Key Points:
- CMIP5 models have a wide range of biases in the frequency of polar vortex splits and displacements
- Splits occur more barotropically than displacements in almost all models
- There are consistent differences in the sea level pressure response to splits and displacements

Correspondence to:
W. J. M. Seviour,
wseviou1@jhu.edu

Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models

William J. M. Seviour[1], Lesley J. Gray[1,2], and Daniel M. Mitchell[1,2]

[1]Department of Physics, University of Oxford, Oxford, UK, [2]National Centre for Atmospheric Science, University of Oxford, Oxford, UK

Abstract Sudden stratospheric warming (SSW) events can occur as either a split or a displacement of the stratospheric polar vortex. Recent observational studies have come to different conclusions about the relative impacts of these two types of SSW upon surface climate. A clearer understanding of their tropospheric impact would be beneficial for medium-range weather forecasts and could improve understanding of the physical mechanism for stratosphere-troposphere coupling. Here we perform the first multimodel comparison of stratospheric polar vortex splits and displacements, analyzing 13 stratosphere-resolving models from the fifth Coupled Model Intercomparison Project (CMIP5) ensemble. We find a wide range of biases among models in both the mean state of the vortex and the frequency of vortex splits and displacements, although these biases are closely related. Consistent with observational results, almost all models show vortex splits to occur barotropically throughout the depth of the stratosphere, while vortex displacements are more baroclinic. Vortex splits show a slightly stronger North Atlantic surface signal in the month following onset. However, the most significant difference in the surface response is that vortex displacements show stronger negative pressure anomalies over Siberia. This region is shown to be colocated with differences in tropopause height, suggestive of a localized response to lower stratospheric potential vorticity anomalies.

1. Introduction

Variability of the winter polar stratosphere often precedes significant circulation and temperature anomalies at the Earth's surface. This influence has been shown to be particularly strong following the rapid breakdown of the stratospheric polar vortex; events known as sudden stratospheric warmings (SSWs) [e.g., *Baldwin and Dunkerton*, 2001; *Thompson et al.*, 2002; *Kidston et al.*, 2015]. Recent studies [e.g., *Nakagawa and Yamazaki*, 2006; *Charlton and Polvani*, 2007; *Mitchell et al.*, 2013] have assessed the surface impact of two types of SSW; vortex splits, in which the vortex divides into two separate vortices, and vortex displacements, in which it moves far away from the pole. A better understanding of the relative impacts of vortex splits and displacements would serve to improve medium-range weather predictions following such events and could provide insight into the physical mechanism by which the stratosphere influences the troposphere. This study presents the first multimodel analysis of the simulation of vortex splits and displacements and their surface influence.

Vortex splits and displacements were identified in a 40 year reanalysis data set by *Charlton and Polvani* [2007], who conclude that there is no statistically significant difference in their surface impacts. On the other hand, *Mitchell et al.* [2013] use the same data set but a different method to classify events and conclude that vortex splits have a significantly larger surface impact. *Mitchell et al.* [2013] discussed the possibility that differences in the two methods have led to these contradictory results. Both studies, however, have relatively large statistical uncertainty because of the few SSWs (about 40) in the observational record; uncertainty which is further amplified when subdividing events into vortex splits and displacements.

Mitchell et al. [2013] tested the significance of the surface northern annular mode (NAM) following vortex splits and displacements against randomly chosen winter periods, finding the difference following splits to be significant, but not following displacements (see their Figure 5). However, they did not test whether anomalies following vortex splits and displacements are significantly different from each other. Such a test is shown in Figure 1 which uses events calculated by *Seviour et al.* [2013] (the majority of which coincide with those of *Mitchell et al.* [2013]). It can be seen that while the NAM signal following splits is larger than displacements,
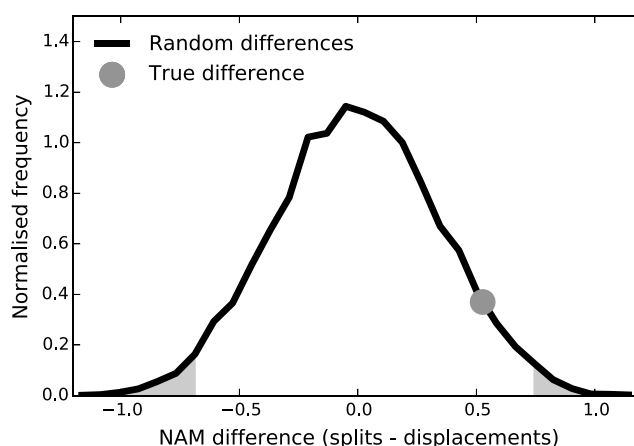
Figure 1. Difference in the surface NAM between vortex splits and displacements over the 30 days following their onset. A probability density function for this difference is calculated by a bootstrap method, randomly sampling 10^4 composites from the joint distribution of splits and displacements. The 95% significant region (according to a two-tailed test; i.e., <2.5% and >97.5%) is shaded and the real composite difference (circle) is at the 94th percentile. Data are taken from the ERA-40 and ERA-Interim reanalyses.

the difference is not statistically significant at the 95% level. This further illustrates uncertainties analyzing the small number of observed events. Modeling studies can compliment these observational results in providing much larger sample sizes of these events.

In this study we identify more than 900 vortex splits and displacements in 13 climate model simulations from the fifth Coupled Model Intercomparison Project (CMIP5) ensemble [*Taylor et al.*, 2012]. SSW events in the CMIP5 ensemble were previously analyzed by *Charlton-Perez et al.* [2013], who found a larger, more realistic frequency of events in "high-top" models (those with a lid height above the stratopause). Motivated by these results and other studies comparing high- and low-top models [e.g., *Cagnazzo and Manzini*, 2009; *Osprey et al.*, 2013], we limit this analysis to high-top models only. *Charlton-Perez et al.* [2013] used the conventional zonal-mean definition of SSW events (a reversal to easterly of the zonal-mean zonal wind at 60°N, 10 hPa). Their analysis is extended here to consider the two-dimensional variability of the polar vortex using vortex moment (or "elliptical") diagnostics [e.g., *Waugh*, 1997], see further details in section 2.2.

Mitchell et al. [2012] previously compared vortex moment diagnostics in climate model simulations from the second Chemistry-Climate Model Validation (CCMVal-2) project, but their analysis was limited because only 3 of the 18 models in CCMVal-2 provided the daily potential vorticity (PV) necessary for the calculation of moment diagnostics. They also did not classify vortex splits and displacements per se, instead focusing on the mean state of the vortex. *O'Callaghan et al.* [2014] have since identified vortex splits and displacements in a primitive equation model using vortex moment diagnostics applied to the PV field. They found significantly stronger positive surface NAM anomalies in the 30 days following vortex splits than following vortex displacements, in agreement with *Mitchell et al.* [2013]. In this study we use the method of *Seviour et al.* [2013] to calculate moment diagnostics based on geopotential height (a quantity output by all CMIP5 models), and we use a simple threshold method to distinguish between vortex splits and displacements.

There are three main objectives to this investigation. First, we wish to evaluate the current state of models' representation of the stratospheric polar vortex and stratosphere-troposphere coupling, including whether there are any consistent biases among models. Second, we aim to determine if there is a relationship between model parameters and biases in the representation of vortex variability, which may motivate future model improvements. Third, we will investigate whether the increased sample size of the CMIP5 ensemble can be used to better understand the surface response to vortex splits and displacements.

2. Data and Methods

2.1. CMIP5 Models

The analysis in this paper considers only climate models with a lid height above the stratopause from the CMIP5 ensemble. In total, 13 such models (listed in Table 1) were available from eight different modeling

Table 1. Parameters of the CMIP5 Models Studied in This Paper[a]

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

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

centers. Although another two (CESM1-WACCM and MIROC-ESM) are listed in the CMIP5 ensemble, appropriate data were not found to be available for these models in the CMIP5 archive (http://cmip-pcmdi.llnl.gov/cmip5/). Of the 13 models, 12 have an uppermost level which is in the upper mesosphere (70–80 km), but CanESM2 has a significantly lower lid which lies close to the stratopause.

Historical simulations have been used throughout this analysis. These include observed climate forcings, such as from greenhouse gases, ozone depletion, land use change, tropospheric and stratospheric aerosols, and solar variability [*Taylor et al.*, 2012]. The simulation period considered is limited to 1958–2005, so that it coincides with the historical reanalysis period (CMIP5 historical simulations end at 2005). Limiting the model simulation analysis to the same period as reanalysis may be important because several studies have suggested that external forcing, such as volcanic eruptions and solar variability, has a significant impact on stratospheric variability [e.g., *Robock*, 2000; *Gray et al.*, 2010]. In order to achieve the largest possible ensemble size, all available ensemble members have been used for each model, which leads to different numbers of years entering the ensemble from different models. This does, however, necessitate that any results appearing in the ensemble mean should also be checked for consistency among the models to ensure that it is not biased by a particular model.

Following *Seviour et al.* [2013], model simulations are compared with the European Centre for Medium-Range Weather Forecasts ERA-40 reanalysis [*Uppala et al.*, 2005] from 1958-1979 and the more recent ERA-Interim [*Dee et al.*, 2011] from 1979 to 2005 (ERA-Interim does not include the presatellite era, so no data is available before 1979). This combination is chosen to maximize the length of the historical period analyzed and is hereafter referred to as "ERA-40/I."

2.2. Moment Diagnostics

The two-dimensional moment diagnostics, M_{nm}, of a distribution, $f(x, y)$, are given in Cartesian coordinates by

$$M_{nm} = \iint_S x^n y^m f(x, y) \, dxdy, \tag{1}$$

where S represents the extent of the distribution and n and m give the order of the moment in the x and y directions, respectively. Applied to the stratospheric polar vortex, these describe the two-dimensional (longitude-latitude) shape of the vortex in terms of an "equivalent ellipse" [*Waugh*, 1997]. Here we focus on two such moment diagnostics; the latitude of the vortex centroid (calculated by setting $n + m = 1$) and the aspect ratio (ratio of major to minor axes; calculated by setting $n + m = 2$). For the full mathematical details of the calculation of these diagnostics, readers are referred to *Matthewman et al.* [2009].
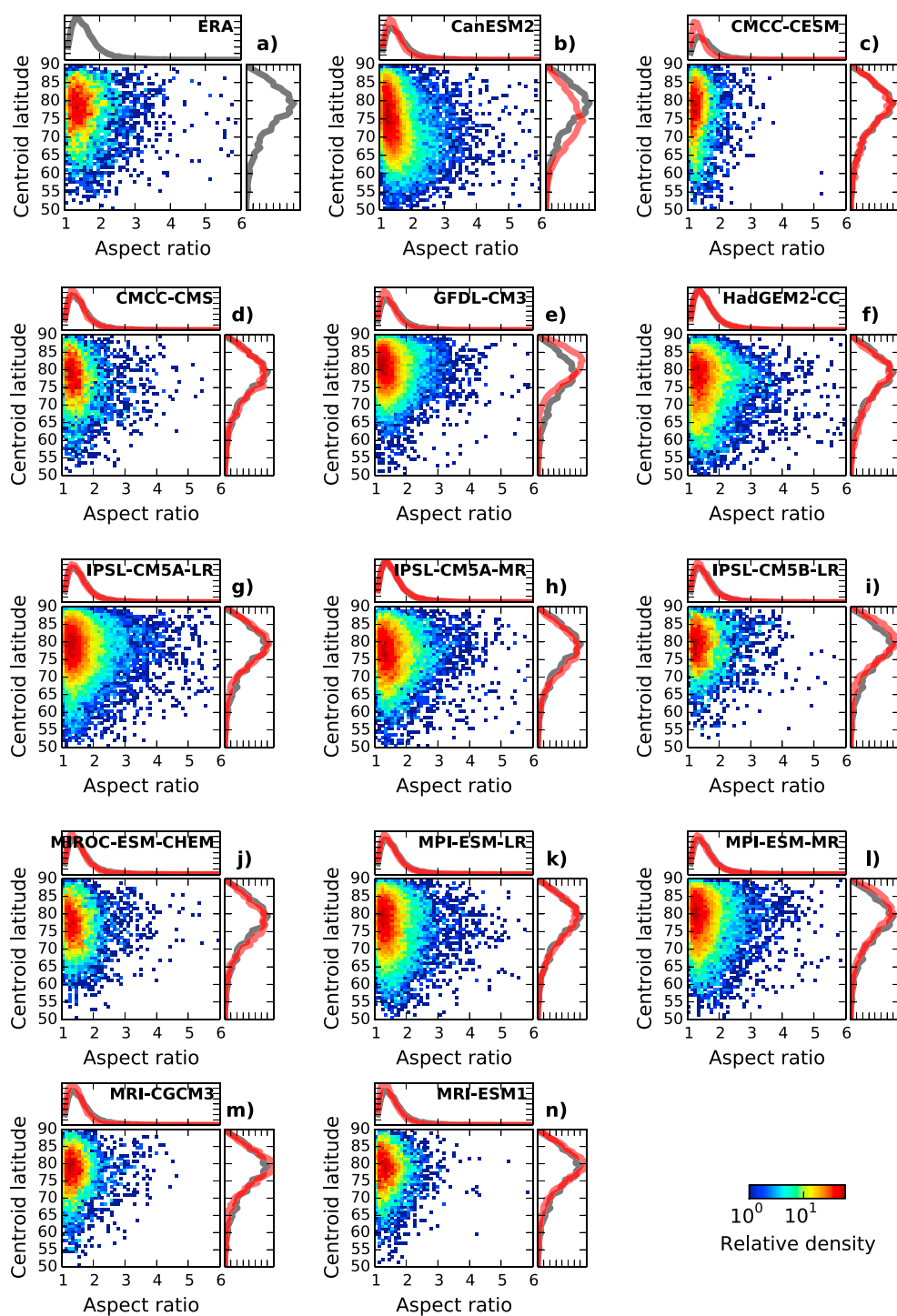
Figure 2. (a–n) Distributions of centroid latitude (*y* axis) and aspect ratio (*x* axis) for ERA-40/I (Figure 2a and grey lines in Figures 2b–2n) and the CMIP5 models (Figures 2b–2n; red lines). Joint distributions are shown with a logarithmic scale such that red squares represent the densest regions.

Previous studies which have calculated moment diagnostics for the stratospheric polar vortex have used PV on an isentropic surface, a quantity which is conserved for adiabatic processes. However, this is not commonly output by climate models and is computationally expensive to calculate, leading the majority of models being excluded from previous studies [*Mitchell et al.*, 2012]. Motivated by this, *Seviour et al.* [2013] developed a method to calculate the moment diagnostics using geopotential height on isobaric levels, a quantity archived by almost all climate models. They showed this to be highly correlated with the PV-derived diagnostics, and so geopotential height is used in the present study. In order to calculate the moment diagnostics, it is necessary to define a contour representing the vortex edge. For each model, this is taken to be the value of the December–March mean zonal-mean geopotential height at 60°N and 10 hPa. This allows for any biases in the mean geopotential height between different models.

Vortex splits and displacements are defined according to the threshold method described by *Seviour et al.* [2013]. A vortex split is identified when the aspect ratio remains higher than 2.4 for 7 days or more, and a displacement requires the centroid latitude to remain equatorward of 66°N for 7 days or more. These thresholds were chosen to give a similar frequency of vortex splits and displacements as the previous studies of *Charlton and Polvani* [2007] and *Mitchell et al.* [2013]. *Seviour et al.* [2013] showed that of 35 events identified using this method in reanalyses from 1958 to 2009, just two events were not captured by either of the past studies. In calculating events for the CMIP5 models the same thresholds are used for each model, so as to identify, as much as possible, geometrically equivalent events. As in *Seviour et al.* [2013], events are limited to December–March.

3. Results

3.1. Vortex Mean State and Variability

Joint distributions of daily centroid latitude and aspect ratio from each of the models are shown in Figure 2, along with those from ERA-40/I. For each model the joint distribution histogram is plotted with a logarithmic color scale which is normalized according to the number of days in each model simulation. The joint distribution for ERA-40/I 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. The shape of this distribution is well replicated by most of the models, although CanESM2 has a significantly different shape, with the high aspect ratio/equatorward centroid latitude being more common.

There are a range of biases among models. CanESM2 has a modal centroid latitude which is about 5° more equatorward than ERA-40/I. Contrastingly, GFDL-CM3 has a modal centroid latitude about 2.5° more poleward. CMCC-CESM displays a clear bias in the aspect ratio, with a distribution much less skewed toward high values than in reanalysis. The majority of these biases are consistent throughout the winter season, with the exception of CanESM2, for which the equatorward bias is stronger in early winter. The seasonally varying averages of the moment diagnostics for each model are shown in the supporting information (Figure S1).

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

As well as the large differences in the combined frequency of vortex splits and displacements, Figure 3 shows that the ratio of frequencies of vortex splits to displacements varies significantly between models. For instance, CanESM2 and CMCC-CESM simulate almost entirely vortex displacements, while IPSL-CM5B-LR and GFDL-CM3 simulate almost entirely vortex splits. In the multimodel mean (MMM) these biases largely cancel to give an approximately equal ratio of splits to displacements, which is in agreement with reanalysis.

The seasonal distribution of these displaced and split vortex events is illustrated in Figure 4. Some models (CMCC-CMS, HadGEM2-CC, and IPSL-CM5A-LR) replicate the observed distribution, with split vortex events
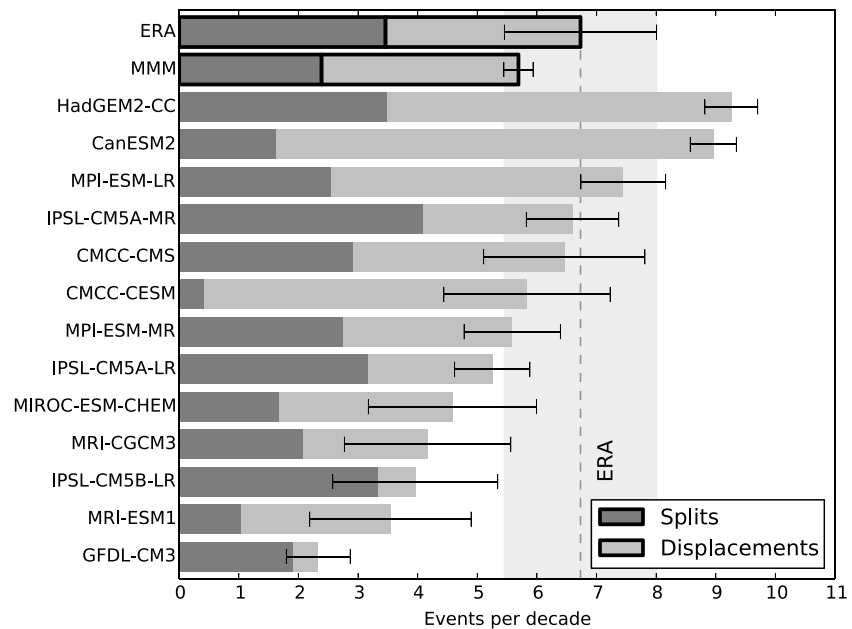
Figure 3. Frequency of split and displaced vortex events in the CMIP5 models, ERA-40/I, and the multimodel mean (MMM). Error bars are for the frequency of all events and represent 1σ range, assuming a binomial distribution of events. The grey shaded region represents the 1σ range for ERA-40/I, along with the mean (dashed line).

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.

Composites of 10 hPa geopotential height over the 10 days following the onset of splits and displacements are shown in Figure 5 for ERA-40/I and the MMM. The mean shape of splits and displacements is very similar in the MMM and in ERA-40/I, with splits occurring approximately along the 90°W to 90°E axis and displacements with a vortex shifted toward Scandinavia and Siberia. The same features can also be seen in the majority of individual models (see supporting information, Figure S2). This confirms that the method used here succeeds in capturing similar events to those seen in observations.

It is now considered how model biases in the climatology of the stratospheric polar vortex affect the frequency of vortex splits and displacements. 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. This quantity represents the most likely state of the vortex, and unlike the mean, is not affected by extreme values (the relationship between the mean and mode of aspect ratio and centroid latitude is shown in the supporting information, Figure S3). The peak can be estimated by the maximum value of a histogram; however, this introduces significant random errors and is sensitive to the selection of bin size. A more accurate estimation of the mode can be made by fitting the aspect ratio and centroid latitude with an analytic distribution and then finding the peak of that distribution. Following *Mitchell et al.* [2011], we fit the aspect ratio with a generalized extreme value (GEV) distribution of the form

$$f(x; \mu, \sigma, \xi) = \frac{a^{(-1/\xi)-1}}{\sigma} e^{-a^{-1/\xi}}, \qquad (2)$$

with

$$a = 1 + \xi \frac{x - \mu}{\sigma}, \qquad (3)$$

where μ is the location parameter, σ the scale parameter, and ξ the shape parameter. These parameters are determined using the method of maximum-likelihood estimation [*Wilks*, 2006]. This method is also used to fit a Gaussian distribution to the cube of the centroid latitude and then the cube root taken to return the original distribution (this is carried out because an analytic distribution does not fit the unscaled centroid latitude). The use of these distributions is statistically, rather than physically motivated. *Mitchell et al.* [2011]
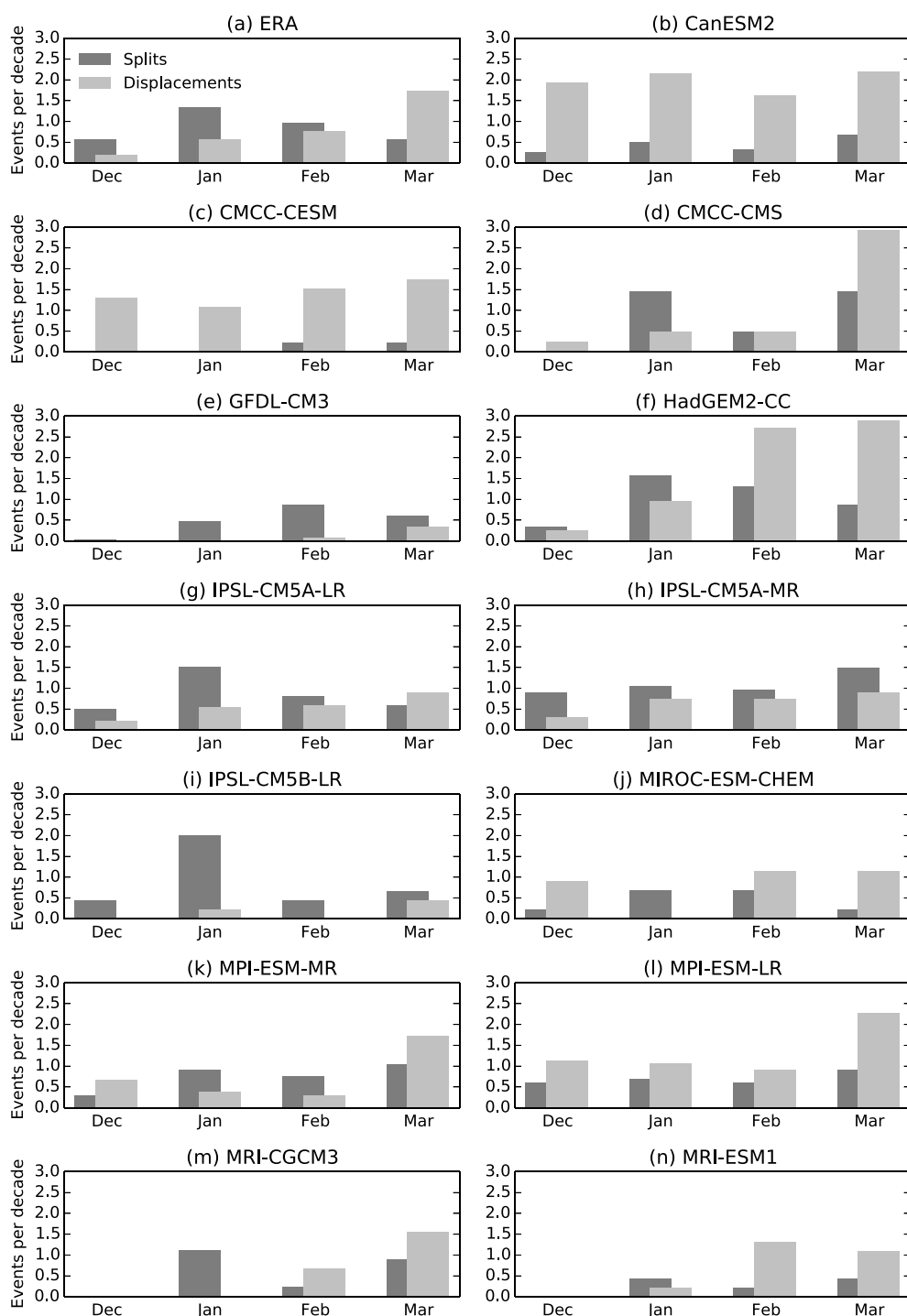
Figure 4. Seasonal distribution of the occurrence of vortex splits and displacements in (a) ERA-40/I and (b–n) the CMIP5 models.

found them to accurately fit the histograms of centroid latitude and aspect ratio in reanalysis data, except for 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 6 shows the relationship between the modal aspect ratio and centroid latitude and frequency of vortex splits and displacements. A strong linear relationship exists in both cases, demonstrating that biases in the average state of the vortex account for the vast majority of intermodel spread in the representation of
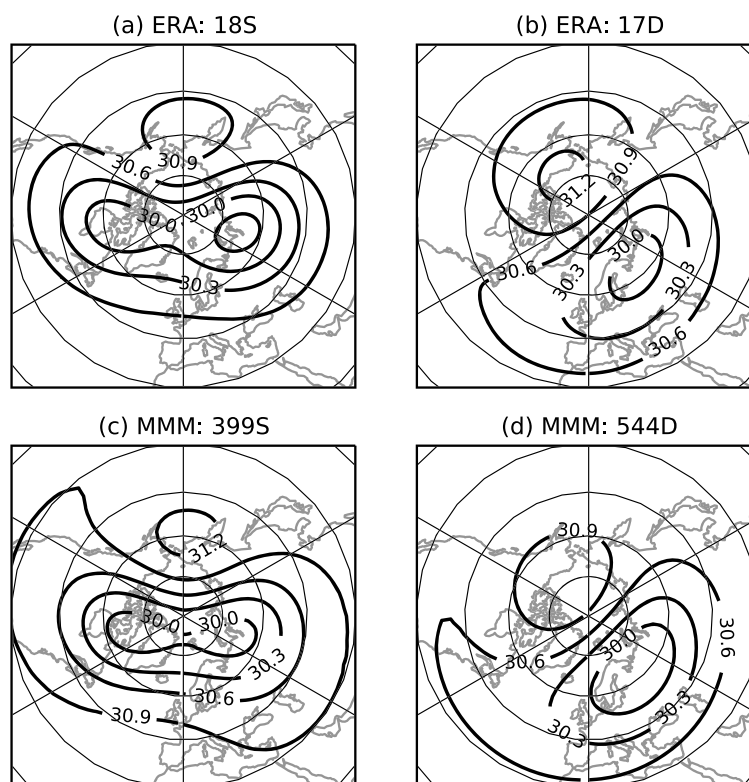
Figure 5. Composites of 10 hPa geopotential height over the 10 days following the onset of split (S) and displaced (D) vortex events in (a,b) ERA-40/I and (c,d) the CMIP5 multimodel mean. The number of events entering the composite and their type are shown in the title of each panel. The contour interval is 0.3 km.

extremes. Hence, models are consistent in their representation of the variability of aspect ratio and centroid latitude, relative to the model climatology. Figure 6 also shows that the values for ERA-40/I 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 vortex splits and displacements can be significantly improved by a more accurate average vortex state. This is an extension of the result of *Butchart et al.* [2011], who related the model mean polar stratospheric state to the frequency of traditionally defined SSWs. Furthermore, while the ERA-40/I 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-40/I, indicating that a too circularly symmetric vortex is a common bias among models.
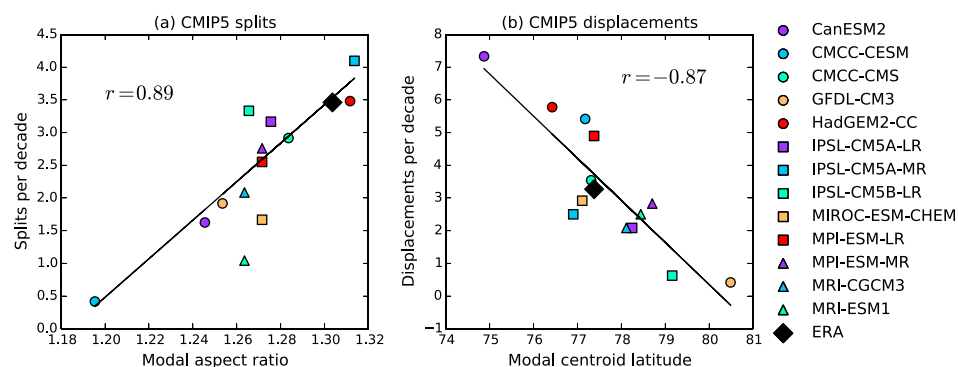

Figure 6. Comparison of the (a) DJFM mean aspect ratio with frequency of split vortex events and (b) DJFM mean centroid latitude with frequency of displaced vortex events in the CMIP5 ensemble and ERA-40/I. Linear regression lines and the corresponding correlation coefficients are also shown.
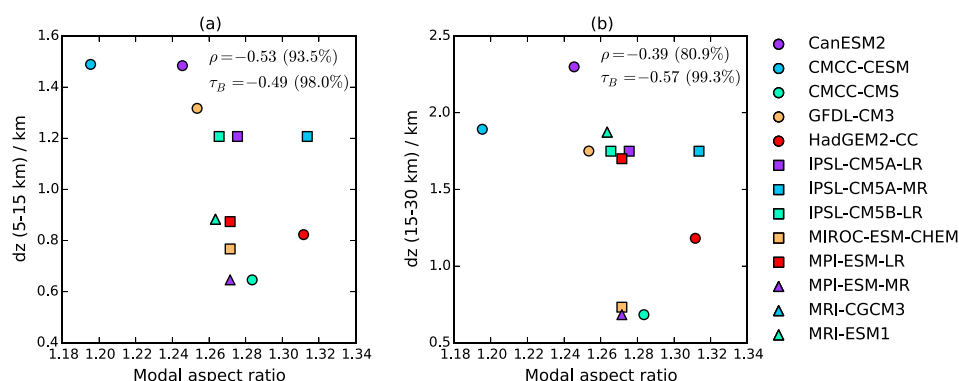
Figure 7. Correlation between modal aspect ratio and vertical resolution over two regions: (a) 5–15 km and (b) 15–30 km. Linear (ρ) and rank (τ_B) correlations are shown, along with their significance levels (for the null hypothesis of zero correlation).

Given that the modal aspect ratio and centroid latitude are closely related to the frequency of splits and displacements, it is of interest whether any model parameters in turn control the modal aspect ratio and centroid latitude. Relationships with the parameters listed in Table 1 have been tested, but there are no statistically significant correlations between the modal aspect ratio or centroid latitude and horizontal resolution or between the centroid latitude and vertical resolution. However, a stronger relationship is found between vertical resolution and the modal aspect ratio and this is shown in Figure 7. These relationships appear quite nonlinear, with aspect ratio being relatively more sensitive to changes in resolution when the resolution is coarse (high dz) and less sensitive when the resolution is finer (low dz). This can be seen in that the Spearman's rank correlations (which test the monotonicity of the relationships) are more statistically significant than the linear correlations. Even so, the relatively wide scatter of points indicates that vertical resolution fails to account for a substantial fraction of intermodel variability in the modal aspect ratio.

The two measures of vertical resolution are themselves correlated ($r = 0.79$) so it is difficult to interpret which of the two regions (if any) has the largest impact on the modal aspect ratio. However, previous studies have shown there to be strong vertical gradients in the static stability and vertical shear of zonal wind near the tropopause [*Chen and Robinson*, 1992; *Grise et al.*, 2010], both of which affect the planetary wave refractive index [*Matsuno*, 1970] and so the propagation of planetary waves into the stratosphere. Hence, the ability to resolve sharp vertical gradients in this region could significantly affect the simulation of stratospheric polar vortex variability. Other factors not considered here, such as the model representation of gravity waves, are also likely to play an important role.

3.2. Stratosphere-Troposphere Coupling

The time-height evolution of the atmosphere before and after vortex splits and displacements in ERA-40/I and the CMIP5 multimodel mean are displayed in Figure 8. This shows composites of polar cap (60°–90°N) average 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. Polar cap Z is highly correlated ($r > 0.95$) with the northern annular mode (NAM) (calculated from zonal mean Z according to the method of *Baldwin and Thompson* [2009]) over the levels shown in Figure 8. *Kushner* [2010] also demonstrated composites of the NAM and polar cap Z following SSWs to be very similar.

The MMM is calculated so as to give each event an equal weight (rather than each model) and so does not give undue weight to models with only a small number of events. On the other hand, this does mean that greater weight is given to models with more ensemble members and more events (almost one third of all displaced vortex events come from CanESM2). Hence, it is important to check features seen in the MMM for consistency among models. Individual composites for each model are shown in the supporting information (Figure S4).

There are large intermodel differences in the evolution of polar cap Z following vortex splits and displacements, both in the stratosphere and troposphere. In the MMM, these combine to give a persistence of lower stratospheric anomalies slightly greater than reanalysis and relatively weak stratospheric precursors for both vortex splits and displacements. Models vary most in the tropospheric anomalies over 10–90 days following events, with some showing weaker anomalies than reanalysis and others stronger.
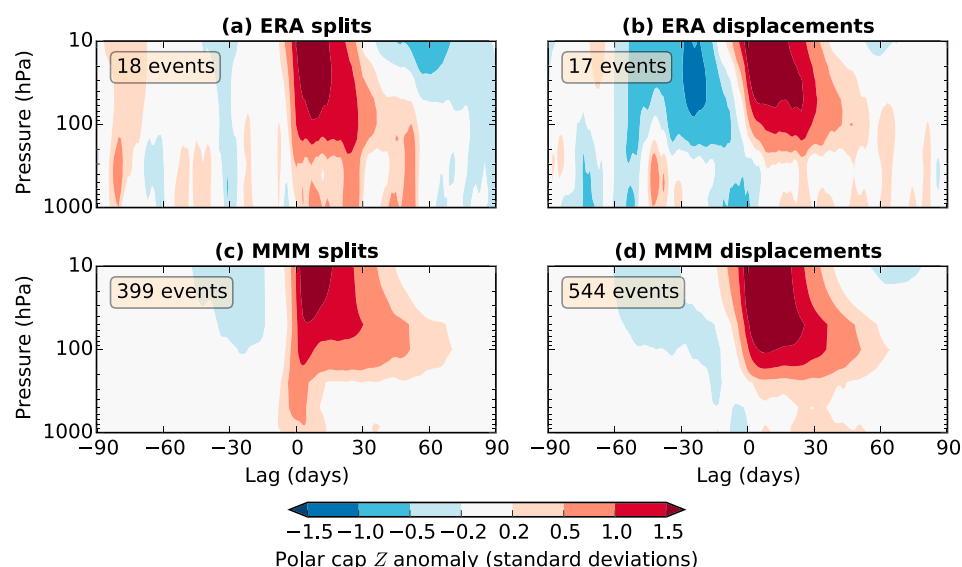
Figure 8. Composites of normalized polar cap averaged geopotential height following splits and displacements in (a,b) ERA-40/I and (c,d) the CMIP5 multimodel mean. Numbers in the upper left of each panel indicate the number of events entering each composite.

As well as these large intermodel differences, there are also some consistent features among models which are apparent in the MMM. Almost all models show a barotropic onset to vortex splits, with anomalies occurring at the same time throughout the depth of the stratosphere. In contrast, vortex displacements appear more baroclinic, with onset occurring first in the upper or midstratosphere and descent to the lower stratosphere taking about 2 weeks. The same difference in baroclinicity is found in reanalysis, indicating that it is likely to be a robust difference.

Mean sea level pressure (MSLP) anomalies averaged over the 30 days following vortex splits and displacements are shown in Figure 9 for ERA-40/I and the CMIP5 MMM, again calculated so as to give each event an equal weight. The climatology from which anomalies are calculated is the average for each day of the year at each spatial location, smoothed with a 10 day running mean. Composites for individual models are shown in the supporting information (Figure S5).

Following both vortex splits and displacements, all models show a positive MSLP anomaly near the North Pole and a negative anomaly centered over Western Europe and the North Atlantic. This pattern gives a negative projection onto the North Atlantic Oscillation and is on average slightly (about 2 hPa) stronger following splits than displacements, although this difference is not statistically significant (see Figure 10). Less consistent among models are anomalies over the North Pacific; many models (e.g., MRI-CGCM3 and 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 MMM, a weakly positive North Pacific anomaly is seen.

This inconsistency in the Pacific anomalies has important consequences for the interpretation of the zonal mean anomalies following vortex split and displacement events shown in Figure 8. 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 but a relatively strong North Atlantic oscillation (NAO) signal, particularly following split vortex events. The reason for this difference is that the model also shows relatively strong positive North Pacific anomalies that to some extent cancel the North Atlantic anomalies in the polar cap average. Such an effect would also be seen in the NAM, even if not calculated from zonally averaged Z, since the surface NAM pattern has centers of action of the same sign in the North Atlantic and North Pacific [e.g., *Ambaum et al.*, 2001].
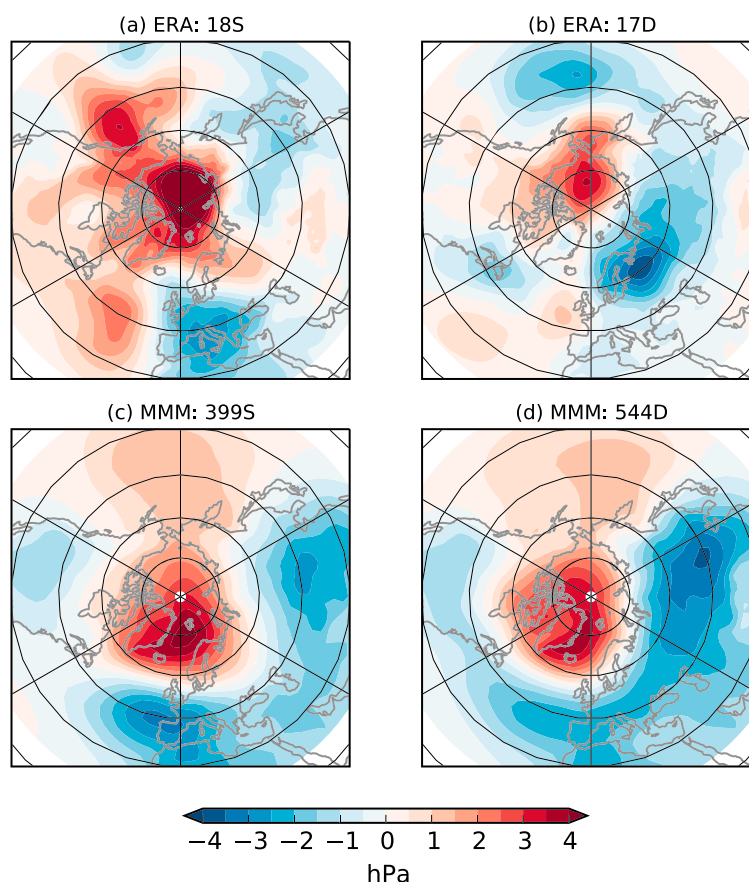
Figure 9. Composites of mean sea level pressure anomalies averaged over 0–30 days following splits (S) and displacements (D) in (a,b) ERA-40/I and the (c,d) CMIP5 multimodel mean. The number of events entering the composite and their type are shown in the title of each panel.

Figure 10 shows the vortex split minus displacement composite difference for MSLP averaged 0–30 days following event onset for both ERA-40/I 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 vortex splits and displacements are populations from the same probability distribution. The bootstrap is carried out by randomly resampling with replacement from the distribution of all events, to create 5000 random composite differences. For the case of ERA-40/I, 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-40/I.

In the CMIP5 MMM difference the most significant feature is the large positive anomaly (a result of a more negative anomaly following vortex displacements) 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 asymmetric and so does not project strongly onto the polar cap average, therefore explaining the small difference in polar cap averaged Z during this period (Figures 8c and 8d).

In order to investigate the origin of the different surface response following vortex splits and displacements, lower stratospheric anomalies are studied. Figure 11 shows composites of 100 hPa geopotential height averaged over the 10 days following vortex splits and displacements for ERA-40/I and the CMIP5 MMM. The difference of splits minus displacements is also shown in each case. Because this is a mean of model absolute values, and models have different climatologies, the MMM for vortex splits and displacements 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. The MMM difference is seen to be remarkably similar to the reanalysis with positive values over most of Siberia and Europe and negative over Canada. This pattern is also consistent among individual models (Figure S6 in the supporting information).
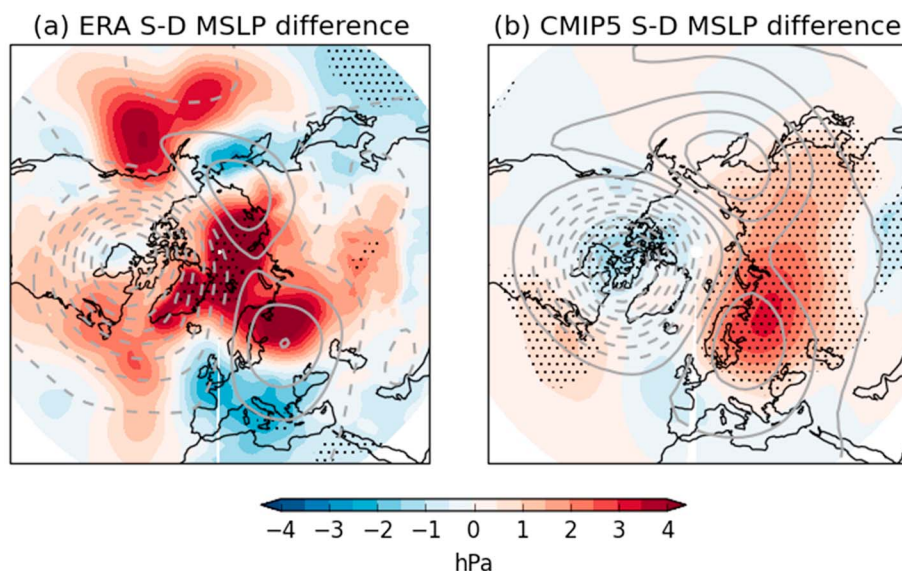
Figure 10. Difference (split minus displacement) of composites of mean sea level pressure averaged 0–30 days following split (S) and displaced (D) vortex events in ERA-40/I and the CMIP5 ensemble. Stippling indicates regions that are >95% significant according to a two-tailed bootstrap test. Grey contours represent the difference in 100 hPa geopotential height averaged 0–10 days following events, the contour interval is 40 m, dashed contours represent negative values, and the lowest magnitude contours are ±20 m.

This 100 hPa geopotential height difference is also overlaid on Figure 10. It can be seen that the positive 100 hPa Z anomalies over Siberia overlie the positive MSLP anomalies, while the negative 100 hPa Z over northern Canada overlies negative MSLP anomalies. A somewhat similar but not statistically significant pattern is seen in ERA-40/I, although the Siberian anomaly is more poleward and the negative anomaly over Canada is much weaker. The implications of this result for mechanisms of stratosphere-troposphere coupling are discussed in section 4.2.
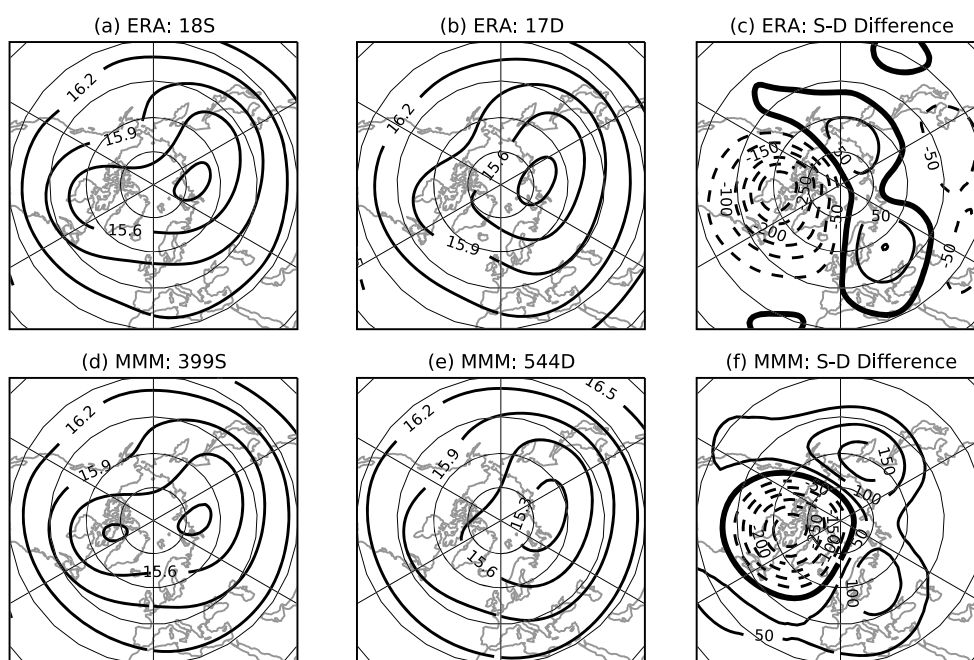

Figure 11. Composites of 100 hPa geopotential height over the 10 days following the onset of split (S) and displaced (D) vortex events in (a, b) ERA-40/I and the (d, e) CMIP5 multimodel mean. (c,f) The difference of splits minus displacements is also shown in each case. The number of events entering the composite and their type are shown in the title of each panel. The contour interval is 0.3 km and dashed contours represent negative values.

4. Discussion

4.1. Measures of Stratosphere-Troposphere Coupling

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

Many studies of stratosphere-troposphere coupling have focused on the lag-height behavior of 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 [e.g., *Manzini et al.*, 2014] make further approximations as to the zonal nature of the coupling by calculating the NAM based on zonal-mean geopotential height, according to the method of *Baldwin and Thompson* [2009]. Our results suggest that because of the difference in model consistency over the two ocean basins, zonal-mean diagnostics or the NAM alone are not good descriptors of intermodel variability. Therefore, we suggest that the NAO index or the full two-dimensional surface fields should be considered when making intermodel comparisons.

4.2. Difference Between Vortex Splits and Displacements

Among the CMIP5 models, the most significant difference in the response to vortex splits and displacements is the more barotropic nature of splits relative to displacements (Figure 8). The same difference is seen in reanalysis data, as shown in this study and previously by *Mitchell et al.* [2013]. This is consistent with the idea that resonant excitation of the barotropic mode [*Esler and Scott*, 2005] plays a significant role in the occurrence of vortex splits. Contrastingly, the more baroclinic nature of vortex displacements is suggestive of the descent of a Rossby wave critical layer [*Matsuno*, 1970, 1971].

We have also found that there are some consistent differences in surface anomalies following the two types of event. In particular, MSLP anomalies following displaced vortex events are more negative over Scandinavia and Siberia than following split vortex events. From the fact that these MSLP differences are colocated with 100 hPa Z (Figure 10), which is in turn related to tropopause height, it may be possible to gain some understanding of the mechanism behind the difference in the surface response to split and displaced vortex events. Specifically, the colocation of surface anomalies and tropopause height is suggestive of a localized increase/decrease of relative vorticity caused by stretching/compression of the tropospheric column (see discussion in *Mitchell et al.* [2013]). Changes in tropopause height are, in turn, caused by the bending of isentropic surfaces toward PV anomalies resulting from the movement of the stratospheric polar vortex [*Ambaum and Hoskins*, 2002]. This argument relates only to the mechanism underlying the difference between the surface responses to split and displaced vortex events and not to the overall responses. It is important to note that there are many similarities in the responses, especially in the NAO region.

5. Summary

This study has analyzed the simulation splits and displacements of the stratospheric polar vortex in 13 high-top CMIP5 models. The main conclusions are as follows:

1. There are a wide range of biases among models in the average state of the stratospheric polar vortex. Some models have a vortex which is too equatorward, others too poleward. The majority of models have a vortex which is too circularly symmetric. Of the model parameters considered in this study, this bias is most related to vertical resolution in the upper troposphere/lower stratosphere, but the relationship is not highly statistically significant and requires more systematic investigation.

2. There is also a wide spread in the frequency of split and displaced vortex events, although the multimodel mean is in agreement with the observed frequency. Importantly, biases in the average state of the vortex relate closely to biases in the frequency of splits and displacements. Hence, an improvement in the average state of the vortex is likely to lead to an improvement in the representation of extremes. Furthermore, a model's frequency of splits and displacements can be closely estimated from its mean state, which requires far fewer years of simulation to determine.

3. In agreement with reanalysis, almost all models show vortex splits to occur barotropically, whereas anomalies associated with vortex displacements descend through the stratosphere over a period of weeks.

This results in a stronger instantaneous annular mode-like signal at the onset of vortex splits. It also points toward different physical mechanisms driving vortex splits and displacements; in particular, wave resonance [*Esler and Scott*, 2005] playing an important role in vortex splits.

4. Robust differences in sea level pressure anomalies following splits and displacements in the CMIP5 models have been found to be colocated with the difference in near-tropopause geopotential height anomalies. This is consistent with a localized tropospheric response to stratospheric PV anomalies [*Ambaum and Hoskins*, 2002] being the mechanism behind the different responses to the two events. This result only applies to the difference between responses to splits and displacements, not the overall responses which share many similarities.

References

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

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

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*, 3956–3971, doi:10.1002/jgrd.50231.

Baldwin, M. P., and T. J. Dunkerton (2001), Stratospheric harbingers of anomalous weather regimes, *Science*, *294*, 581–584, doi:10.1126/science.1063315.

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.479.

Butchart, N., et al. (2011), Multimodel climate and variability of the stratosphere, *J. Geophys. Res.*, *116*, D05102, doi:10.1029/2010JD014995.

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

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

Charlton-Perez, A. J., et al. (2013), On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models, *J. Geophys. Res. Atmos.*, *118*, 2494–2505, doi:10.1002/jgrd.50125.

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

Davini, P., C. Cagnazzo, and J. A. Anstey (2014), Blocking view of the stratosphere-troposphere coupling, *J. Geophys. Res. Atmos.*, *119*, 11,100–11,115, doi:10.1002/2014JD021703.

Dee, D. P., et al. (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.

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

Gray, L. J., et al. (2010), Solar influences on climate, *Rev. Geophys.*, *48*, RG4001, doi:10.1029/2009RG000282.

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*, 2275–2292, doi:10.1175/2009JCLI3369.1.

Kidston, J., A. A. Scaife, S. C. Hardiman, D. M. Mitchell, N. Butchart, M. P. Baldwin, and L. J. Gray (2015), Stratospheric influence on tropospheric jet streams, storm tracks and surface weather, *Nat. Geosci.*, *8*, 433–440, doi:10.1038/ngeo2424.

Kushner, P. J. (2010), Annular modes of the troposphere and stratosphere, in *The Stratosphere: Dynamics, Transport, and Chemistry, Geophys. Monogr. Ser.*, vol. 190, edited by L. M. Polvani, A. H. Sobel, and D. W. Waugh, pp. 59–91, AGU, Washington, D. C.

Manzini, E., et al. (2014), Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere-Troposphere coupling, *J. Geophys. Res. Atmos.*, *119*, 7979–7998, doi:10.1002/2013JD021403.

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

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

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. Clim.*, *22*, 1566–1585, doi:10.1175/2008JCLI2365.1.

Mitchell, D. M., A. J. Charlton-Perez, and L. J. Gray (2011), Characterizing the variability and extremes of the stratospheric polar vortices using 2D moment analysis, *J. Atmos. Sci.*, *68*, 1194–1213.

Mitchell, D. M., et al. (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. Clim.*, *26*, 2668–2682, doi:10.1175/JCLI-D-12-00030.1.

Nakagawa, K. I., and K. Yamazaki (2006), What kind of stratospheric sudden warming propagates to the troposphere?, *Geophys. Res. Lett.*, *33*, L04801, doi:10.1029/2005GL024784.

O'Callaghan, A., M. Joshi, D. Stevens, and D. Mitchell (2014), The effects of different sudden stratospheric warming types on the ocean, *Geophys. Res. Lett.*, *41*, 7739–7745, doi:10.1002/2014GL062179.

Osprey, S. M., L. J. Gray, S. C. Hardiman, N. Butchart, and T. J. Hinton (2013), Stratospheric variability in twentieth-century CMIP5 simulations of the met office climate model: High top versus low top, *J. Clim.*, *26*, 1595–1606, doi:10.1175/JCLI-D-12-00147.1.

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

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

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*, 485–498, doi:10.1175/BAMS-D-11-00094.1.

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

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.

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

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