be extracted, such as vortex area and kurtosis, but these were not found necessary for identifying split and displaced vortex events, and are therefore omitted from the analysis.

Figure 2.1 compares the distribution of aspect ratio and centroid latitude, calculated using daily  $PV_{850}$  and  $Z_{10}$ . The centroid latitude distributions are almost identical, with a peak around 80°N. 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 to be expected, since the PV field contains more small-scale, filamentary structures than geopotential height. Neither distribution shows bi-modality, suggesting that the application of clustering techniques, as in *Coughlin and Gray [2009]* and *Hannachi et al. [2010]*, may be inappropriate. As well as having similar distributions, the timeseries of the PV and geopotential height derived diagnostics



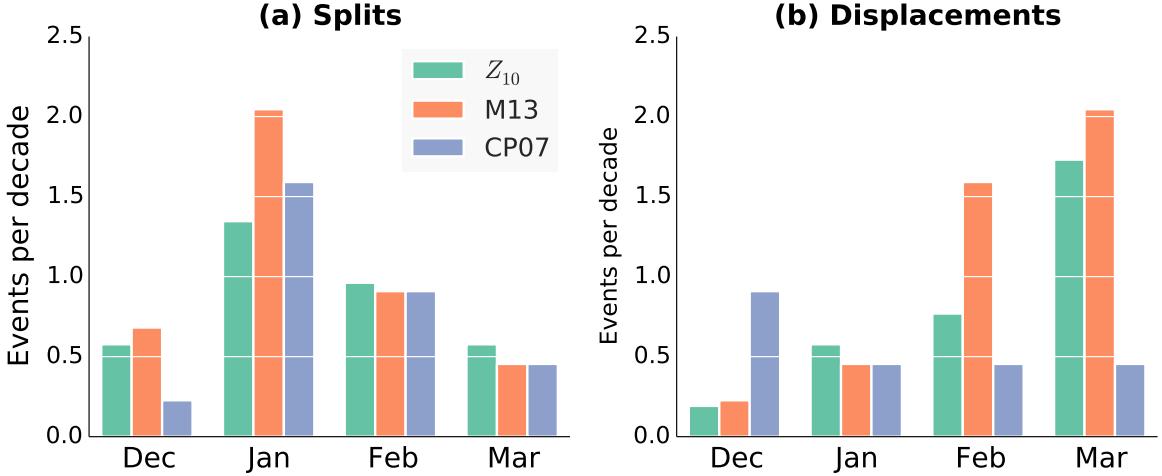
**Figure 2.1:** Distributions of the December-March centroid latitude (a) and aspect ratio (b), of the Northern Hemisphere stratospheric polar vortex over 1958-2009. Diagnostics are calculated from geopotential height at 10 hPa ( $Z_{10}$ ) and potential vorticity at 850 K (PV). Thresholds of  $66^{\circ}\text{N}$  in centroid latitude and 2.4 in aspect ratio are used to define events, and are indicated by the black vertical lines.

(not shown) are well correlated, with correlation coefficients of 0.9 for centroid latitude and 0.6 for aspect ratio.

## 2.2.2 Event definition

In order to identify displaced and split vortex events, a threshold criterion is introduced and applied to the geopotential height derived diagnostics. A displaced vortex event requires the centroid latitude to remain equatorward of  $66^{\circ}\text{N}$  for 7 days or more. A split vortex event requires the aspect ratio to remain higher than 2.4 for 7 days or more. These thresholds are indicated in Figure 2.1, and were selected to give a similar frequency of displacement and splitting events as M13. This choice is somewhat subjective, but the results presented below are not sensitive to the exact choice of threshold. There were no occasions on which both criteria were met simultaneously. The onset date is defined as the day that the appropriate threshold is first exceeded, and to ensure that no events are counted twice, the events are required to be spaced at least 30 days apart, chosen to reflect radiative timescales in the lower stratosphere [Newman and Rosenfield, 1997]. Mitchell *et al.* [2011a] found that the aspect ratio and centroid latitude follow an extreme value distribution [Coles, 2001] and we note that both thresholds chosen here lie beyond the extreme value thresholds of their respective distributions. Using this method, 17 displaced and 18 split vortex events (listed in Table 2.1) are

### 2.3. ANALYSIS



**Figure 2.2:** Histogram of the seasonal distribution of displaced and split vortex events, from the  $Z_{10}$  method, M13 and CP07.

identified over the 52 winters, an average of 7 per decade. These events are all mid-winter events – the DJFM data were used explicitly to avoid counting final warmings. This frequency lies between the values of CP07 (6/decade) and M13 (8/decade).

There are significant differences in the seasonal distribution of displaced and split vortex events, as shown in Figure 2.2. Split vortex events are more frequent in early-mid winter, with a peak in January, while displaced vortex events are skewed towards late winter. Figure 2.2 also shows the seasonal distribution of displaced and split vortex events from CP07 and M13. The shape of the distribution agrees well with the seasonal distribution found by M13 for both types of event. However, there is less similarity with the CP07 distribution of displaced vortex events (Figure 2.2b). CP07 indicates an approximately flat distribution throughout winter, and many fewer displaced vortex events overall. It should be noted that the seasonal distribution of displaced and split vortex events from the moment based methods does not arise from the underlying climatology of centroid latitude or aspect ratio, which remains approximately constant throughout winter [Mitchell *et al.*, 2011a].

## 2.3 Analysis

To evaluate how well the new method captures displaced and split vortex events, Figure 2.6 shows composites of PV in the mid-stratosphere (850 K) following their onset dates. These are averaged over the 5 days following the onset date for split

No.	Event onset	Event type	$\Delta T_{10}$ (K)	$\bar{U}_{10}$ (m s <sup>-1</sup> )
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 2.1:** A summary table of displaced (D) and split (S) vortex events, identified using the Z10 method for data from 1958-2009.  $\Delta T_{10}$  represents the mean area-weighted 50°-90°N cap temperature anomaly at 10 hPa calculated 5 days either side of the event onset date.  $\bar{U}_{10}$  represents  $\bar{U}$  at 60°N and 10 hPa averaged over the same period. Stars (\*) represent those numbers that do not coincide (within 10 days) with events defined by CP07.

#### 2.4. SUMMARY

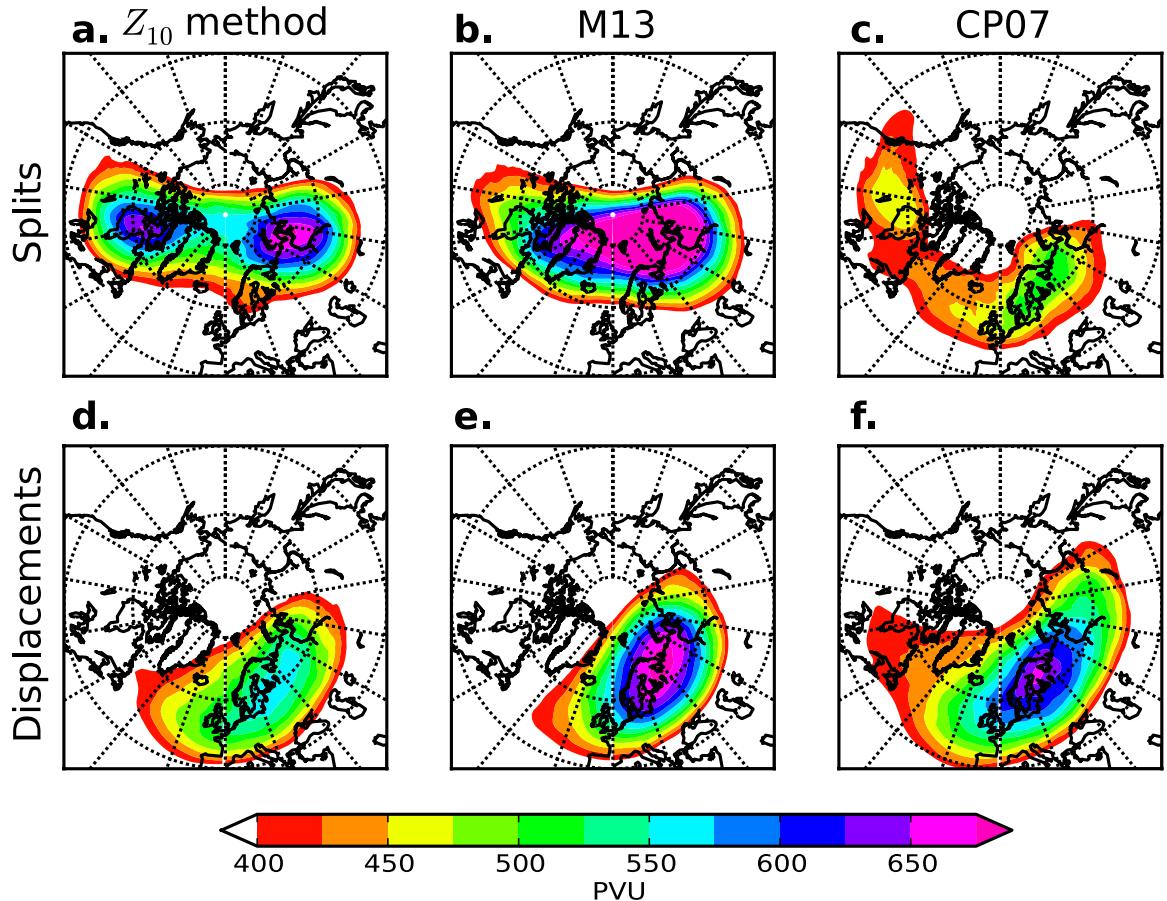
events and 7 days for displaced events (these averaging periods reflect the different timescales of the events). The composites are compared with the corresponding composites following the events identified by M13 and CP07. For the split vortex events, the  $Z_{10}$  method clearly shows two separated vortices, one centred over Canada and the other over Siberia. This characteristic direction reflects the climatological wave-2 pattern at this altitude. For the M13 events the split vortex composite shows the vortex stretched across the same  $90^{\circ}\text{W}$ - $90^{\circ}\text{E}$  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, and is similar 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.

Figure 2.6 demonstrates that the  $Z_{10}$  method succeeds in identifying displaced and split vortex events events as well as, and in some cases better than, the methods 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 the extreme events (M13).

## 2.4 Summary

Recent research has demonstrated the need to distinguish between displaced and split stratospheric polar vortex events, due to their different impacts on surface weather patterns. However, current methods to identify these events are complex or require non-standard variables. In this chapter, a new, robust method has been developed to identify displaced and split vortex events, which requires only geopotential height at 10 hPa. The method is briefly summarised as follows:

1. To identify the vortex region, a single contour of 10 hPa geopotential height is selected: the value of the zonal mean at  $60^{\circ}\text{N}$ .
2. Using this contour and the geopotential height field, the centroid latitude and aspect ratio are calculated, following the methodology of *Matthewman et al. [2009]*.
3. Events are identified using a threshold criterion: Displaced events are identified if the centroid latitude remains equatorward  $66^{\circ}\text{N}$  for 7 days or more.

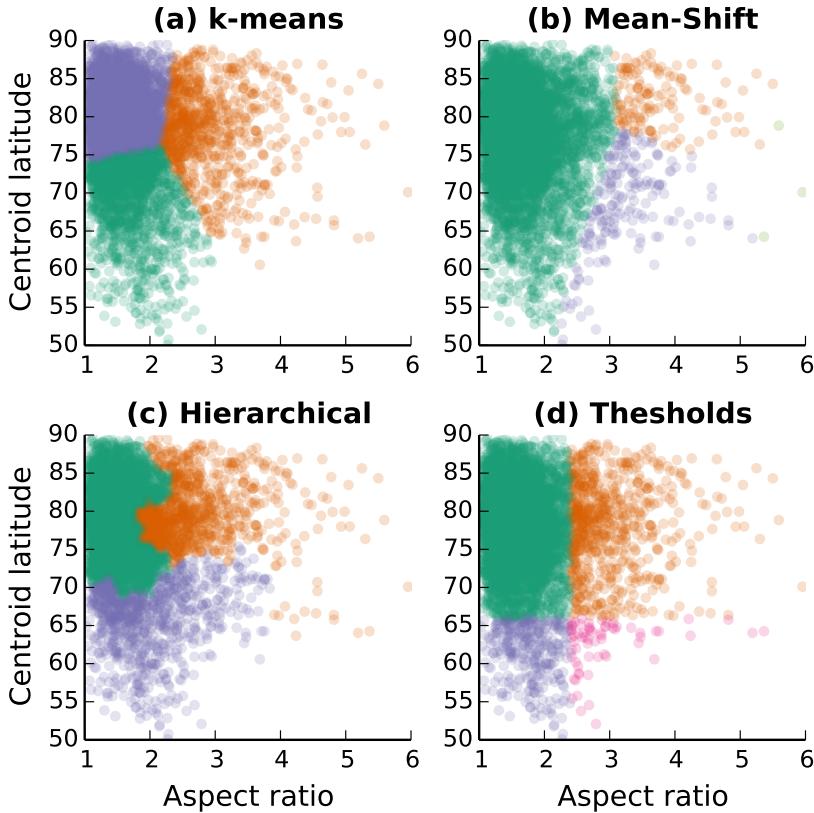


**Figure 2.3:** Composites of potential vorticity at the 850 K isentropic surface from the ECMWF reanalyses over 1958-2009. Composites are taken over the 5 days following the onset date of split vortex events (a,b,c) and 7 days following displaced vortex events (the difference is due to the different timescales of these events). We compare the current ( $Z_{10}$ ) method (a,b) with that of M13 (b,e) and CP07 (c,f).

#### 2.4. SUMMARY

Split events are identified if the aspect ratio remains above 2.4 for 7 days or more. No two events may occur within 30 days.

Results using this method demonstrate that it is able to identify split and displaced vortex events at least as effectively as previous methods.

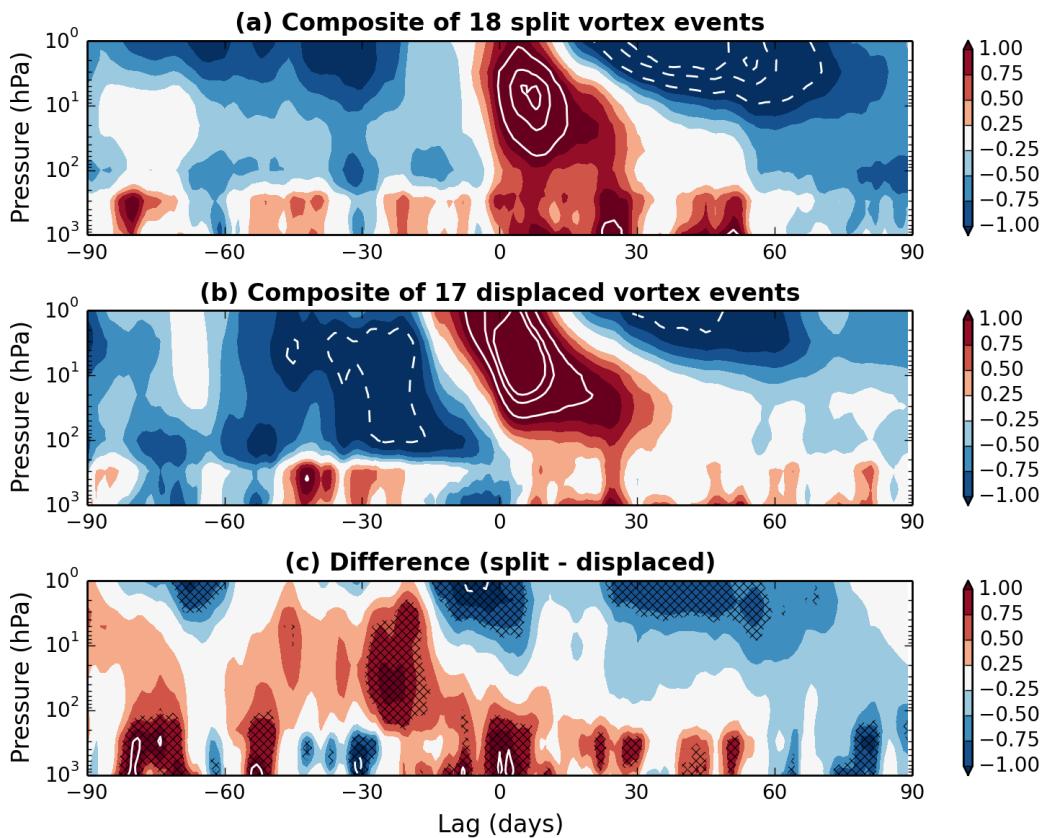


**Figure 2.4:** Three clustering algorithms and a threshold division applied to the moment diagnostics in centroid latitude-aspect ratio space. For the *k*-means and hierarchical algorithms three clusters were specified. The mean-shift algorithm determined the number of clusters to be 4.

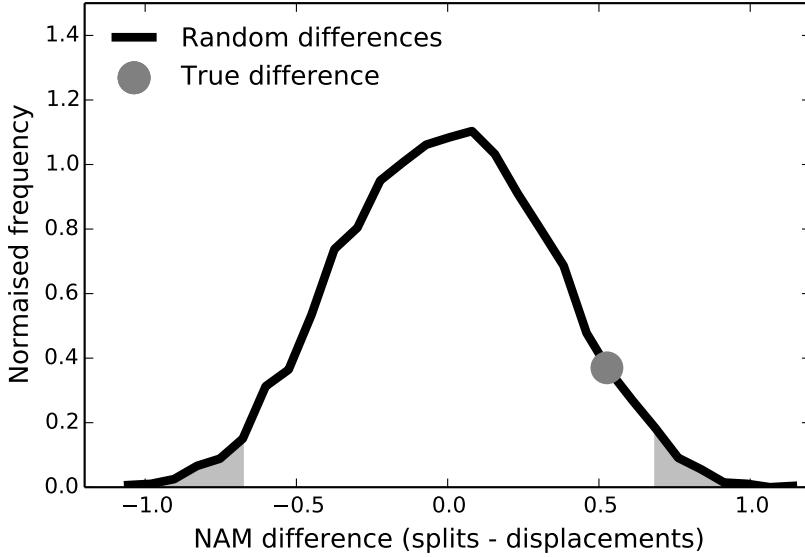
M13 found that the mean sea-level pressure (MSLP) anomalies are different before and after displaced and split vortex events. In Figure ?? we present composites of MSLP 30 days before and 30 days following the onset dates of displaced and split vortex events, calculated using the  $Z_{10}$  method, from the ECMWF re-analyses. Statistical significance is estimated from a Monte-Carlo method, using  $10^5$  composites of equal size, formed from randomly sampled winter dates. The strongest precursor is found for displaced vortex events, with a wave-1 pattern that is similar to the 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 North Atlantic Oscillation (though with a southern centre of action shifted towards Europe).

The evolution of anomalies throughout the depth of the atmosphere can be investigated through the use of annular modes [e.g. *Baldwin and Dunkerton, 2001*].

## 2.4. SUMMARY



**Figure 2.5:** Composites of the time-height evolution of the NAM during (a) 17 vortex displacement events and (b) 18 splitting events. 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. Data is from the ECMWF Reanalyses 1958-2009.

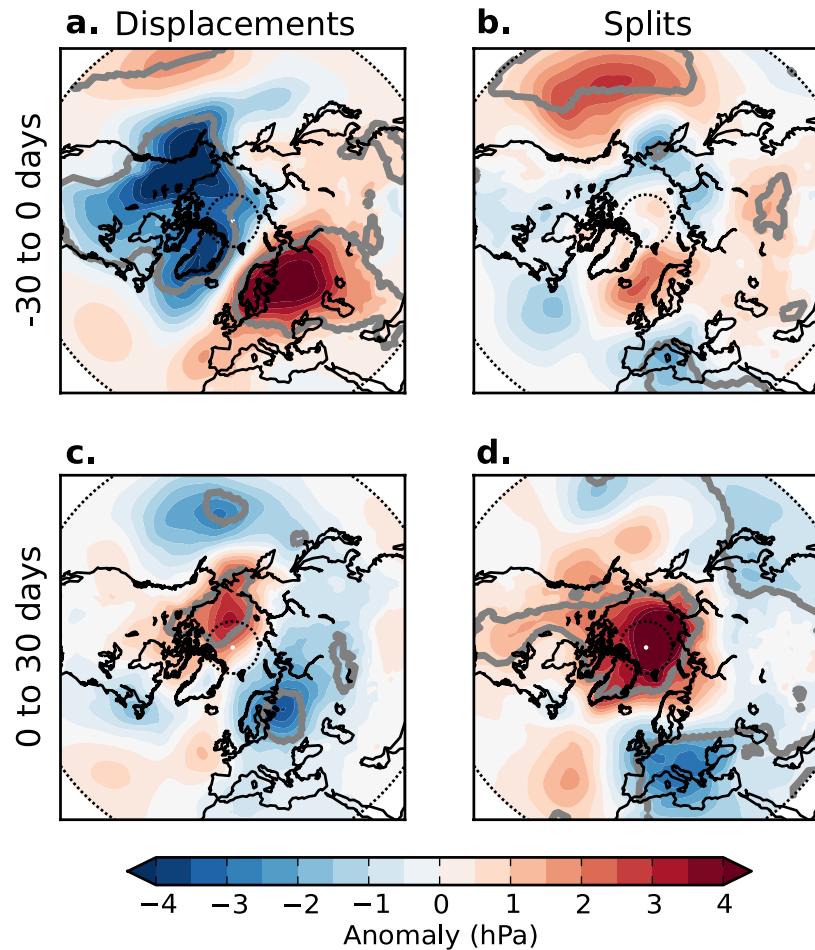


**Figure 2.6:** Distribution of 0-30 day mean NAM composite differences between split and displaced vortex events, formed by randomly shuffling the labels ‘split’ and ‘displacement’ between events. The 95% significant region (according to a two-tailed test) is shaded and the true composite difference is at the 94th percentile.

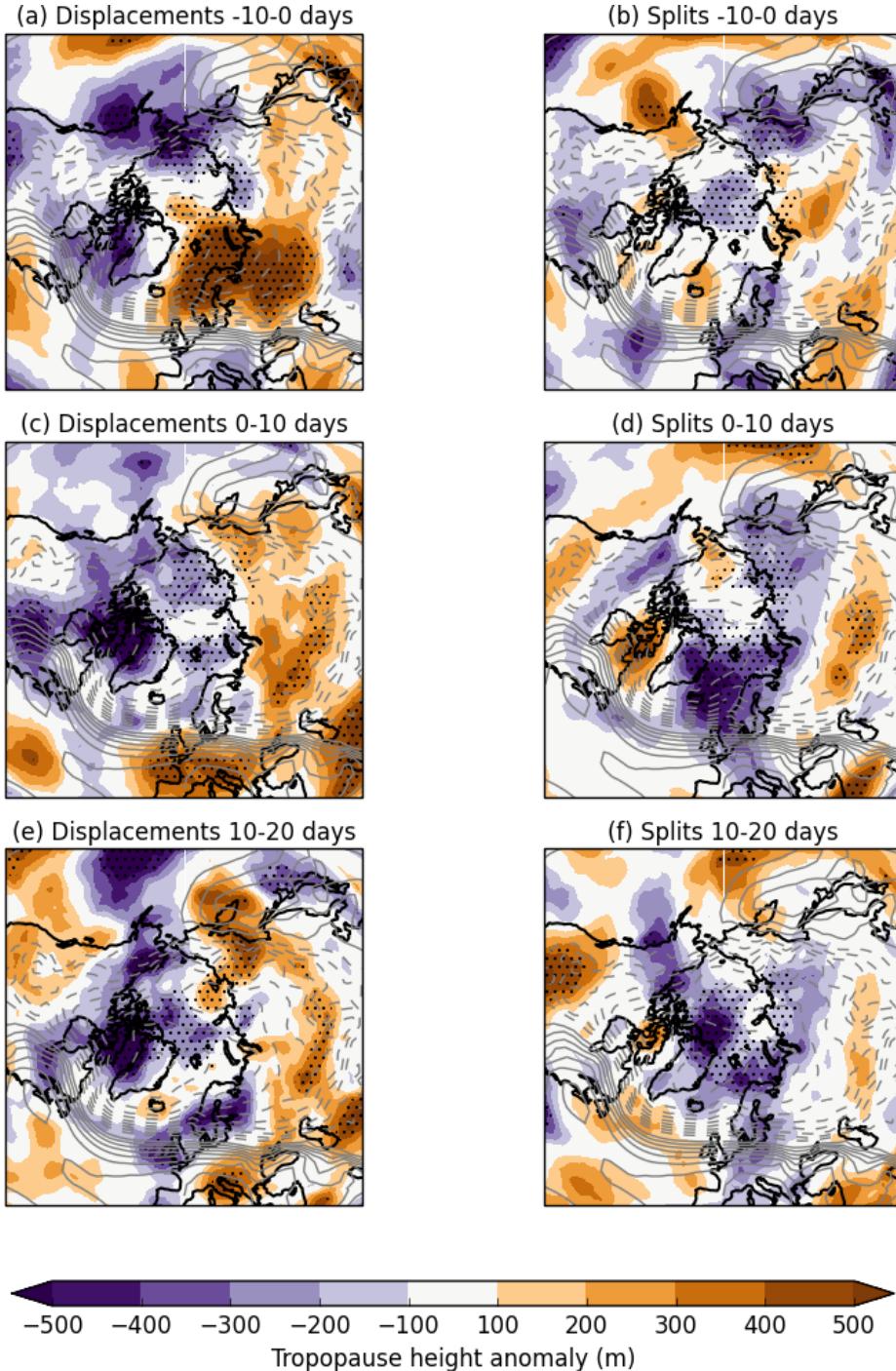
The Northern Annular Mode (NAM) (known as the Arctic Oscillation at the surface) is the leading mode of variability in the wintertime variability of the Northern Hemisphere circulation. Here we use the method of [Baldwin and Thompson, 2009], who define the NAM as the leading empirical orthogonal function (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 at each level so that the entire time series has unit variance.

Figure ?? presents time-height composites of NAM anomalies 90 days either side of the onset date of displaced and split vortex events. Despite the larger stratospheric anomaly following displaced vortex events, the tropospheric signal is larger following split vortex events (as in Figure ??). The vertical evolution of these events also differs greatly, with anomalies descending from the upper stratosphere over a period of weeks for displacement events, while split vortex events are near barotropic. This suggests an excitation of the barotropic mode, which supports the idea of the wave resonance view of SSWs [Esler and Scott, 2005].

## 2.4. SUMMARY



**Figure 2.7:** 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 is from the ECMWF reanalyses (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 Monte-Carlo significance test.



**Figure 2.8:** 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. 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.

#### 2.4. SUMMARY

Overall, these results support the findings of M13, but use a new method to identify split and displaced vortex events, and extend the analysis to 2009. The fact that the results are consistent suggests the findings of M13 are robust and highlights the importance of distinguishing split and displaced vortex events in understanding stratosphere-troposphere coupling following SSWs.

CHAPTER 2. A GEOMETRICAL DESCRIPTION OF VORTEX VARIABILITY

# CHAPTER 3

---

## Representation of vortex variability in climate models

---

### 3.1 Introduction

Over the past decade an increasing number of climate models have included a well-resolved stratosphere, with model lids above the stratopause. For example, the fifth Coupled Model Intercomparison Project (CMIP5) [Taylor *et al.*, 2012] includes 15 models with an uppermost level above the stratopause, whereas the previous intercomparison project, CMIP3, includes only 5 [Cordero and Forster, 2006]. The CMIP5 model simulations are significant in that they are motivated by, and evaluated in the the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) [Stocker *et al.*, 2013]. This change in model stratospheric resolution has been largely motivated by an increased understanding of the stratosphere's influence on tropospheric climate (discussed in Gerber *et al.* [2012] and Chapters ?? and 2).

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. They attribute 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.

These results are supported by similar studies which compare 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 extratropics as well as decadal trends in the NAO were more realistically simulated by the high-top than the low-top model. These differences have again been linked to the different simulation of SSW events in high- and low-top models [*Sassi et al.*, 2010].

Other studies have compared simulations of climate change with high- and low-top models. *Huebener et al.* [2007] linked a 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. Interestingly, *Manzini et al.* [2014] also show that global surface temperature trends under climate change (and so climate sensitivity) are larger for high-top compared to low-top models. However, when comparing pairs of high- and low-top versions of the same (or similar) models, little in climate sensitivty difference is detected, and they conclude that this difference is unlikely to have a physical basis.

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. Indeed, *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 2 to evaluate the representation of stratospheric polar vortex variability in contemporary climate models. We wish to perform as much as possible, a like-for-like comparison, and so 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, including split and displaced vortex events. We also consider the influence of these events on the troposphere in each model.

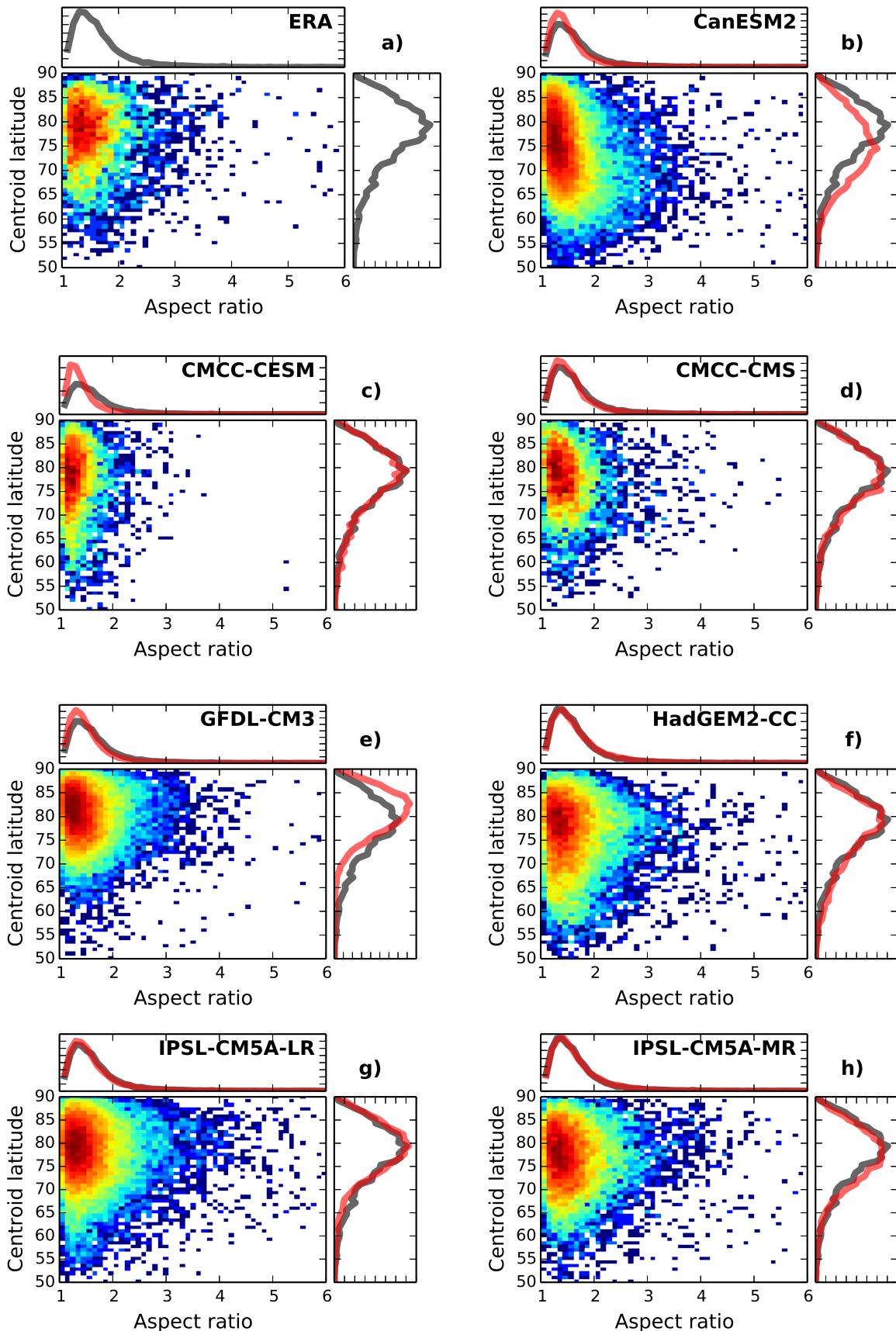
### 3.2. VORTEX MEAN STATE AND VARIABILITY

There are three main objectives to this investigation. First, we wish to evaluate the current state of model's 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 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 2.

#### 3.1.1 CMIP5 model simulations

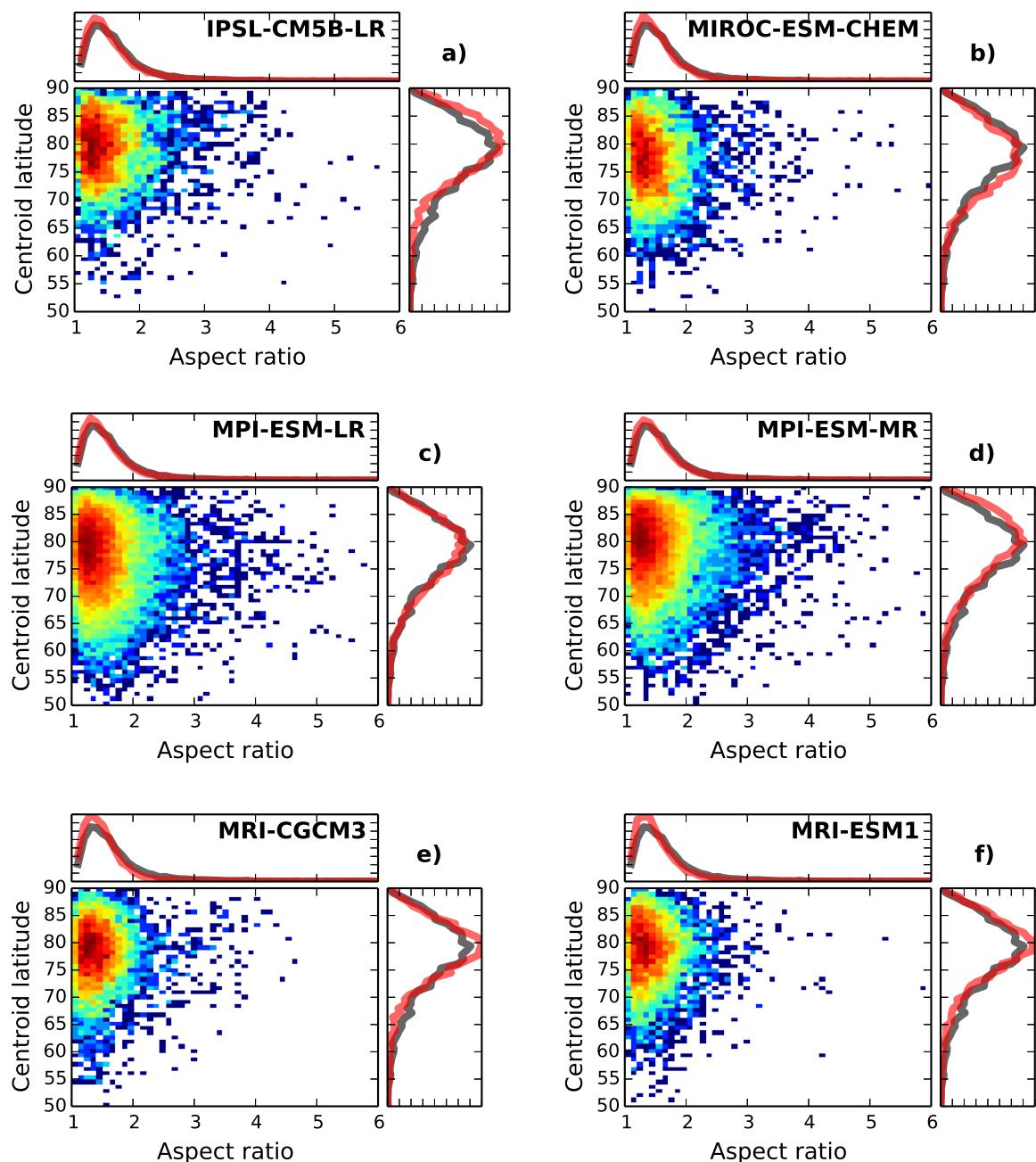
For this analysis only climate models with a lid height above the stratopause are selected from the CMIP5 ensemble. In total, 13 such models were available from 8 different modelling centres, although another two (CESM1-WACCM and MIROC-ESM) are listed in the CMIP5 ensemble, appropriate data was not found to be available for these models in the CMIP5 archive (<http://pcmdi3.llnl.gov/esgcet/home.htm>). These models are listed in Table 3.1. 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 is very close to the stratopause.

Historical simulations have been used throughout this analysis. These include observed climate forcings, such as from greenhouse gasses, 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), therefore allowing for a more direct comparison with observations. In order to achieve the largest possible ensemble size, all available ensemble members have been used for each model, which leads to a differing number of years entering the ensemble from different models.



**Figure 3.1:** Distributions of centroid latitude and aspect ratio for the ERA (grey lines) (a) and the CMIP5 models (red lines). Joint distributions are shown with a logarithmic scale such that red squares represent the densest regions.

3.2. VORTEX MEAN STATE AND VARIABILITY



**Figure 3.1:** (Continued)

Model	Ensemble size	Lid/ km	Levels	dh/km	dz <sub>1</sub> /km	dz <sub>2</sub> /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 3.1:** 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. Horizontal resolution, dh, is estimated at 45°N and vertical resolution is shown averaged over two regions; 5-15 km (dz<sub>1</sub>) and 15-30 km (dz<sub>2</sub>).

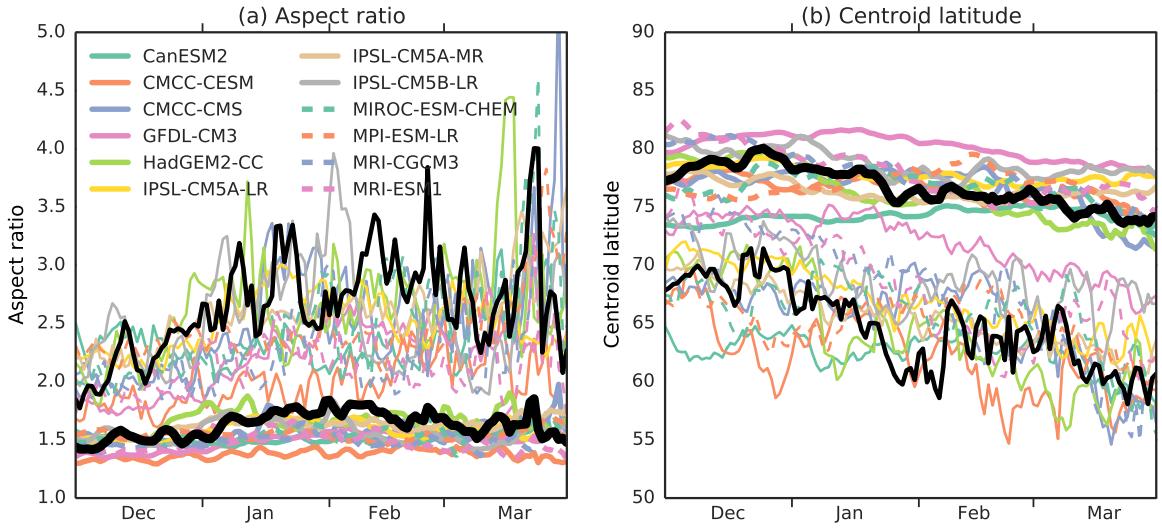
## 3.2 Vortex mean state and variability

### 3.2.1 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 2.2.1. 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 3.1, along with that from the ERA-40/ERA-Interim reanalysis (hereafter ERA) calculated in Chapter 2. 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 2, it can be seen that the joint distribution for ERA has an approximately triangular distribution with high 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

### 3.2. VORTEX MEAN STATE AND VARIABILITY



**Figure 3.2:** 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.

latitude being more common.

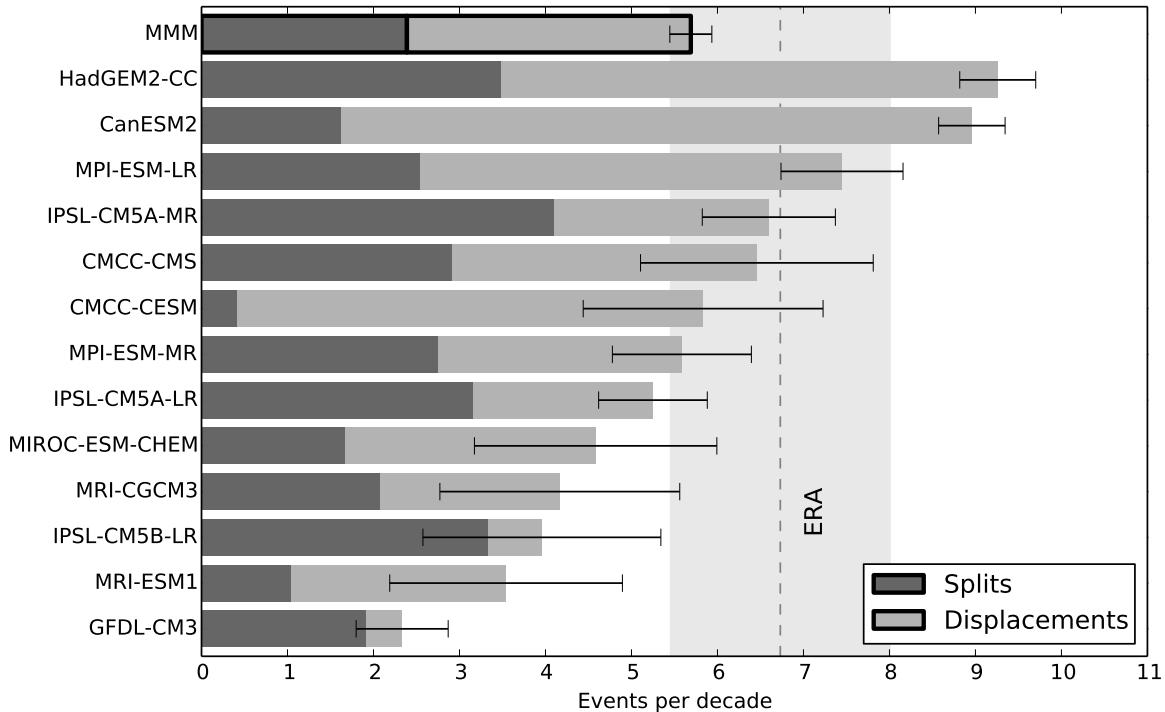
No clear consistent biases among models emerge from this analysis. CanESM2 has a modal centroid latitude which is about  $5^{\circ}$  too far equatorward compared to reanalysis. Contrastingly, GFDL-CM3 has a modal centroid latitude about  $2.5^{\circ}$  more poleward than observed. CMCC-CESM displays 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 3.2. 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 aspect ratio

#### 3.2.2 Displaced and split vortex events

Displaced and split vortex events are identified within the CMIP5 ensemble using the threshold-based method described in Section 2.2.2. The same thresholds as

used for ERA ( $66^{\circ}\text{N}$  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 3.3.



**Figure 3.3:** 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  $\sigma$  range, assuming a binomial distribution of events. The grey shaded region represents the one  $\sigma$  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  $60^{\circ}\text{N}$  and 10 hPa. They also found HadGEM2-CC to have the highest frequency of events within the CMIP5 ensemble, while MRI-CGCM3 is the 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, MRI-CGCM3 is 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 2, also holds for climate models (although this is not explicitly studied here).

### 3.2. VORTEX MEAN STATE AND VARIABILITY

As well as the large differences in the total frequency of displaced and split vortex events, it can be seen in Figure 3.3 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 3.4. 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 3.2.1, affect the frequency of split and displaced vortex 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 average, undisturbed, 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 a theoretical distribution and then finding the peak of that distribution. Following *Mitchell et al.* [2011a], we fit the aspect ratio with a generalized extreme value (GEV) distribution of the form

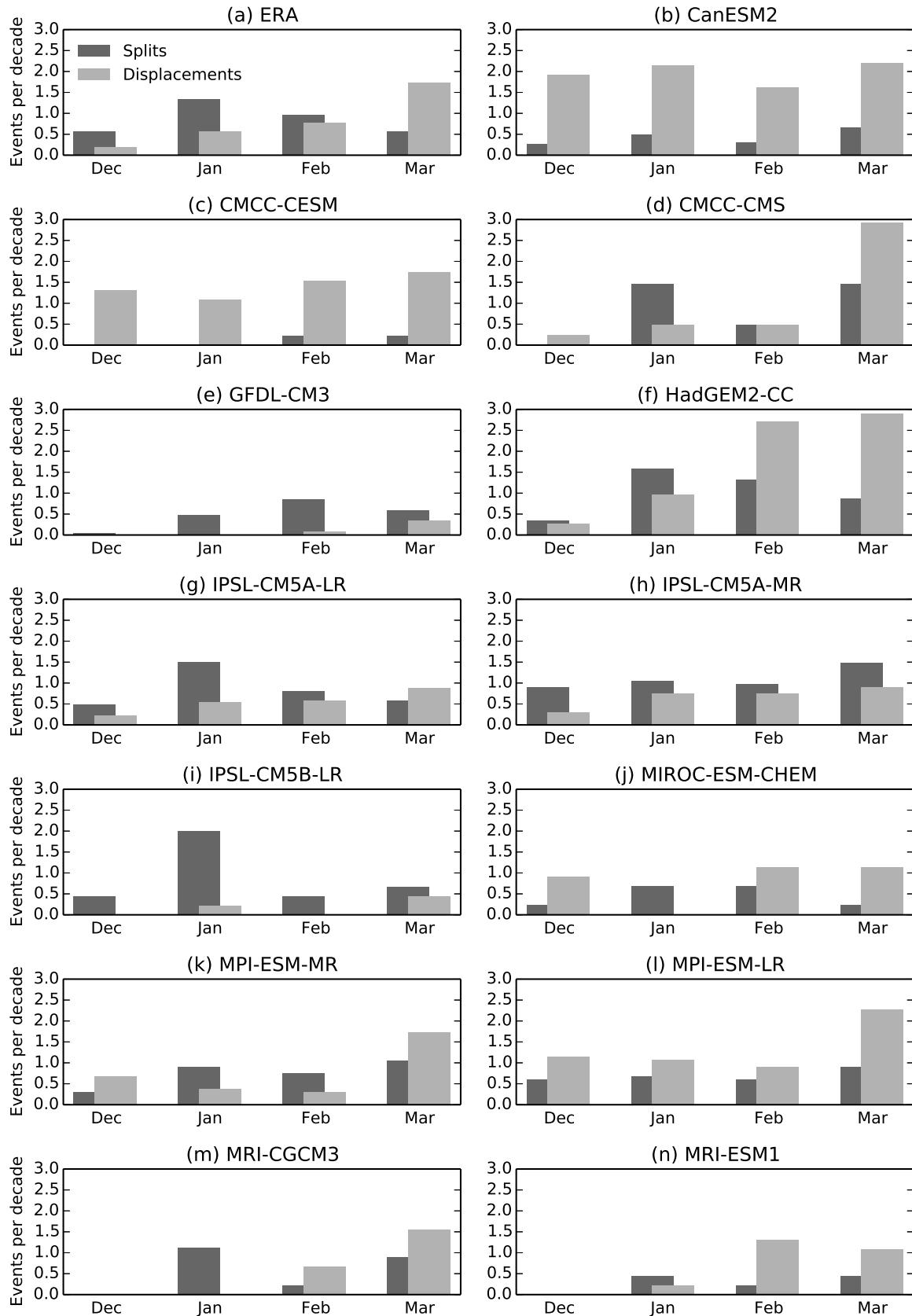
$$f(x; \mu, \sigma, \xi) = \frac{a^{(-1/\xi)-1}}{\sigma} e^{-a^{-1/\xi}} , \quad (3.1)$$

with

$$a = 1 + \xi \frac{x - \mu}{\sigma} , \quad (3.2)$$

where  $\mu$  determines the position of the peak along the  $x$ -axis,  $\sigma$  determines the variance of the distribution and  $\xi$  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

$$f(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) , \quad (3.3)$$



**Figure 3.4:** Seasonal distribution of the occurrence of split and displaced vortex events in ERA (a) and the CMIP5 models.

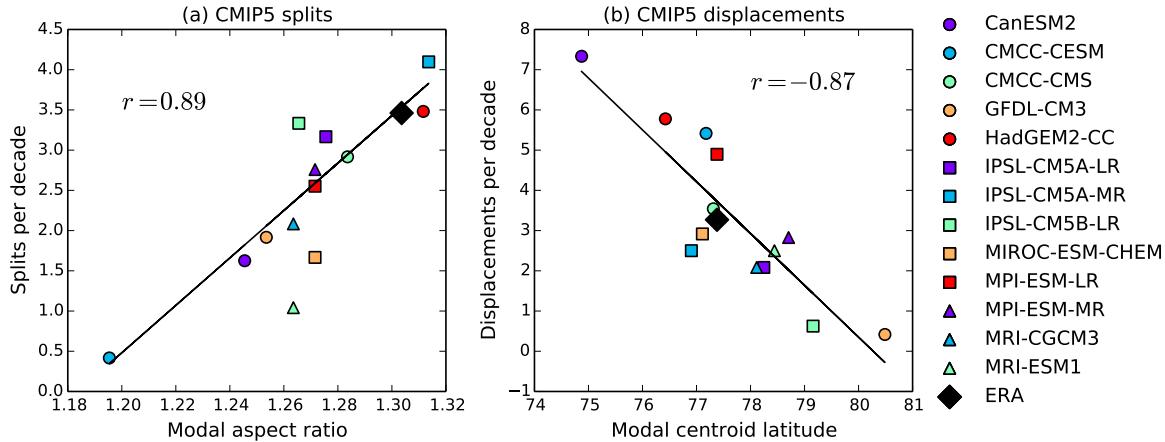
### 3.2. VORTEX MEAN STATE AND VARIABILITY

where  $\mu$  determines the position of the peak along the  $x$ -axis and  $\sigma$  is the standard deviation, to the cube of the centroid latitude, and then the cube root taken to return the original distribution. *Mitchell et al.* [2011a] 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 3.5 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, with the modal aspect ratio accounting for 79% of the spread in the frequency of split vortex events and the modal centroid latitude accounting for 76% of the spread in the frequency of displaced vortex events. This demonstrates that biases in the average undisturbed 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 accurately represent the variability of aspect ratio and centroid latitude, relative to the model climatology.

It can also be seen in Figure 3.5 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 undisturbed 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 circular vortex is a common bias among models.

The structure of the stratospheric polar vortex during split and displaced vortex events in the CMIP5 ensemble is shown in Figure 3.6. This shows composites of 10 hPa geopotential height at the onset date of the events for each model. The absolute values of the contours are different for each model due to differences in climatology, although the spacing between contours remains constant. It can be seen that the majority of models accurately reproduce splitting events 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 also has a very different appearance of displaced vortex events, although this composite only consists of three events so differences are unlikely to be statistically significant. There is also significant inter-model



**Figure 3.5:** 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.

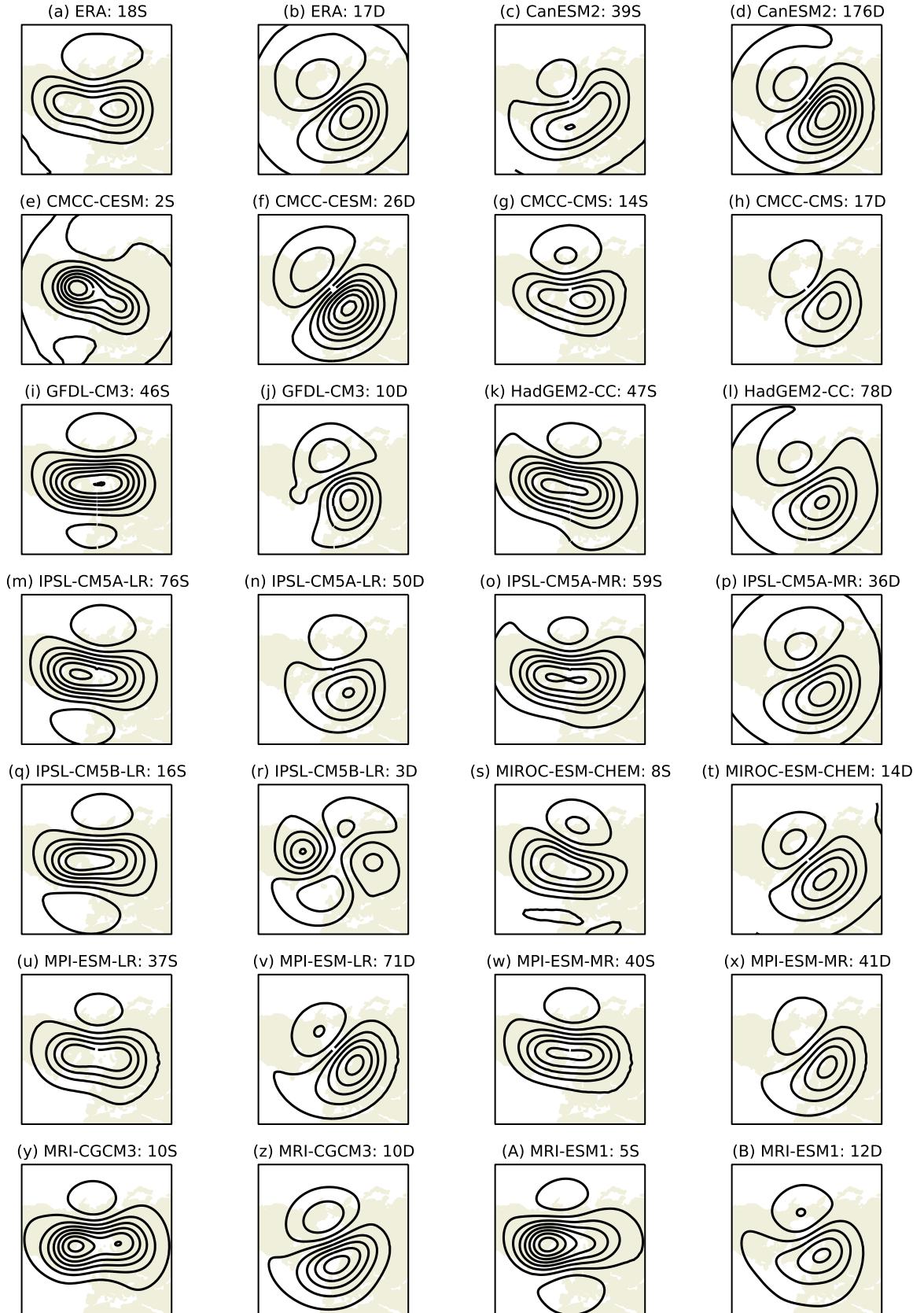
spread in the relative strengths of the Aleutian and Azores highs during split vortex events. Several models (GFDL-CM3, IPSL-CM5A-LR, IPSL-CM5B-LR, MRI-ESM1) show an approximately equal strength Aleutian and Azores highs, while others (CMCC-CMS, HadGEM2-CC, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM2) show a weaker Azores high, which is more in 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 [Mitchell *et al.*, 2013].

### 3.3 Stratosphere-troposphere coupling

#### 3.3.1 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 3.7. This shows composites of polar cap ( $60^{\circ}$ - $90^{\circ}$ 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

### 3.3. STRATOSPHERE-TROPOSPHERE COUPLING



**Figure 3.6:** 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.

height is occasionally below ground level. It should be noted that polar cap Z is highly correlated with the NAM over the levels shown in Figure 3.7. Indeed, comparing Figures 3.7 (a) and (b) with Figure 2.5 shows that the ERA composites for polar cap Z and the NAM are very similar. In Figure 3.7 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, which is more in 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 10 hPa. 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 apparent in the multi-model mean (MMM) (Figures 3.7 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

### 3.3. STRATOSPHERE-TROPOSPHERE COUPLING

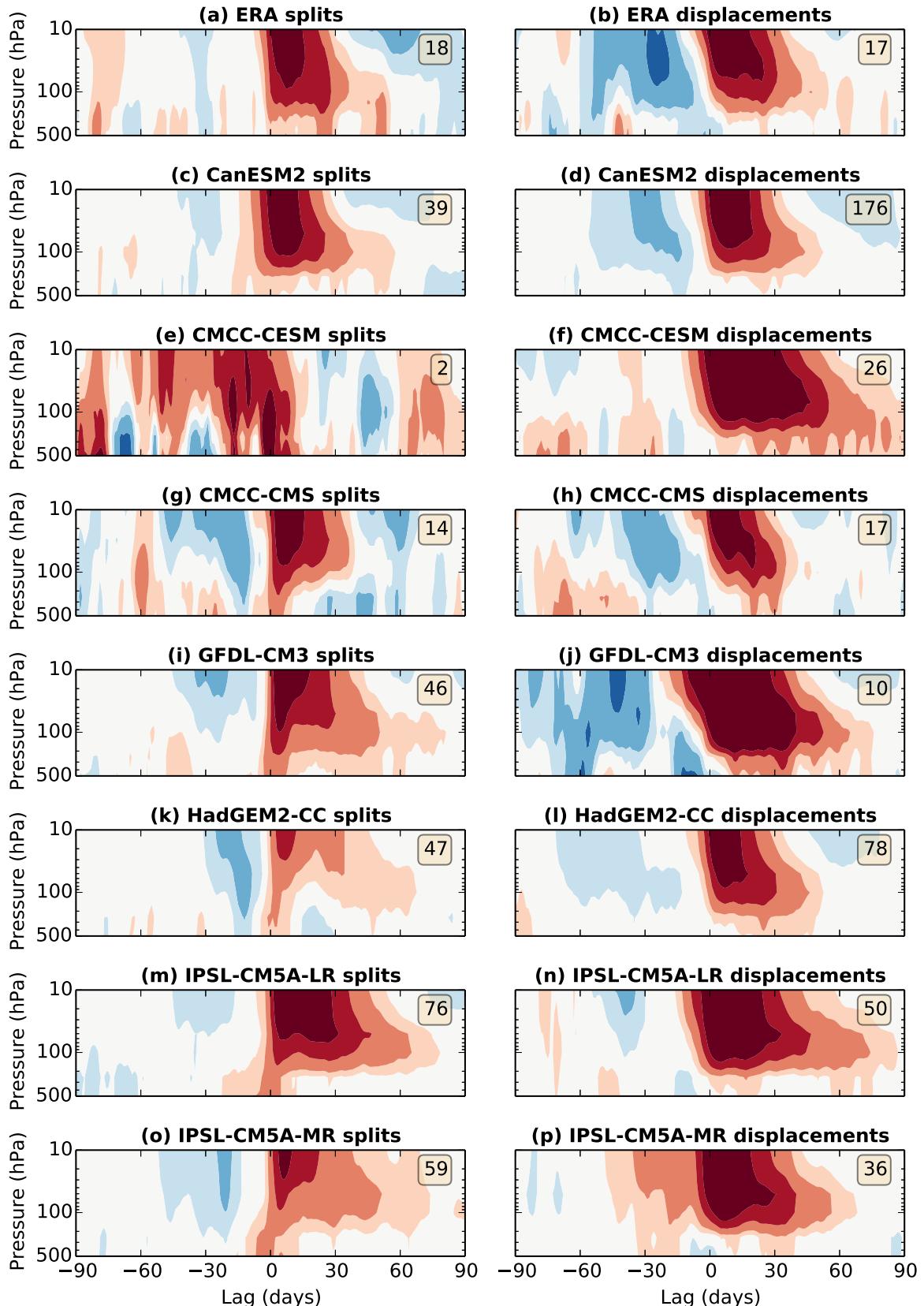
hand this does mean that greater weight is given to models with more ensemble members and more events (almost one third of all displaced vortex events come from CanESM2), but the difference in baroclinicity of split and displaced vortex events is observed to be very consistent among models.

#### 3.3.2 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 3.8. Also shown are the anomalies for ERA (a,b) (the same as Figure 2.7) 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, 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 3.7 (m,n)), but a relatively strong NAO signal (Figure 3.8 (m,n)), particularly following split vortex events. This is because 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, since the surface NAM pattern (the leading EOF of MSLP) has the same sign and similar magnitude in the North Atlantic and North Pacific [e.g., *Ambeam et al., 2001*].



**Figure 3.7:** 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.

### 3.3. STRATOSPHERE-TROPOSPHERE COUPLING

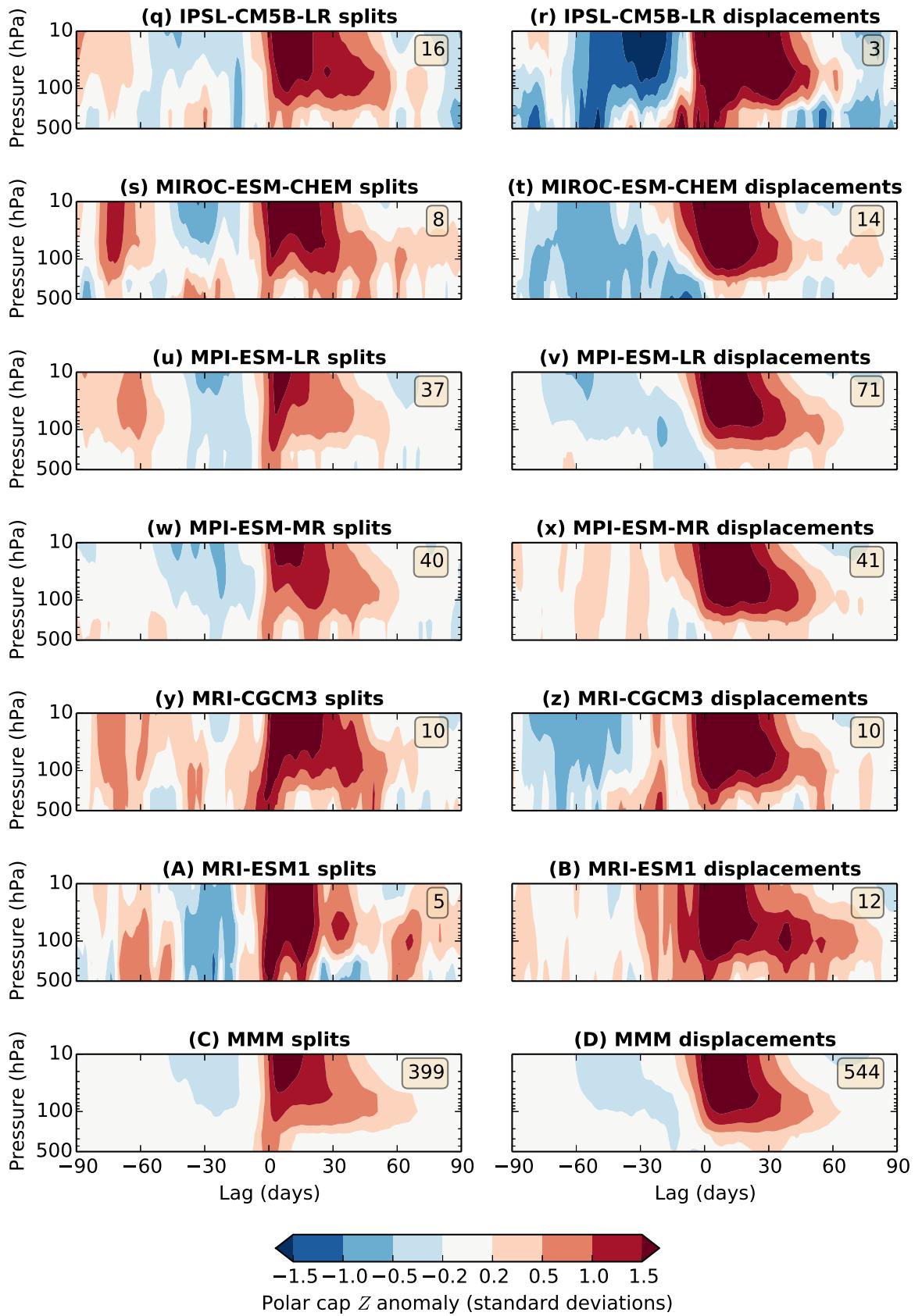
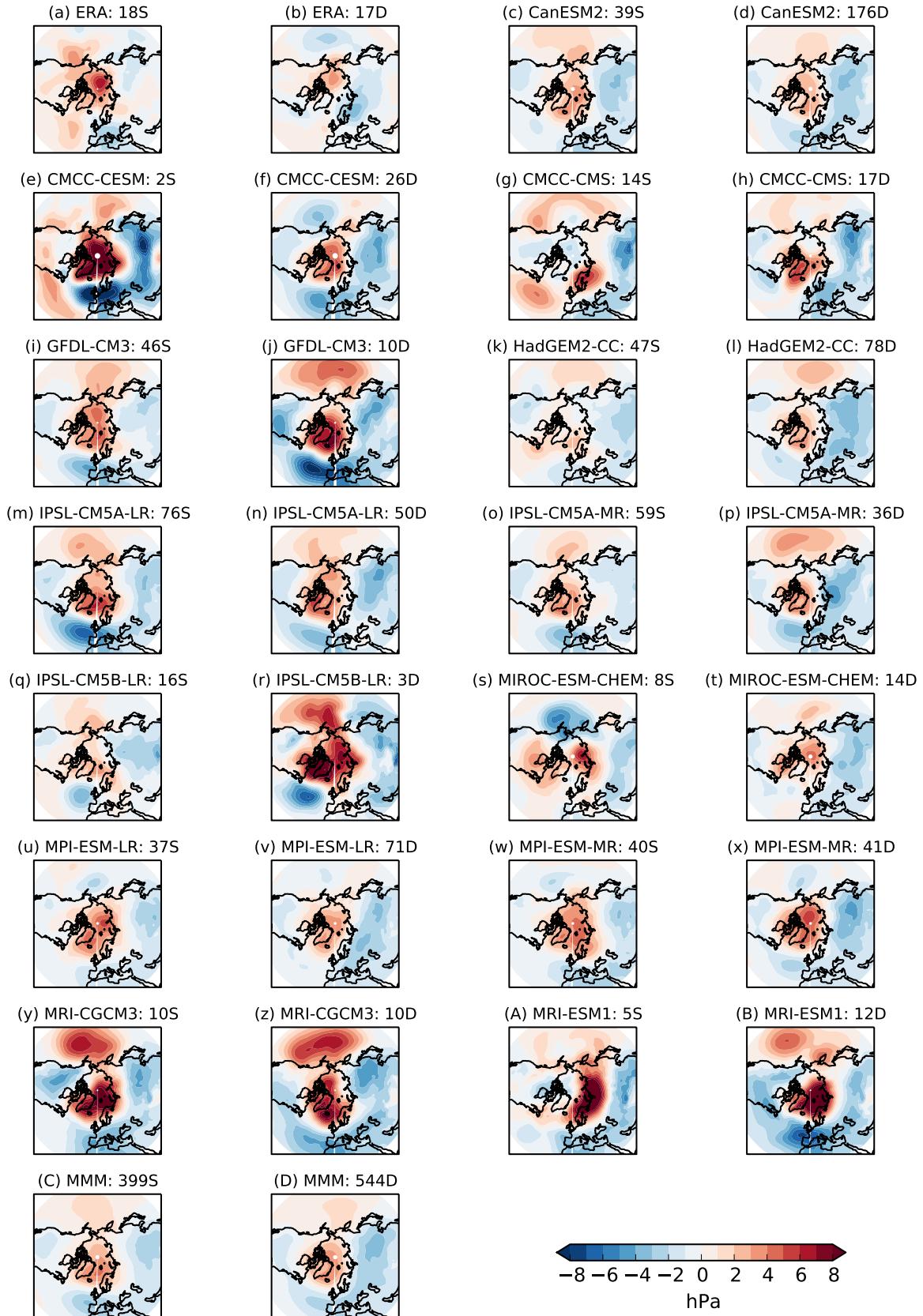


Figure 3.7: (Continued)

CHAPTER 3. REPRESENTATION OF VORTEX VARIABILITY IN CLIMATE MODELS



**Figure 3.8:** 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.

### 3.3. STRATOSPHERE-TROPOSPHERE COUPLING

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. 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 3.9 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 3.9 (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 3.6). 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 baroclinic nature of displaced vortex events compared to split vortex events.

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 minimum, which is negative in some cases, located near 90°E, which is consistent with the position of minimum in the split vortex composite. It can be seen that the multi-model mean difference (Figure 3.9 (S)) is remarkably similar to the reanalysis difference (Figure 3.9), 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.

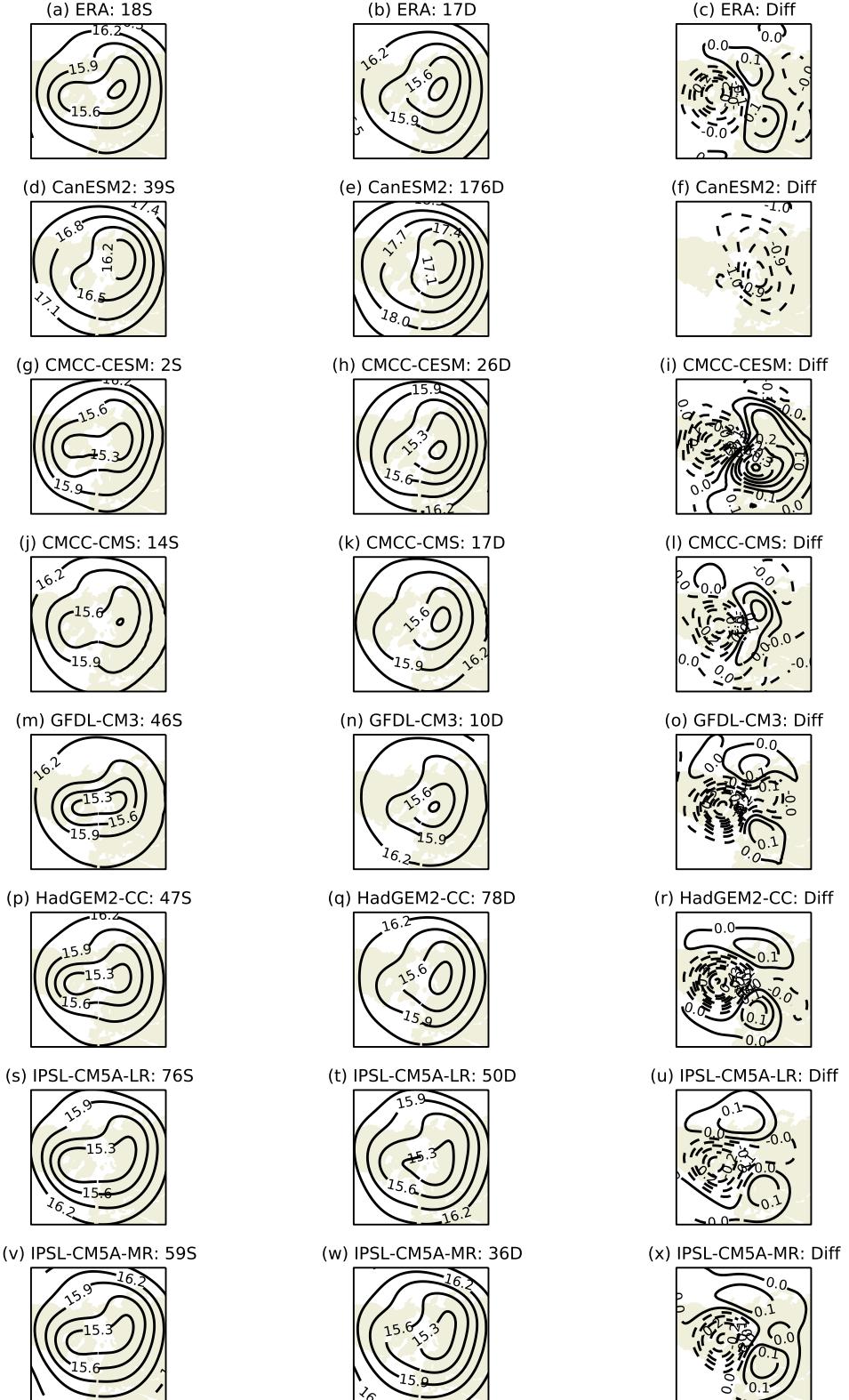
Figure 3.10 shows the split minus displaced vortex composite difference for MSLP averaged 0-30 days following onset and the 100 hPa Z composite difference averaged 0-10 days following onset, for both ERA and the CMIP5 MMM. Statistical significance in the MSLP difference is calculated by a two-tailed 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. Group all events together and select from them, with replacement, two random subsets which are equal in size to the total number of split and displaced vortex events respectively.
2. Take the difference of the averages of these two subsets.
3. Repeat the above 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.

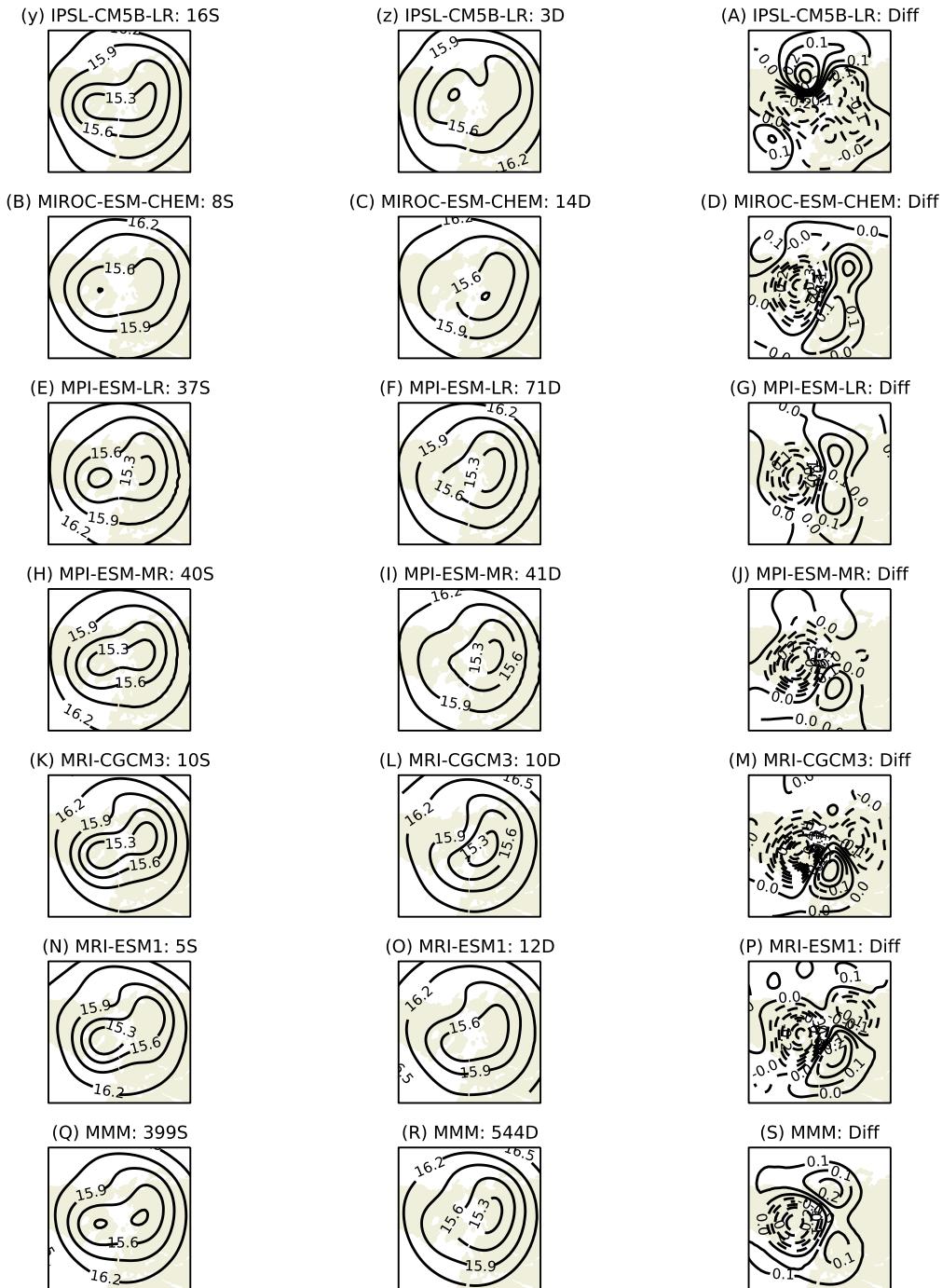
For the case of ERA, very little statistical significance in the composite difference is seen, while in the CMIP5 MMM there are large statistically significant regions. This is due to the greatly increased sample size in CMIP5; a total of 943 events compared to just 35 in ERA.

In the CMIP5 MMM difference the largest feature is the large positive anomaly (a result of a more negative anomaly following displaced vortex events) over 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 assymetric and so does not project strongly onto the polar cap average, therefore explaining the small difference in polar cap averaged Z (Figure 3.7 (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 3.8 (C,D)). For the ERA MSLP difference there are stronger

### 3.3. STRATOSPHERE-TROPOSPHERE COUPLING



**Figure 3.9:** 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 3.9:** (Continued)

### 3.4. DISCUSSION

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 2.3.

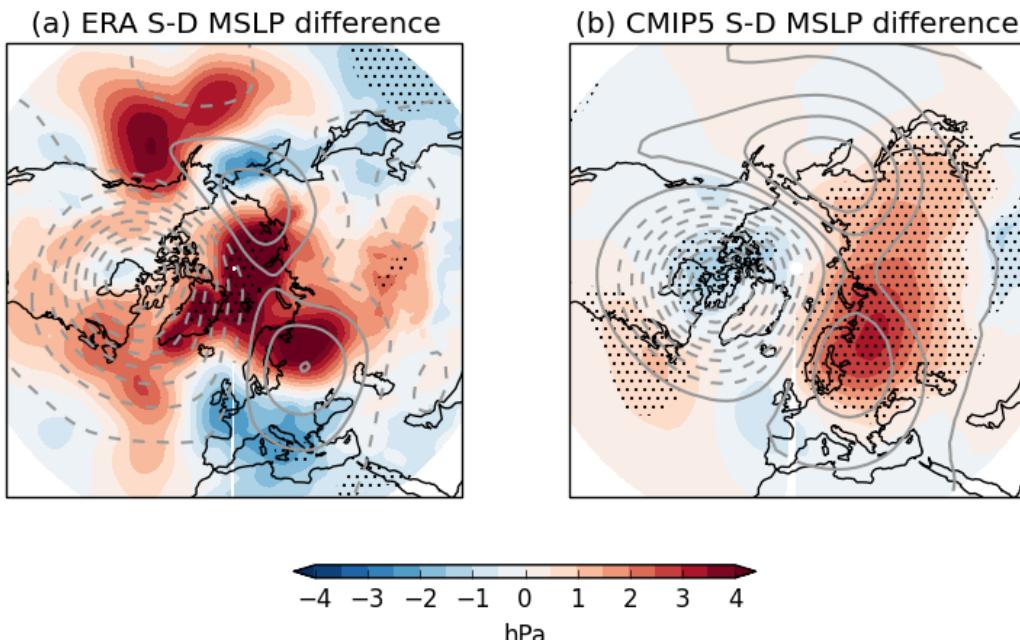
The CMIP5 difference of Figure 3.10 shows the positive 100 hPa Z anomalies over Siberia over-lie the positive MSLP anomalies, while the negative 100 hPa Z over northern Canada over-lies negative MSLP anomalies. A somewhat similar, but not statistically significant pattern is seen in ERA, although the Siberian anomaly is more polewards and the negative anomaly over Canada is much weaker. Significantly, 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 3.9 (a,b) and 2.8 (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 fact that anomalous MSLP differences are co-located with 100 hPa Z (which is in turn related to tropopause height) lends weight to the idea that the mechanism behind the *differences* in the surface response to split and displaced vortex events is a localised geostrophic and hydrostatic adjustment to stretching caused by changes in tropopause height. This is the mechanism discussed by *Ambaum and Hoskins [2002]* and in Section ???. It is important to note that this does not imply that this is the dominant mechanism in the *individual* responses to split and displaced vortex events, which share many similarities, for instance in their influence on the NAO. This result is discussed in more detail in Section 3.4.3.

## 3.4 Discussion

### 3.4.1 Effect of model resolution

We have demonstrated that there are large differences in the representation of the average undisturbed 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 than  $5^\circ$  and in the aspect ratio of 0.12, which corresponds to (?% and ?% of the ERA standard deviation respectively). Because a similar number of models have a poleward as an equatorward centroid latitude bias, the multi-model mean is approximately accurate. In contrast, only two of 14 models have higher modal aspect ratio than reanalysis.



**Figure 3.10:** 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  $\pm 20$  m.

### 3.4. DISCUSSION

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. As a corrolory, models therefore have a relatively 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 3.1 are related to models' polar vortex climatology. Figure 3.11 shows the correlations between model resolution and modal aspect ratio and centroid latitude.

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 more detail in Figure 3.12. The relatively wide scatter of points as well as the small correlation coefficient values indicate that vtical resolution fails to accound for a substantial fraction of inter-model variability in the modal aspect ratio. However, the  $p$ -values shown indicate a relatively high level of statistical significance for both correlations. It should be noted that these correlations are not highly influenced by outliers. This can be seen by the bootstrap test  $p$ -values, as well as the rank correlations which are  $-0.60$  (96.9%) and  $-0.75$  (99.7%) for vertical resolutions over 5-15 km and 15-30 km respectively.

The two measures of vertical resolution,  $dz$  (5-15 km) and  $dz$  (15-30 km), are themselves correlated (see Figure 3.11), so it is difficult to interpret which region (if any) has the largest impact on the modal aspect ratio. 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. They also found that the strongest relationship to be with upper-troposphere/lower-stratosphere (UTLS) vertical resolution, which is where we also find the strongest correlation with modal aspect ratio (although slightly lower statistical significance; see Figure 3.12). Blocking events are known to influence upwards wave propagation into the stratosphere [??], as well as stratospheric anomalies influencing the frequency of blocking events [??]. Hence, 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.

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 3.12(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. A similar effect is also seen in comparing vertical resolution and blocking frequency [Anstey *et al.*, 2013]. 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 3.12.

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 in Figure 3.11 (treating the measures of vertical resolution as a single parameter since they are highly correlated). Assuming these parameters to be independent (i.e. a binomial distribution), there is an approximately 26% chance of finding at least one 95% significant correlation among those tested if the data are sampled from an uncorrelated distribution.

On the other hand, there is 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], with planetary waves tending to propagate towards regions of high  $n_s$  and becoming evanescent in regions where  $n_s < 0$ . It is given by

$$n_s^2 = \frac{\bar{q}_\phi}{a\bar{u}} - \frac{s^2}{a^2 \cos^2 \phi} - \frac{f^2}{4N^2 H^2}, \quad (3.4)$$

where  $N^2$  is the static stability,  $H$  scale height,  $s$  wavenumber, and

$$\bar{q}_\phi = 2\Omega \cos \phi - \left[ \frac{(\bar{u} \cos \phi)_\phi}{a \cos \phi} \right]_\phi - \frac{a}{\rho_0} \left( \frac{\rho_0 f^2}{N^2} \bar{u}_z \right)_z. \quad (3.5)$$

It is therefore apparent that  $n_s$  is sensitive both to the vertical gradient in static stability and the vertical shear of zonal wind. 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

### 3.4. DISCUSSION

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.

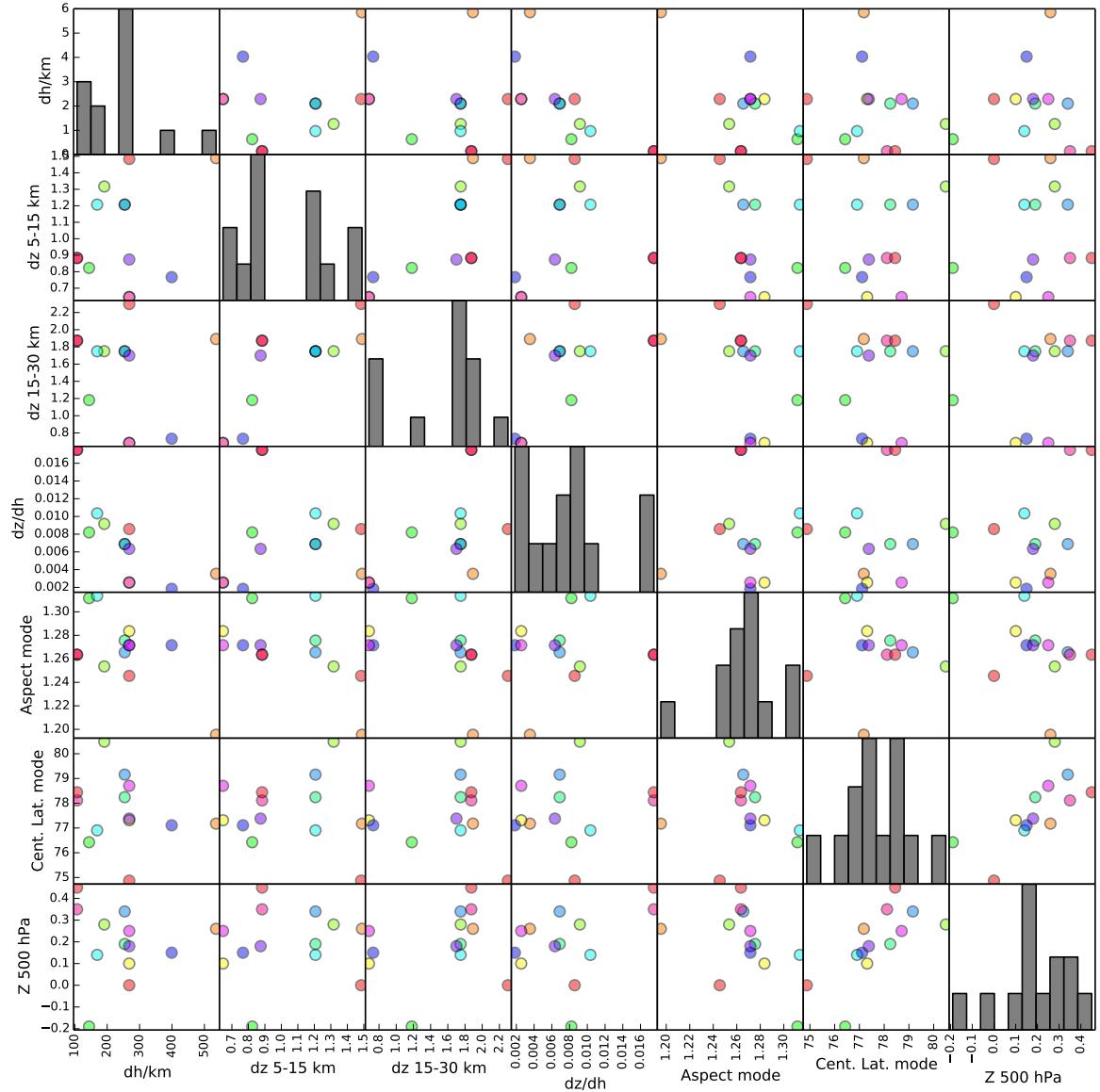
The fact that there is not a significant relationship between the modal centroid latitude and vertical resolution may suggest that wave-2 propagation is more sensitive to vertical resolution than wave-1. This is because aspect ratio is highly correlated with wave-2 activity and centroid latitude with wave-1 [*Waugh and Randel, 1999*]. 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 ??). Hence smaller differences in vertical gradients are required to exclude wave-2 propagation than wave-1 propagation.

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.

#### 3.4.2 Measures of stratosphere-troposphere coupling

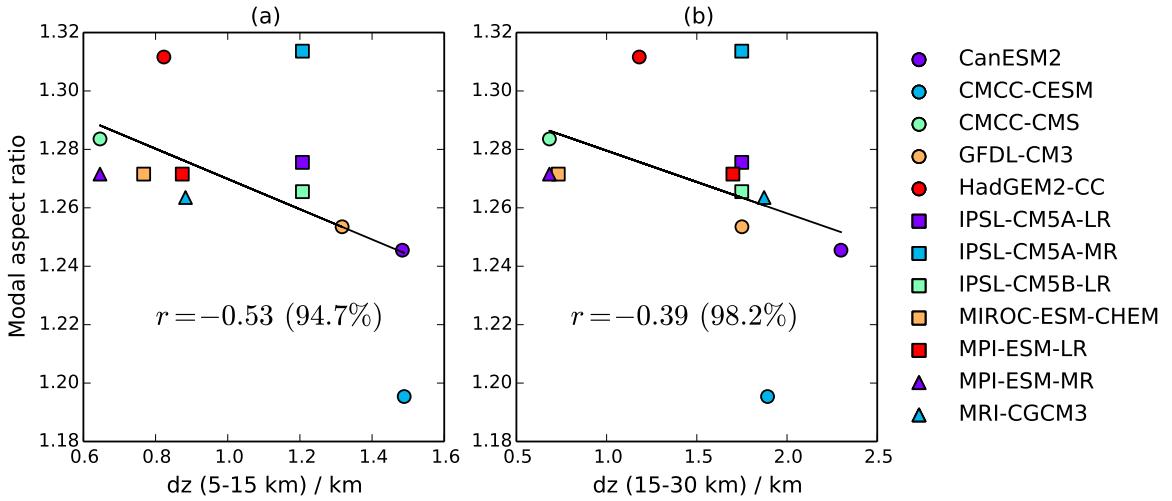
We have shown that zonally-averaged quantities such as polar cap Z (which is highly correlated with the NAM) following split and displaced vortex events are much less consistent among models than the NAO. This inconsistency is dominated by differences in the North Pacific, with some models showing positive MSLP anomalies and others negative.

Following *Baldwin and Dunkerton [2001]*, many studies of stratosphere-troposphere coupling have focussed 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 a good descriptors of inter-model variability. Therefore, we suggest that the



**Figure 3.11:** Relationship between model resolution, moment diagnostics, and stratosphere-troposphere coupling in the CMIP5 models. Model resolution is shown as horizontal resolution ( $dh$ ), vertical resolution ( $dz$ ) over 5-15 km and 15-30 km, and aspect ratio ( $dz(5-15\ km)/dh$ ). Also shown is the modal aspect ratio and centroid latitude and the mean polar cap-averaged  $Z$  at 500 hPa 0-30 days following split and displaced vortex events. Histograms for the relevant quantities are shown along the leading diagonal.

### 3.4. DISCUSSION



**Figure 3.12:** Expansion from Figure 3.11 of the 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  $r = 0$ .

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., ??]. 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 over the North Atlantic and Pacific are not significantly correlated, rather that the annular NAM pattern is a statistical artifact. ? 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, 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.

### **3.4.3 Difference between split and displaced vortex events**

## **3.5 Conclusions**

# CHAPTER 4

---

## The role of the stratosphere in seasonal prediction

---

### 4.1 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) [Pascoe *et al.*, 2006; 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 [Kodera and Kuroda, 2002; 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 [e.g., Baldwin and Dunkerton, 2001], the initial conditions of the stratosphere itself can act as a source of enhanced predictability [Baldwin *et al.*, 2003; Charlton *et al.*, 2003; Hardiman *et al.*, 2011]. The effect of the stratosphere on the troposphere is most pronounced following a rapid midwinter breakdown of the strong westerly stratospheric polar vortex (known as a stratospheric sudden warming, SSW), and past work has fo-

cused on the influence of these events on forecast skill [Kuroda, 2008; Sigmund *et al.*, 2013]. However, SSWs are highly nonlinear events which are currently not predictable beyond about two weeks in advance [Marshall and Scaife, 2010], limiting their usefulness in seasonal prediction. SSWs also occur almost exclusively in the Northern Hemisphere (NH), with only one event in the approximately 60 year record having been observed in the Southern Hemisphere (SH), in September 2002 [Roscoe *et al.*, 2005].

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, in turn, comes about because of lesser SH orography and land-sea temperature contrasts which can excite planetary waves. This reduced variability also means that anomalies in the Antarctic stratosphere persist for longer than those in the Arctic [Simpson *et al.*, 2011], so they may be predictable on longer time scales, and so 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. The SAM is the dominant mode of variability of the extratropical Southern Hemisphere and 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. 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), following the ultimate breakdown of the vortex (final warming) in late spring/early summer, this air becomes released to mid-latitudes. The extent of this summertime ozone depletion is largely determined by the total deficit in ozone over the Antarctic during spring [Bodeker *et al.*, 2005]. Interannual variability in springtime ozone depletion can therefore significantly affect the amount of

#### 4.1. INTRODUCTION

harmful ultraviolet radiation reaching the Earth's surface over more populated areas of the Southern Hemisphere.

*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 Eliassen-Palm (EP) flux (a measure of the momentum transmitted by planetary waves) entering the stratosphere 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.

The influence of planetary wave forcing on ozone depletion comes about through both chemical and dynamical mechanisms. Planetary wave breaking causes an increase of the strength of the stratospheric residual mean meridional circulation [*Haynes et al.*, 1991], with a resultant increase in large-scale descent and adiabatic warming over the pole. This warming inhibits the formation of polar stratospheric clouds which have a vital role in the activation of halogen species that cause the chemical depletion of ozone. The increased meridional circulation as well as an enhancement of horizontal two-way mixing caused by planetary wave breaking, also cause an increase in the dynamical transport of tropical ozone-rich air to the polar regions, further increasing ozone concentrations. Breaking planetary waves can also modify the geometry of the stratospheric polar vortex, stripping away elements of ozone-depleted air [*Waugh et al.*, 1994], or in the extreme case of the 2002 SSW cause the ozone hole to split in two [*Charlton et al.*, 2005].

Here, we address directly the influence of the stratosphere on springtime Antarctic seasonal forecast skill using a set of historical hindcasts (or historical re-forecasts) of a new operational system with a fully stratosphere-resolving general circulation model. We find significant skill in the prediction of the Antarctic stratospheric polar vortex up to four months in advance, and even of 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 highly 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 long-lived stratospheric circulation anomalies.

## 4.2 Seasonal forecast system

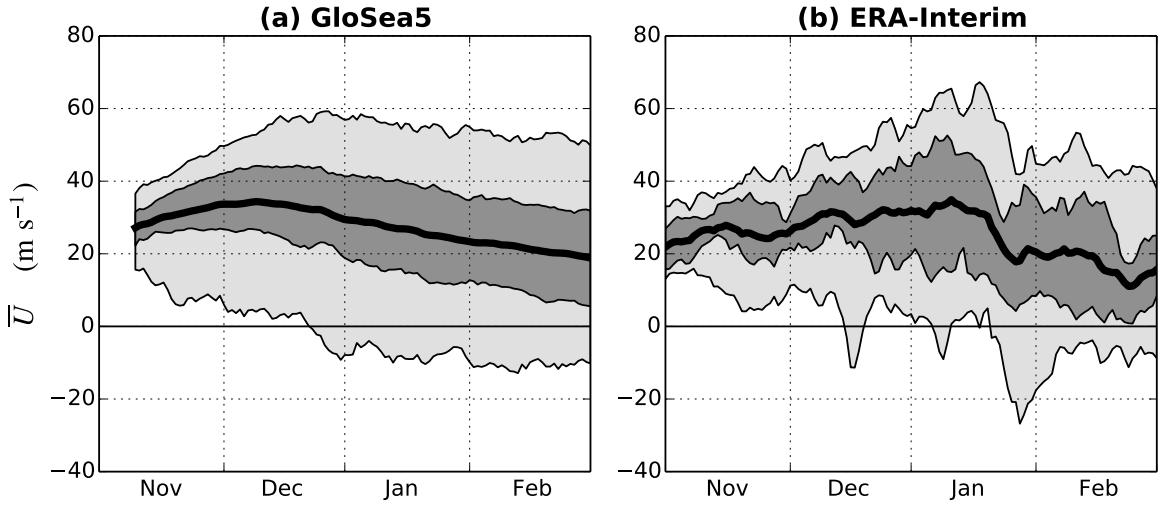
The analysis in this paper is based on results from a set of hindcast predictions produced by the Met Office Global Seasonal Forecast System 5 (GloSea5) [MacLachlan *et al.*, 2014]. This system is based upon the HadGEM3 coupled general circulation model [Hewitt *et al.*, 2011], with an atmospheric resolution of  $0.83^\circ$  longitude by  $0.56^\circ$  latitude, 85 quasi-horizontal atmospheric levels and an upper boundary at 85 km. The ocean resolution is  $0.25^\circ$  in longitude and latitude, with 75 quasi-horizontal levels. A 15-member ensemble of hindcasts was run for each year in the period 1996–2009. The hindcast length is approximately four months from three separate start dates spaced two weeks apart and centered on 1st August (07/25, 08/01, 08/09), with 5 members initialised on each start date. Members initialised on the same start date differ only by stochastic parameterization of model physics [Tennant *et al.*, 2011].

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]. 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 [Cionni *et al.*, 2011].

*Scaife et al.* [2014] have shown that this seasonal forecast system produces unprecedented skillful forecasts of the North Atlantic Oscillation during the Northern Hemisphere winter. The combined effects of ENSO, QBO and sea-ice teleconnections, as well as to the increased ocean resolution, which has improved the representation of Northern Hemisphere blocking events [*Scaife et al.*, 2011] contribute to this skill.

Hindcast accuracy is verified by comparison to the ERA-Interim reanalysis [Dee *et al.*, 2011]. This provides a ‘clean comparison’ since the hindcasts exactly match ERA-Interim at the initialisation date. 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 observa-

#### 4.3. SEASONAL FORECAST RESULTS



**Figure 4.1:** Time series of daily 10 hPa zonal-mean zonal wind ( $\bar{U}$ ) at  $60^{\circ}\text{S}$  for all GloSea5 ensemble members from 1996–2009 (a) and ERA-Interim from 1979–2010 (b). The thick black line indicates the mean, dark grey shading the interquartile range and light grey the 95th percentile range. Individual time series of the two ensemble members of GloSea5 which simulate an SSW (one for 1997 and one for 2002), and the year with an observed SSW (2002) are shown in red.

tions of ozone concentrations, and this assimilation has been demonstrated to be in close agreement with independent satellite data [*Dragani, 2011*].

## 4.3 Seasonal forecast results

### 4.3.1 Stratospheric polar vortex

The climatology of Antarctic stratospheric polar vortex winds in the GloSea5 hindcasts is compared to the ERA-Interim reanalysis climatology in Fig. 4.7. The strength of the stratospheric polar vortex is measured by the zonal-mean zonal wind ( $\bar{U}$ ) at  $60^{\circ}\text{S}$  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 realizations), 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

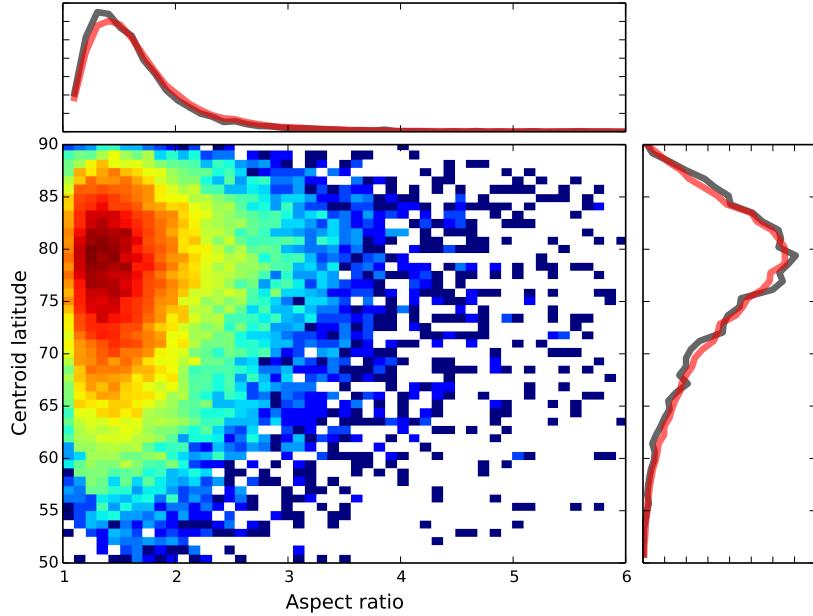


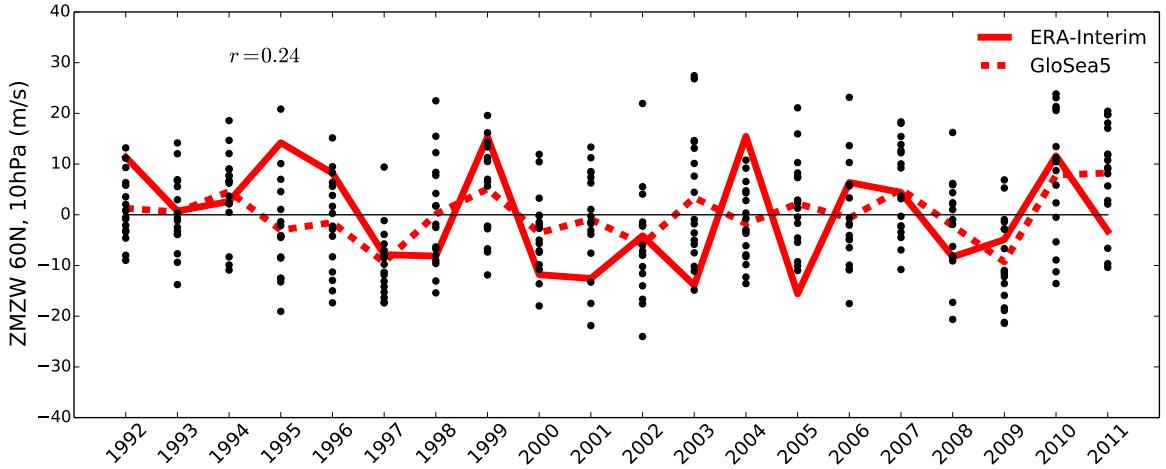
Figure 4.2

the ERA-Interim values are noisier as would be expected from a sample size consisting of fewer years.

The GloSea5 hindcast skill for the prediction of the Antarctic stratospheric polar vortex winds is shown in Fig. 4.8(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 fact it consists of only a single realization). Results are shown for September–November (SON) averages, corresponding to a 1–4 month lead time. The correlation between the GloSea5 ensemble mean and ERA-Interim is 0.73, which is statistically significant at the 99% confidence level. This correlation does not depend strongly on particular years; the correlation remains significant at the 95% level ( $r = 0.57$ ) if the year 2002 is excluded. Significance is calculated using a two-tailed bootstrap test, whereby the percentile of the observed correlation is calculated from a distribution formed by the correlation 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.

Two SSW events were simulated in the GloSea5 hindcasts; in 1997 and 2002. Time series of stratospheric polar vortex winds for these events are shown in Fig.

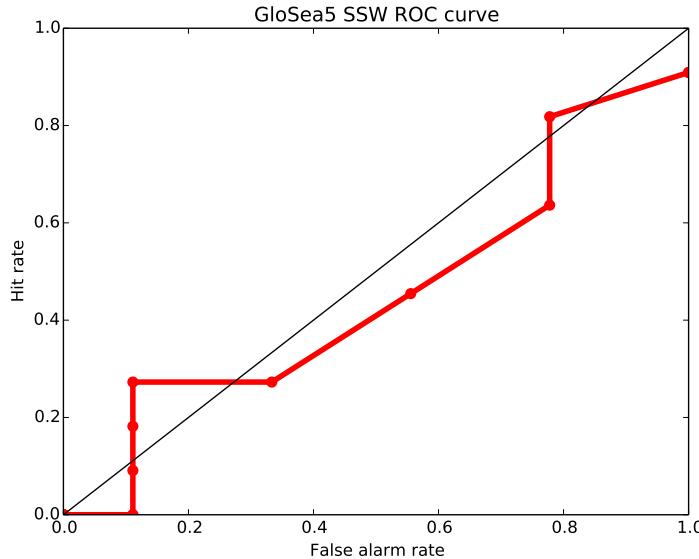
#### 4.3. SEASONAL FORECAST RESULTS



**Figure 4.3**

4.7(a) along with the observed 2002 SSW in Fig. 4.7(b). Note that although  $\bar{U}$  at  $60^{\circ}\text{S}$  and 10 hPa does not quite become easterly for the 1997 event, it does become easterly *poleward* of  $60^{\circ}\text{S}$ , which satisfies the World Meteorological Organization definition of a SSW. Given the total of 210 ensemble hindcasts, these two simulated events suggest a frequency of Southern Hemisphere SSW events of approximately one in 100 years in the current climate (making the assumption that the model can accurately simulate the probability of these events). It can also be seen that 2002 has the most anomalous stratospheric polar vortex in the GloSea5 hindcasts, with 14 of 15 ensemble members simulating negative anomalies, and the most negative ensemble mean. It is therefore possible that the 2002 event was to some degree predictable 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.* [2005], or the result of external forcing.

It should be noted that 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 Fig. 4.9, which shows geopotential height in the mid-stratosphere at the central date (date of minimum at  $\bar{U}$   $60^{\circ}\text{S}$  and 10 hPa) of the two simulated events in GloSea5 and the observed event in ERA-Interim. The distinction between splitting and displacement SSW events is important because it has been observed that tropospheric anomalies are greater following vortex splitting events, at least in the Northern Hemisphere [*Nakagawa and Yamazaki*, 2006; *Mitchell et al.*, 2013].

**Figure 4.4**

The timing of the final warming of the stratospheric polar vortex has a significant effect on stratospheric temperature and ozone concentrations [Yamazaki, 1987], as well as 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. It is likely that shorter lead-time forecasts would be required to produce skillful predictions of the final warming date.

### 4.3.2 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]. This diagnostic is not likely to be 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

#### 4.3. SEASONAL FORECAST RESULTS

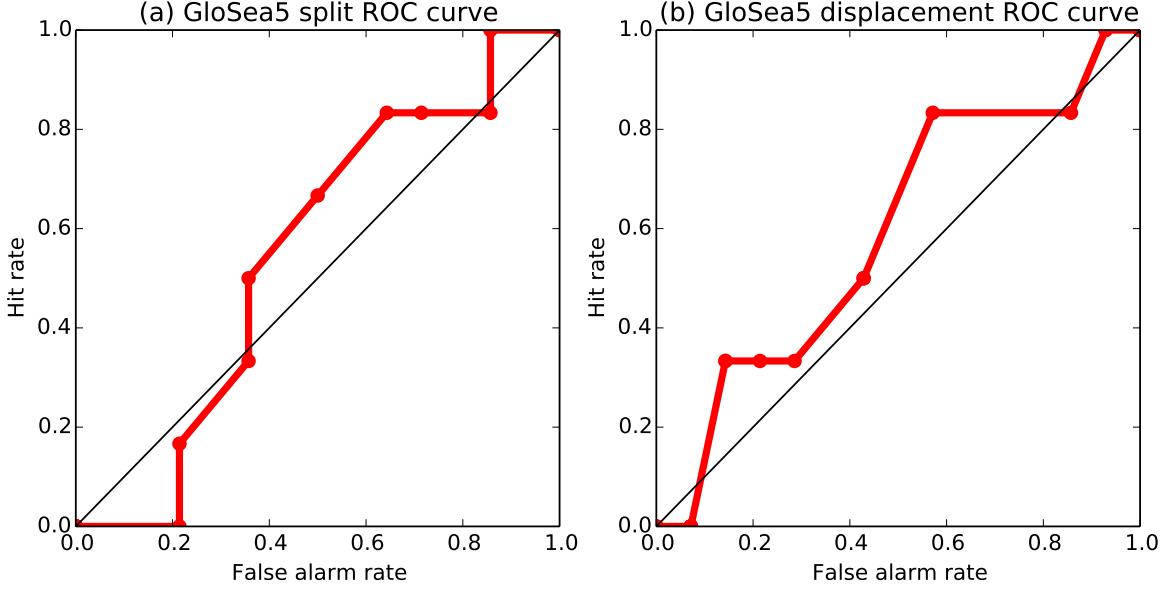


Figure 4.5

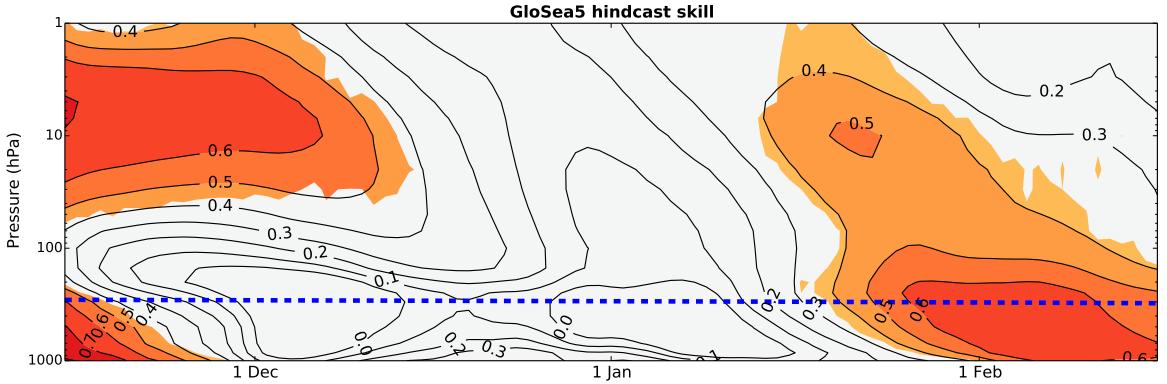
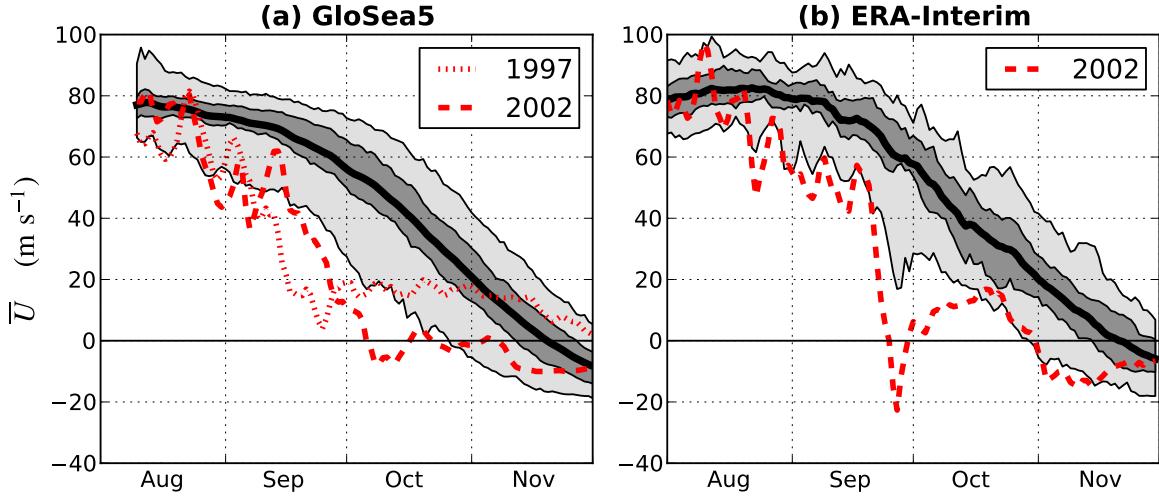


Figure 4.6

quantities.

SON mean total column ozone quantities area-weighted averaged over the polar cap ( $60\text{--}90^\circ\text{S}$ ) are shown in Fig. 4.10(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 Fig. 4.10(b)



**Figure 4.7:** Time series of daily 10 hPa zonal-mean zonal wind ( $\bar{U}$ ) at  $60^\circ\text{S}$  for all GloSea5 ensemble members from 1996–2009 (a) and ERA-Interim from 1979–2010 (b). The thick black line indicates the mean, dark grey shading the interquartile range and light grey the 95th percentile range. Individual time series of the two ensemble members of GloSea5 which simulate an SSW (one for 1997 and one for 2002), and the year with an observed SSW (2002) are shown in red.

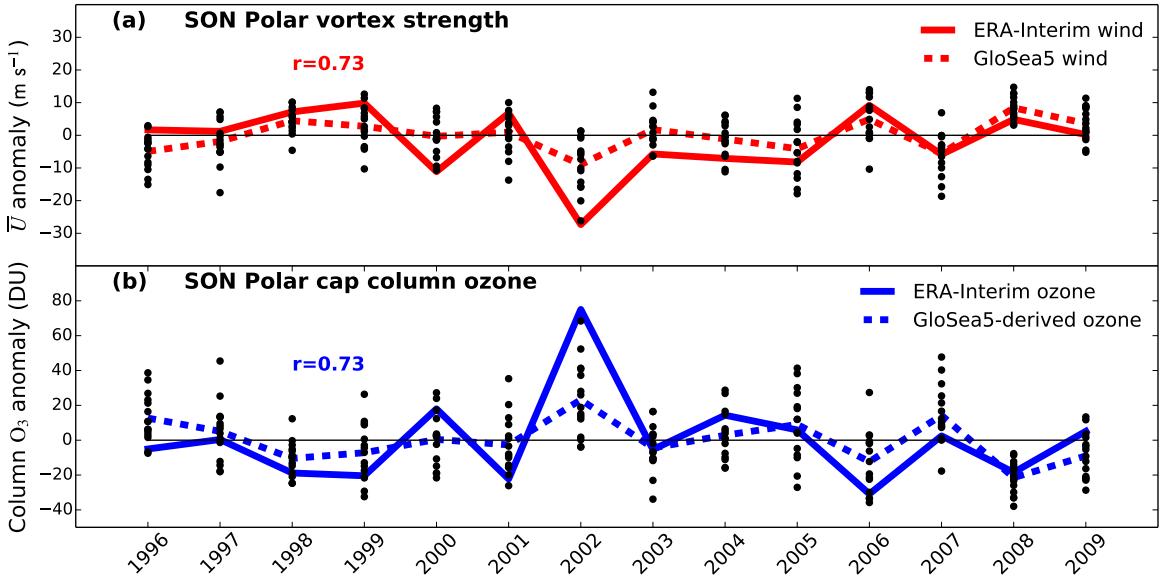
anomalies of polar cap total column ozone from the long-term trend are plotted against anomalies of the SON mean  $\bar{U}$  at  $60^\circ\text{S}$  and 10 hPa. It can be seen that these two quantities are highly correlated ( $r = -0.92$ ), meaning polar vortex variability explains approximately 85% of the variance of polar cap total column ozone anomalies.

This observed correlation can be used with the GloSea5 forecasts of polar vortex winds to produce inferred predictions of polar cap total column ozone quantities. Figure 4.8(b) shows the GloSea5 hindcasts along with the assimilated values from ERA-Interim. The correlation between the GloSea5 ensemble mean and ERA-Interim is 0.72, which is statistically significant at the 99% level. Errors from the regression in Fig. 4.8(b) for the inferred ozone quantities for each ensemble member are small compared to the spread between ensemble members, and so not plotted in this figure.

### 4.3.3 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  $40^\circ\text{S}$  and  $65^\circ\text{S}$  [Gong and Wang, 1999]. These anomalies are calculated from

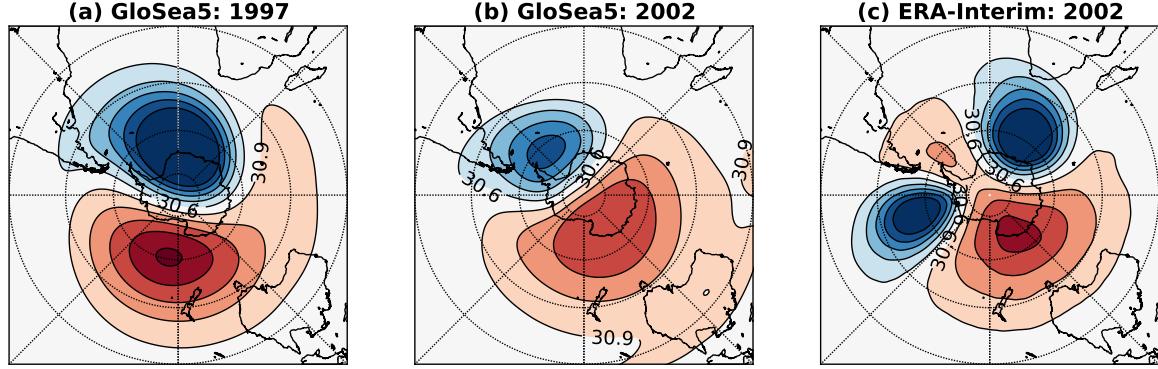
#### 4.3. SEASONAL FORECAST RESULTS



**Figure 4.8:** (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 initialised near 1st August.

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 Fig. 4.11. The correlation of the GloSea5 ensemble mean and ERA-Interim is 0.64, which is statistically significant at the 99% level, confirming skillful prediction of the SAM at 1–4 month lead times. This is similar to the value for the NAO correlation skill of 0.62 found by [Scaife et al. \[2014\]](#) in the same seasonal forecast system.

Figure 4.12(a) shows the correlation of ERA-Interim and GloSea5 SON averaged mean sea-level pressure anomalies at each grid point in the Southern Hemisphere. As would be expected from the low frequency variability of ENSO, correlations are greatest over the tropical Pacific. However, the correlations are also as high as 0.7 across southern Australia and parts of Antarctica. On the other hand, correlations over southern Africa and South America are not found to be significant. It is perhaps unsurprising that there is little skill over the Andes region, since this is significantly above sea-level, so mean sea-level pressure is not well defined.



**Figure 4.9:** Geopotential height at 10 hPa at the central date (date at which  $\bar{U}$  at  $60^{\circ}\text{S}$ , 10 hPa is at its minimum value) of 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.

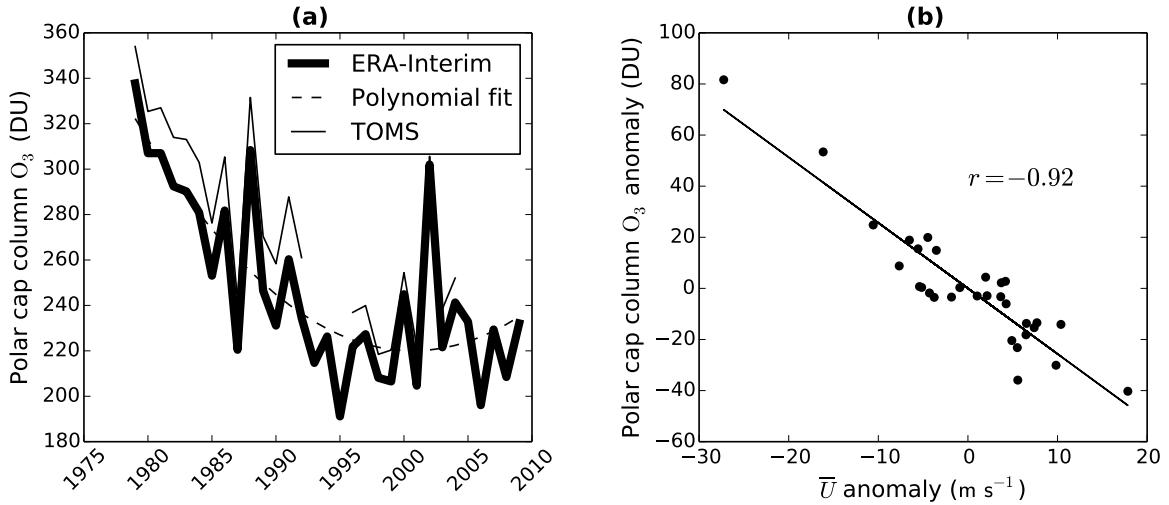
Correlations of GloSea5 SON average near-surface temperature with the gridded station-based data set HadCRUT4 [Morice *et al.*, 2012] are shown in Fig. 4.12(b). This dataset is chosen because of the scarcity of temperature observations in the Southern Hemisphere, which introduces significant biases into reanalysis data. Again, the highest correlations are found near the tropical Pacific, but significant correlations of about 0.5 are found across eastern Australia, New Zealand and Antarctica. There are also significant correlations in southern Africa and South America. The extratropical regions where the greatest forecast skill is found are similar to those which are observed to be most affected by variations in the SAM [Gillett *et al.*, 2006].

### 4.3.4 Stratosphere-troposphere coupling

Given that statistically significant skill in hindcasts of the stratospheric polar vortex is found at the same time of year as skill in predictions of the SAM, the question arises as to whether skill in one may affect the other. In order to investigate this, forecast skill as a function of lead-time and height is studied for polar cap ( $60\text{--}90^{\circ}\text{S}$ ) mean geopotential height anomalies ( $Z'$ )<sup>1</sup>. Figure 4.13(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

<sup>1</sup>At 1000 hPa monthly mean  $Z'$  is highly correlated ( $r = 0.98$ ) with the SAM index

#### 4.3. SEASONAL FORECAST RESULTS

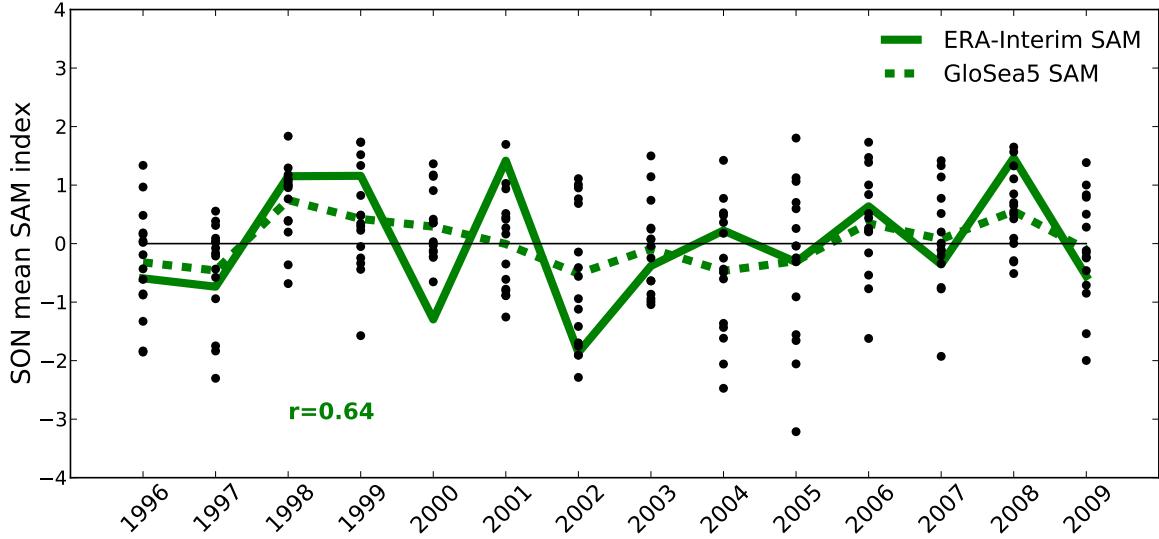


**Figure 4.10:** (a) Time series of SON mean polar cap averaged (60–90°S) total column ozone in ERA-Interim and the TOMS satellite instrument. The ERA-Interim data are fitted with a 2nd-order polynomial. (b) Anomalies of ERA-Interim column ozone from the polynomial fit plotted against SON mean anomalies at 10 hPa and 60°S for each year from 1979–2009.

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).

As would be expected from the initialisation 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 2–3 month lead time). This observed greater stratospheric than tropospheric skill might be expected from the longer ‘memory’ of the stratosphere; SAM decorrelation timescales are about 60–70 days in the stratosphere but only about 10 days in the troposphere during SON [Simpson *et al.*, 2011].

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 tropospheric circulation. An obvious candidate for such a signal is the polar stratosphere, since this remains predictable through-



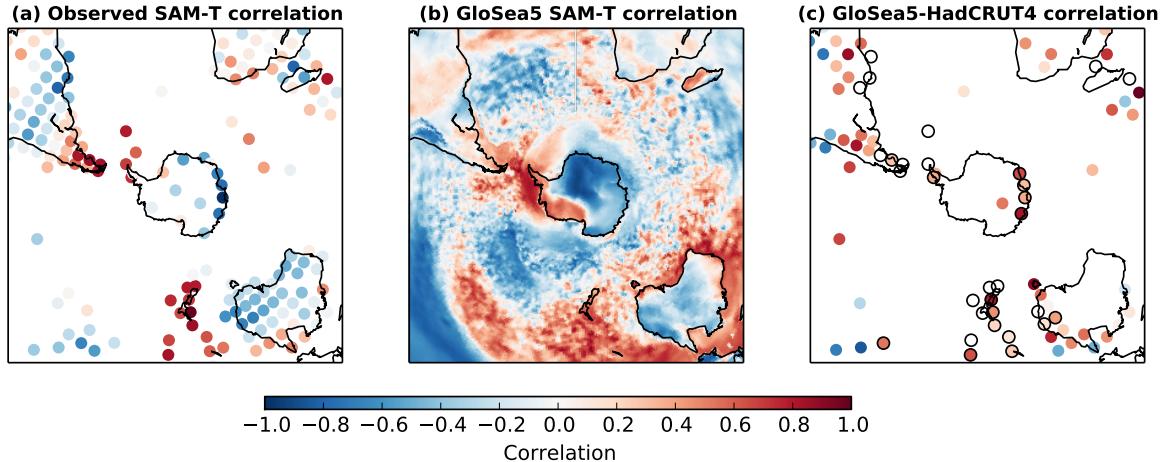
**Figure 4.11:** SON mean Southern Annular Mode (SAM) index in individual GloSea5 hindcast ensemble members (dots), ensemble mean (dashed green curve) and ERA-Interim (solid green curve). The SAM is calculated from mean sea-level pressure data, and hindcasts initialised near 1st August. The correlation of the ensemble mean and ERA-Interim values is 0.64, which is statistically significant at the 99% level.

out 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. ERA-Interim values are used to produce this model based on the linear regression of  $Z'$  at 10 hPa on 1st August with  $Z'$  at all other times and heights for 31 of the 32 years from 1979–2009. This model is then used to produce a hindcast of the 32nd year based on its  $Z'$  at 10 hPa on 1st August. The method ensures that no information from the hindcast year enters the model. The process is then repeated to make hindcasts of all 32 years; a procedure known as leave-one-out cross validation [*Wilks, 2006*].

Figure 4.13(b) shows the average correlation of 30-day running means of these statistical forecasts 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 skilful 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 sug-

#### 4.4. DISCUSSION



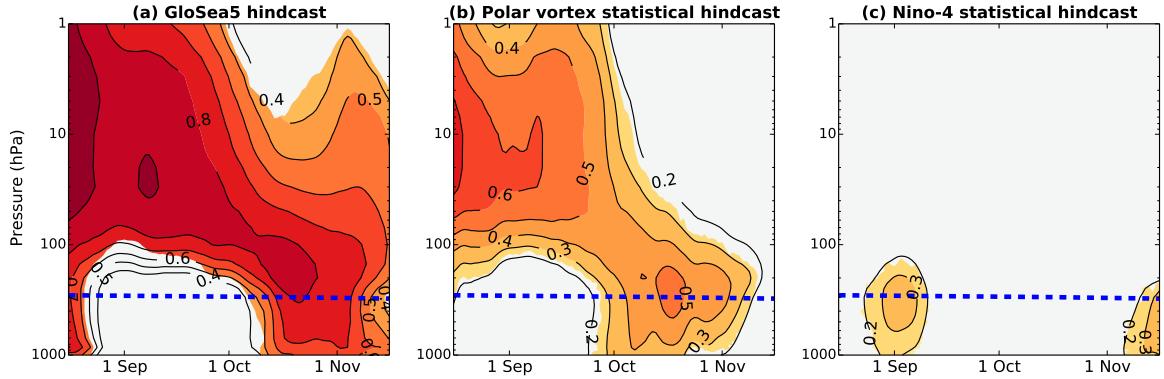
**Figure 4.12:** (a) Correlation of SON mean sea-level pressure in the GloSea5 ensemble mean with ERA-Interim values. (b) Correlation of SON mean near-surface air temperature in the GloSea5 ensemble mean and HadCRUT4 gridded station-based temperature data. Regions with no observations are white and grey shading indicates regions where the correlation is not greater than zero at the 95% confidence level, using a bootstrap test at each gridpoint.

gests that the re-emergence of tropospheric skill in the GloSea5 hindcasts in October is likely to be caused by persistent stratospheric anomalies which descend with time.

The statistical hindcasts in Fig. 4.13(b) show lower skill than the GloSea5 hindcasts at all times, and do not show statistically significant tropospheric correlations, nor the increase in upper-stratospheric skill during November. These observations could potentially be explained by the importance of non-linearities or the influence of external factors, such as ENSO, on the Antarctic stratosphere-troposphere system. Indeed, statistical hindcasts similar to those shown in Fig. 4.13(b) were produced based on the Niño-3 index, and found to have statistically significant tropospheric skill during November, but none at other times or heights. This is consistent with the results of [Lim et al. \[2013\]](#), who find the greatest correlation between tropical Pacific sea-surface temperatures and the SAM during November–January.

## 4.4 Discussion

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



**Figure 4.13:** (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. 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].

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 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 rela-

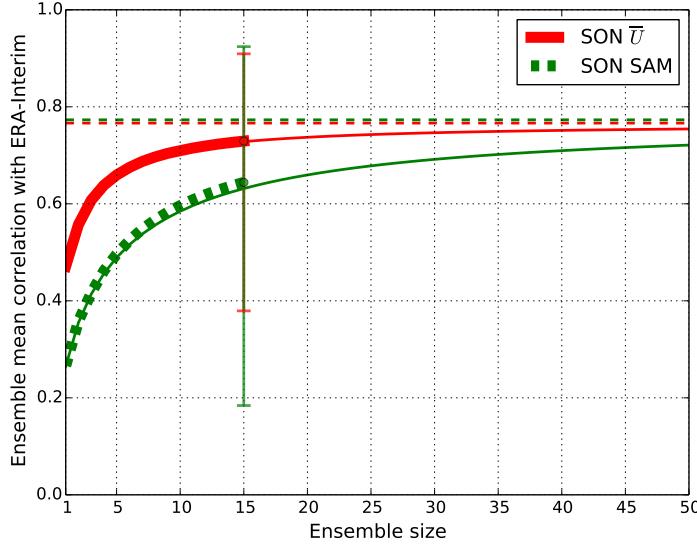
#### 4.4. DISCUSSION

tionship 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].

The correlation skill of 0.64 for the SON mean SAM in the GloSea5 hindcasts is greater than that of other contemporary seasonal forecast systems at similar lead times. For instance, *Lim et al.* [2013] report a correlation of 0.3–0.4 for the SON mean SAM from 1st August initialised forecasts using the Predictive Ocean and Atmosphere Model for Australia, version 2 (POAMA2). Significantly, this system has only two model levels in the stratosphere, and so is unable to simulate the stratosphere-troposphere coupling described here. *Lim et al.* [2013] suggest that the significant SAM predictability found from October–January in their system is the result of the influence of ENSO through a tropospheric teleconnection. These findings are not inconsistent with our results, since this time period is beyond the extent of the GloSea5 hindcasts, and largely after the stratospheric final warming, when the stratosphere is much less variable. *Lim et al.* [2013] were also mostly concerned with longer range forecasts (up to 6-month lead time) which are beyond the persistence time scales of the Antarctic stratosphere, but within those of the tropical Pacific.

Despite this significant correlation skill in hindcasts of the SAM, it is clear from 4.11 that the amplitude of the ensemble mean hindcast 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 and correlation are directly related [Kumar, 2009], so that the expected correlation would be just 0.3. The fact that it is greater than this is because the average correlation between ensemble members and observations is much greater than that between pairs of ensemble members. A similar but smaller difference is also found for the stratospheric polar vortex forecasts. These results mean that individual ensemble members have a smaller predictable signal than observations.

Given this result, it might be expected that more skillful 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 members for each year, following the method of *Scaife et al.* [2014]. This is repeated many times 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 Fig. 4.14. These curves closely



**Figure 4.14:** 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)

follow the theoretical relationship of *Murphy* [1990], which relies only on the mean correlation between pairs of ensemble members,  $\langle r_{mm} \rangle$ , and the mean correlation between individual ensemble members and observations,  $\langle r_{mo} \rangle$ , given by

$$r = \frac{\langle r_{mo} \rangle \sqrt{n}}{\sqrt{1 + (n - 1)\langle r_{mm} \rangle}} \quad (4.1)$$

where  $r$  is the ensemble mean correlation, and  $n$  is the ensemble size. These curves are shown in Fig. 4.14, along with their asymptote for an infinite sized ensemble. Although the stratospheric forecasts cannot be greatly improved with a larger ensemble size in the current system, correlation scores of about 0.7 of the SAM could be achieved with an ensemble size near 30. Both have an asymptote near 0.8, similar to that found by *Scaife et al.* [2014] for the NAO.

## 4.5 Conclusions

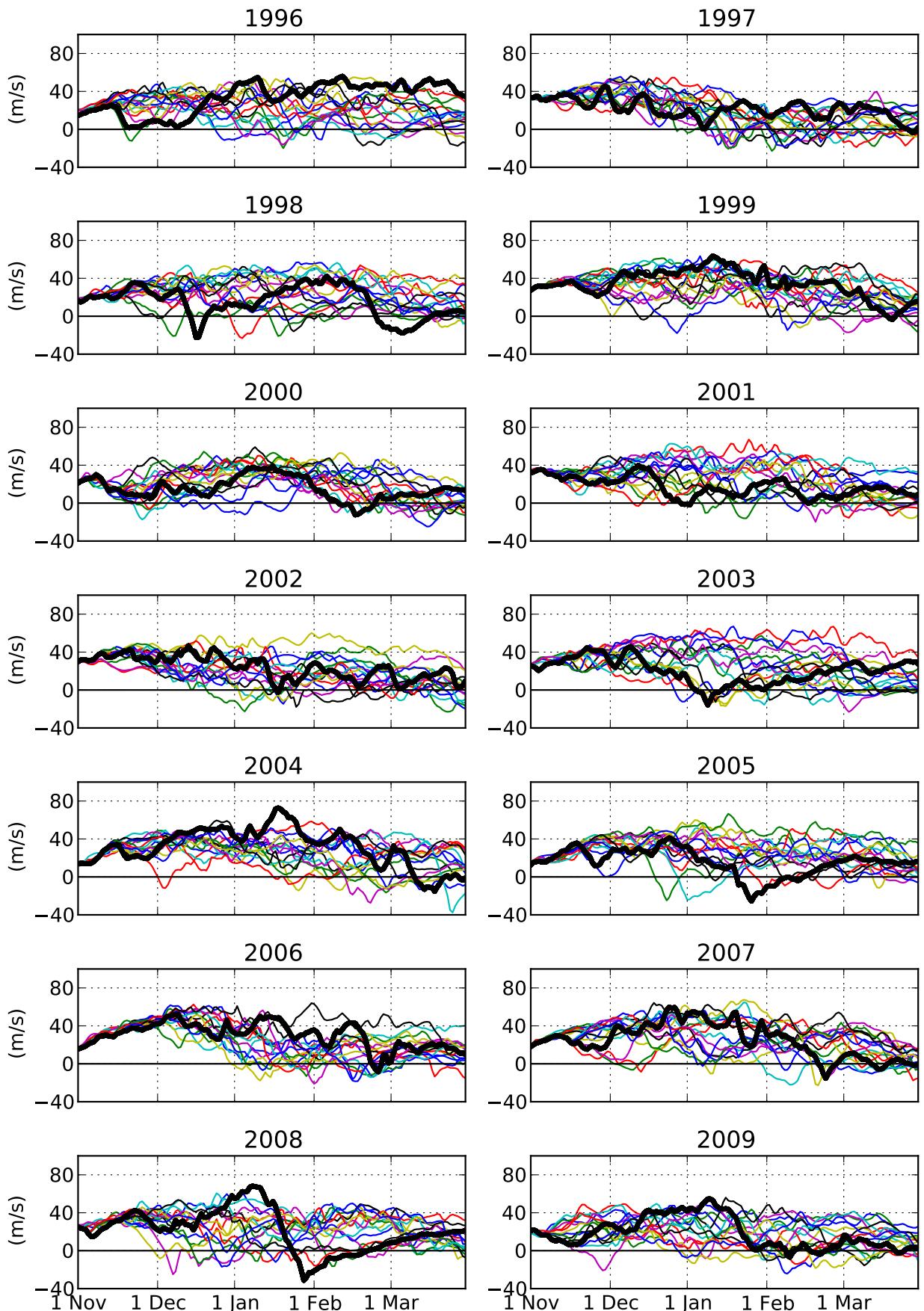
Using a set of seasonal hindcasts initialised at the start of the austral spring, we have demonstrated skillful prediction of the interannual variability of the Antarctic stratospheric polar vortex at 1-4 month lead times. This includes extreme events such as the 2002 SSW, which is the most extreme year in the ensemble

#### 4.5. CONCLUSIONS

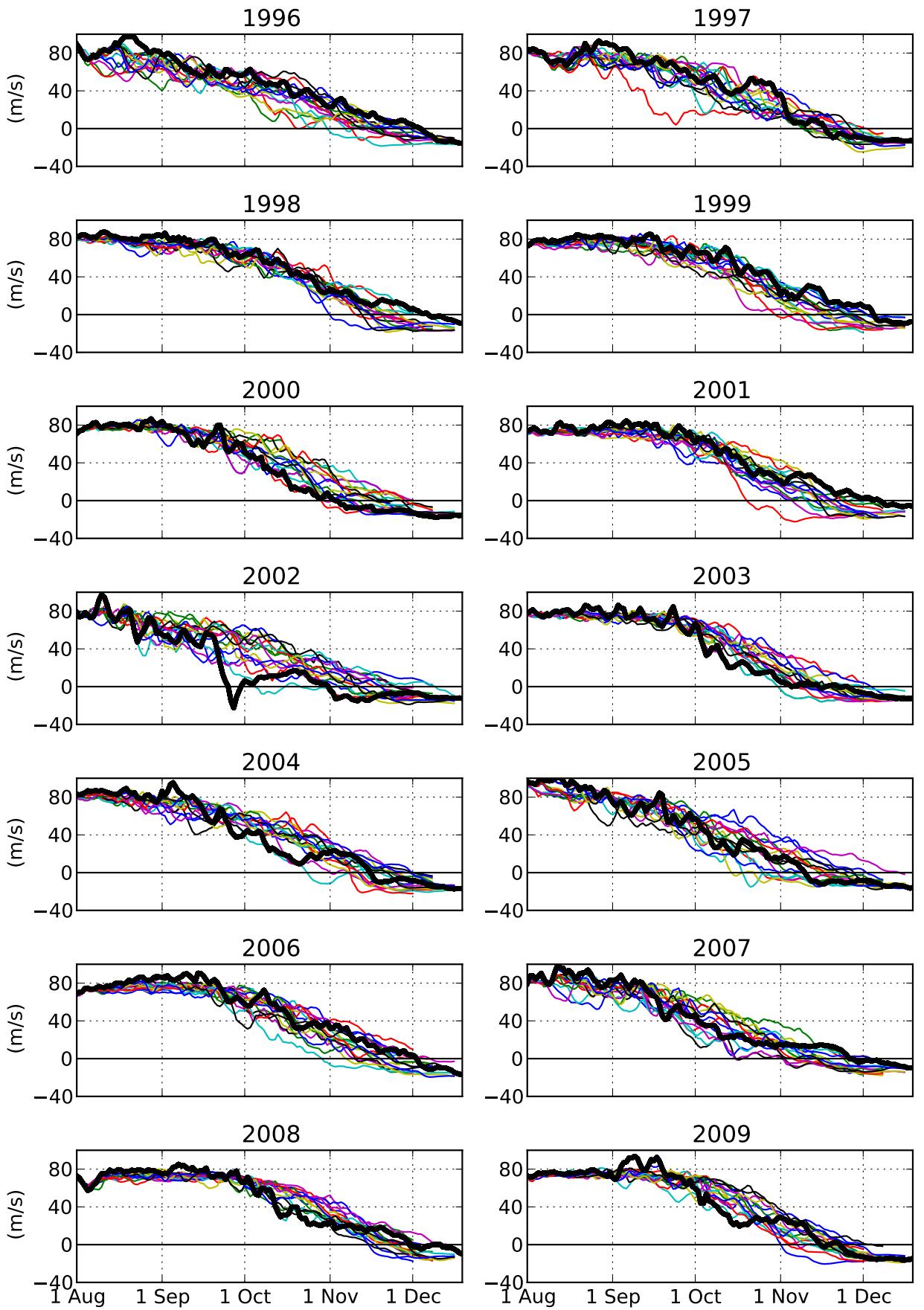
mean, and has one ensemble member which simulates a SSW. Because this variability is observed to be closely correlated with Antarctic column ozone amounts, we are able to infer skillful prediction of interannual variability in Antarctic ozone depletion.

We also find significant skill, which exceeds that of other contemporary seasonal forecast systems, 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.

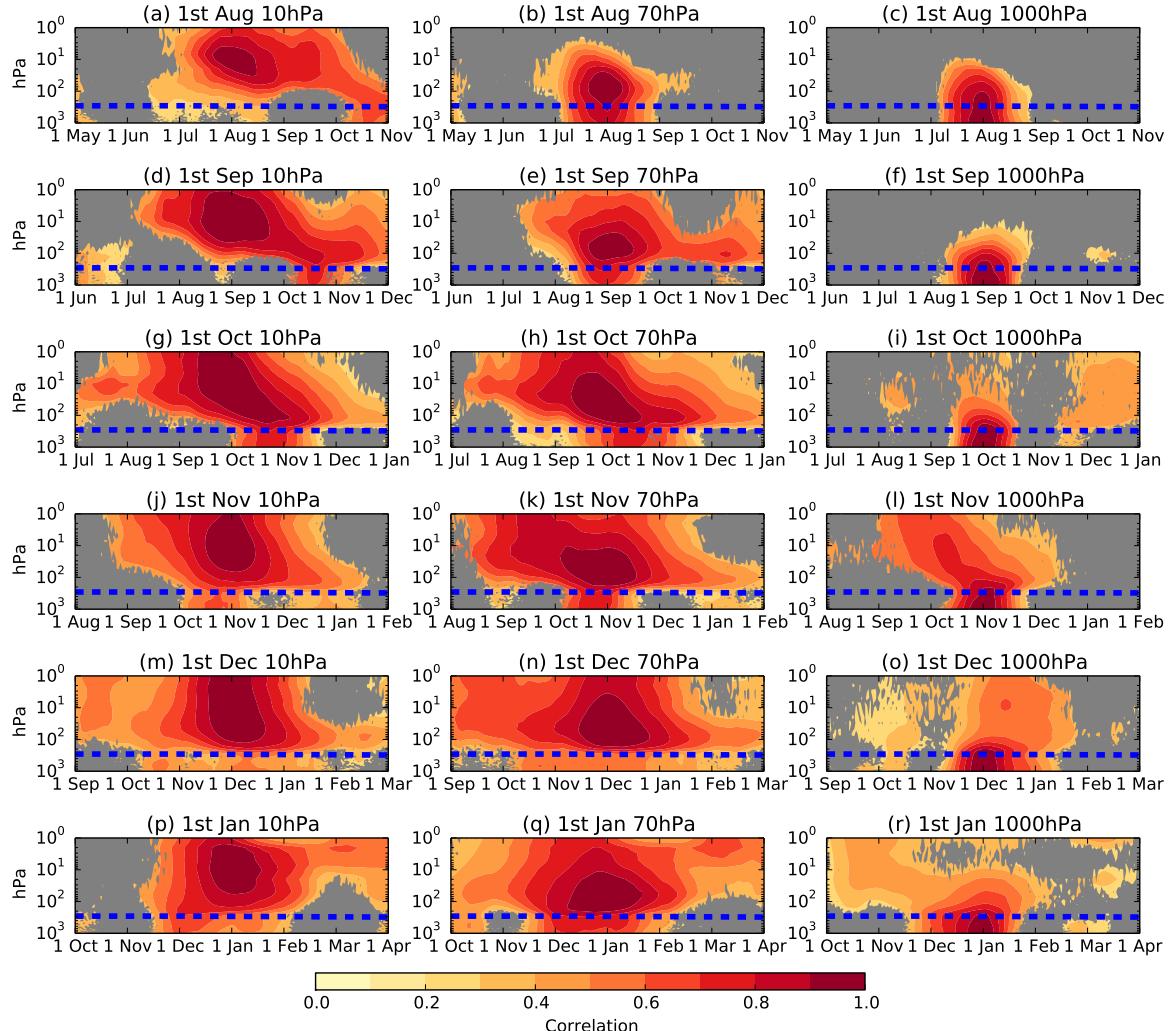
Assuming that the 14 year period studied here is representative of future years, these results suggest that it may now be possible to make skillful seasonal forecasts of interannual variations in ozone depletion and large scale weather patterns across the Southern Hemisphere. They also demonstrate the importance of the inclusion of a well-resolved stratosphere in seasonal forecast models.



**Figure 4.15:** Timeseries of zonal-mean zonal wind in the Arctic polar vortex ( $60^{\circ}\text{N}$ , 10 hPa) in the ERA-Interim reanalysis (thick black lines) and the GloSea5 ensemble hindcasts (thin coloured lines). Individual ensemble members are initialised from dates centred on November 1st.

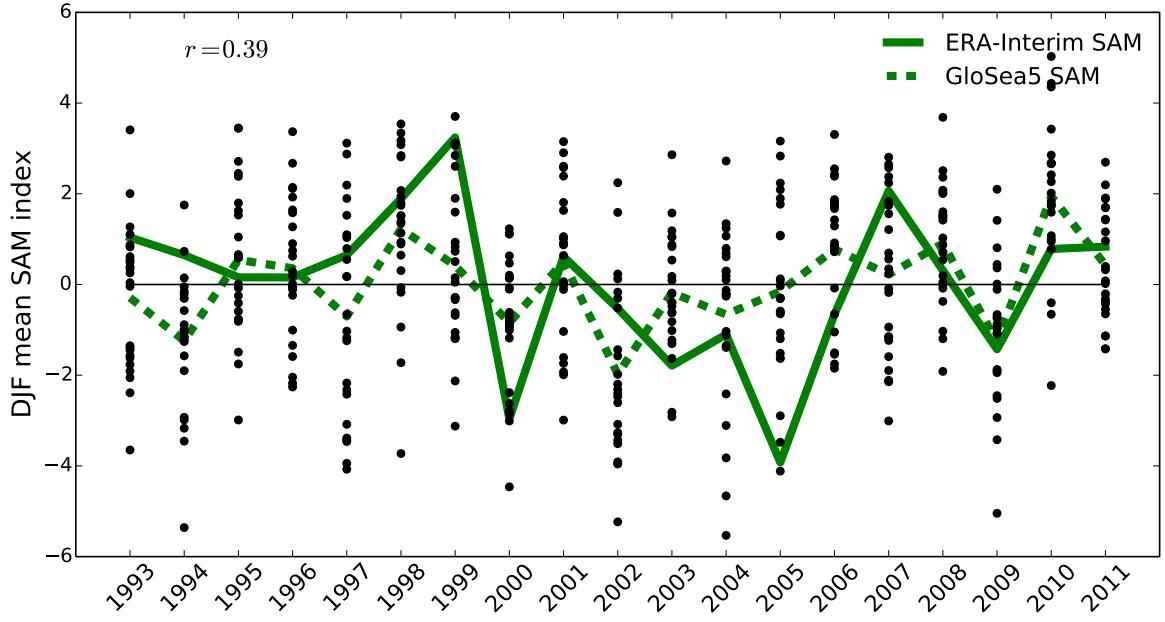


**Figure 4.16:** As figure 4.15 but for the zonal-mean zonal wind in the Antarctic polar vortex ( $60^{\circ}\text{S}$ , 10 hPa), and ensemble members initialised from dates centred on August 1st.

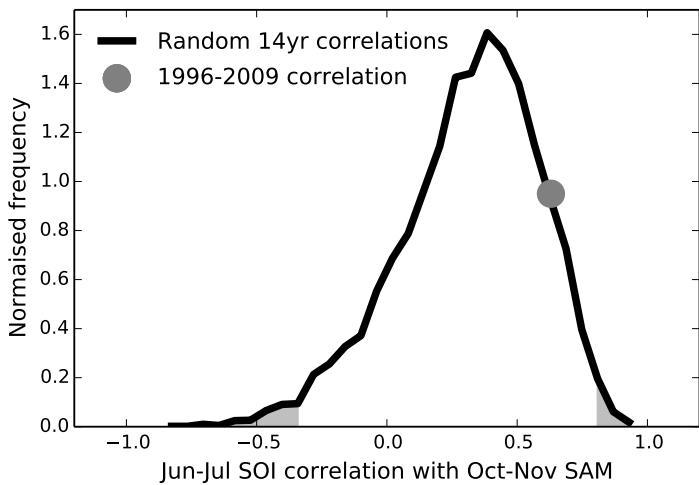


**Figure 4.17:** Correlation of ERA-Interim (1979–2010)  $Z'$  at 10 hPa, 70 hPa, and 1000 hPa with  $Z'$  at other times and lags. Values are smoothed with a 30-day running mean before correlations are calculated, and colours represent correlations that are 95% significant from zero according to a bootstrap test.

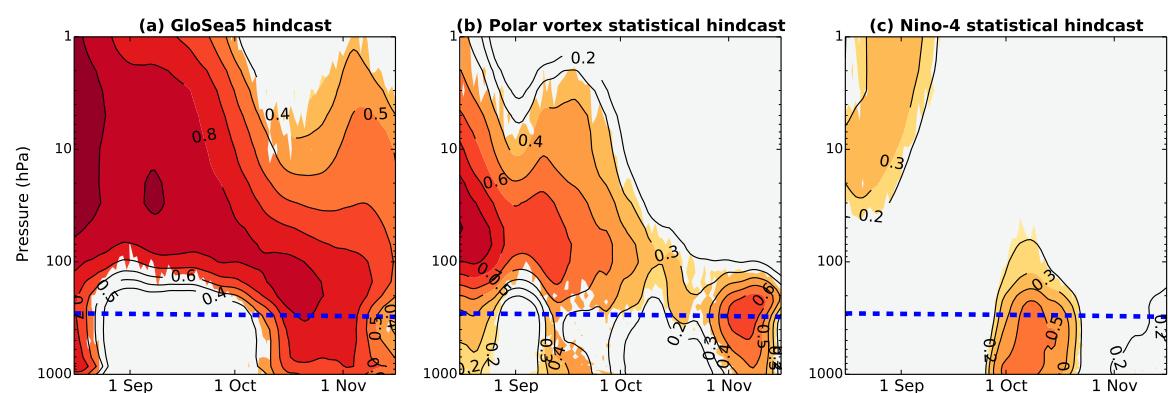
#### 4.5. CONCLUSIONS



**Figure 4.18:** 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



**Figure 4.19:** 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 4.20:** As Figure 4.13 but all analysis restricted to the 1996–2009 period.

---

## Bibliography

---

- 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., 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. Atmos.*, 118(10), 3956–3971, doi:10.1002/jgrd.50231.
- Baldwin, M., and T. Dunkerton (1999), Propagation of the Arctic Oscillation from the stratosphere to the troposphere, *J. Geophys. Res.* ....
- 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., 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.

## BIBLIOGRAPHY

- 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. 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.
- 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.
- 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(5), 1223–1238, doi:10.1175/2008JCLI2549.1.
- Charlton, A., and L. Polvani (2007), A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks, *J. Climate*, 20, 449–469.
- 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

## BIBLIOGRAPHY

- variability in low-top versions of the CMIP5 models, *J. Geophys. Res. Atmos.*, 118(6), 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(1), 83–109.
- Chen, P., and W. Robinson (1992), Propagation of planetary waves between the troposphere and stratosphere, *J. Atmos. Sci.*
- 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.
- Cordero, E., and P. 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., and L. J. Gray (2009), A continuum of sudden stratospheric warmings, *J. Atmos. Sci.*, 66, 531–540.
- 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.
- 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. Kållberg, 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.
- 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.

## BIBLIOGRAPHY

- Esler, J., and R. Scott (2005), Excitation of transient Rossby waves on the stratospheric polar vortex and the barotropic sudden warming, *J. Atmos. Sci.*, 62, 3661–3682.
- 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., A. Butler, N. Calvo, A. J. Charlton-Perez, M. Giorgetta, E. Manzini, and J. Perlitz (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., 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 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.
- 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.

## BIBLIOGRAPHY

- 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.
- 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.
- 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(1857), 2089–101, 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.
- 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.
- Kodera, K., and Y. Kuroda (2002), Dynamical response to the solar cycle, *J. Geophys. Res.*, 107, 4749, doi:10.1029/2002JD002224.

## BIBLIOGRAPHY

- 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. (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.
- 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.
- 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.*, submitted.
- 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. Atmos.*, (1), n/a–n/a, doi:10.1002/2013JD021403.
- 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.

## BIBLIOGRAPHY

- Matsuno, T. (1970), Vertical Propagation of Stationary Planetary Waves in the Winter Northern Hemisphere., *J. Atmos. Sci.*, 27(6), 871–883.
- 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.
- Mitchell, D. M., A. J. Charlton-Perez, and L. J. Gray (2011a), Characterizing the Variability and Extremes of the Stratospheric Polar Vortices Using 2D Moment Analysis, *J. Atmos. Sci.*, 68, 1194–1213.
- Mitchell, D. M., L. J. Gray, and A. J. Charlton-Perez (2011b), The structure and evolution of the stratospheric vortex in response to natural forcings, *J. Geophys. Res.*, 116, D15,110, doi:10.1029/2011JD015788.
- 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.
- 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.
- Murphy, J. (1990), Assessment of the practical utility of extended range ensemble forecasts, *Q. J. R. Meteorol. Soc.*, 116, 89–125.

- 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.
- Newman, P., and J. Rosenfield (1997), Stratospheric thermal damping times, *Geophys. Res. Lett.*, 24, 433–436.
- Pascoe, C. L., L. J. Gray, and A. A. Scaife (2006), A GCM study of the influence of equatorial winds on the timing of sudden stratospheric warmings, *Geophys. Res. Lett.*, 33, L06,825, doi:10.1029/2005GL024715.
- 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.
- 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.
- 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.
- 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(9), 3048–3064, doi:10.1175/2010JAS3255.1.
- Scaife, A. A., D. R. Jackson, R. Swinbank, N. Butchart, H. E. Thornton, M. Keil, and L. Henderson (2005), 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., D. Copsey, C. Gordon, C. Harris, T. Hinton, S. Keeley, A. O'Neill, M. Roberts, and K. Williams (2011), Improved Atlantic winter blocking in a climate model, *Geophys. Res. Lett.*, 38(23), doi:10.1029/2011GL049573.

## BIBLIOGRAPHY

- 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*.
- 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*(19), 5268–5273, doi:10.1002/grl.50927.
- 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/NGE01698.
- 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.
- 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, 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.
- 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.
- 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 (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.

## BIBLIOGRAPHY

- Taylor, K. E., R. J. Stouffer, and G. a. Meehl (2012), An Overview of CMIP5 and the Experiment Design, *Bull. Amer. Meteor. Soc.*, 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Tennant, W. J., G. J. Shutts, A. Arribas, and S. A. Thompson (2011), Using a Stochastic Kinetic Energy Backscatter Scheme to Improve MOGREPS Probabilistic Forecast Skill, *Mon. Weather Rev.*, 139, 1190–1206, doi:10.1175/2010MWR3430.1.
- 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., J. M. Wallace, and G. C. Hegerl (2000), Annular modes in the extratropical circulation. Part II: Trends, *J. Climate*, (689), 1018–1036.
- Thompson, D. W. J., M. P. Baldwin, and S. Solomon (2005), Stratosphere-troposphere coupling in the Southern Hemisphere, *J. Atmos. Sci.*, pp. 708–715, doi:10.1175/JAS-3321.1.
- Uppala, S. M., Kallberg, P. W., 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(612), 2961–3012.
- Waugh, D. (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.

## BIBLIOGRAPHY

- Waugh, D. W., L. Oman, P. A. Newman, R. S. Stolarski, S. Pawson, J. E. Nielsen, and J. Perlitz (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. Kubar, 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.
- WMO (2011), Scientific Assessment of Ozone Depletion: 2010, *Tech. rep.*, World Meteorological Organization, Global Ozone Research and Monitoring Project, Report 52., Geneva.
- Yamazaki, K. (1987), Observations of the Stratospheric Final Warmings in the Two Hemispheres, *J. Meteor. Soc. Japan*, 65, 51–66.