University of BRISTOL

University of Bristol - Explore Bristol Research
General rights

Key Points:
- The frequency of split SSW events in particular is underrepresented in most models compared to reanalysis
- Biases in split SSW frequency are associated with the climatological vortex aspect ratio and lower stratospheric wind representation
- Biases in displacement SSW frequency are associated primarily with the climatological vortex location

Persistent Model Biases in the CMIP6 Representation of Stratospheric Polar Vortex Variability

Richard J. Hall[1] [ID], **Daniel M. Mitchell[1]** [ID], **William J. M. Seviour[2]** [ID], and **Corwin J. Wright[3]** [ID]

[1]School of Geographical Sciences and Cabot Institute for the Environment, University of Bristol, Bristol, UK, [2]Global Systems Institute and Department of Mathematics, University of Exeter, Exeter, UK, [3]Centre for Space, Atmospheric and Oceanic Science, University of Bath, Bath, UK

Abstract Sudden stratospheric warmings (SSWs) can have major impact on surface wintertime weather, especially at mid-high latitudes. We do not yet have a complete understanding of why some of these events influence our weather more than others, but one factor may be the dynamical nature of the SSW; whether it involves a split or a displacement of the polar vortex, and one way to explore this is through comprehensive climate models. Here, we analyze the stratospheric dynamics of SSWs within models from the sixth Coupled Model Intercomparison Project (CMIP6). All CMIP6 models simulate SSWs to some degree, although we find a persistent bias in the relative underrepresentation of split vortex events. When comparing with CMIP5 models, large biases persist despite significant model improvements in resolution and in representing atmospheric processes. We show that the simulated displacement frequency is strongly related to climatological lower stratospheric eddy heat flux. The split frequency, on the other hand, is not related to lower stratospheric eddy heat flux, but is strongly related to both the vortex geometry (aspect ratio) and lower stratospheric zonal winds. This suggests that those models with a large positive bias in zonal winds may inhibit the propagation of zonal wavenumber 2 planetary waves from the troposphere, which are associated with split events. Our results suggest how future model development may address these longstanding biases.

Plain Language Summary In winter, the polar stratosphere is very cold and surrounded by strong westerly winds, known as the stratospheric polar vortex. This vortex can be disrupted, and the polar stratosphere can warm by as much as 40K in a few days. This is known as a sudden stratospheric warming (SSW). These events are characterized as splits, when the vortex splits into two smaller vortices, or displacements, where the vortex is shifted away from the pole. Since SSWs were discovered in the 1950s, there have been roughly equal numbers of splits and displacements. Climate models simulate SSWs, although they usually simulate more displacements than splits. We investigate the reasons for this model bias. Those models which have too-circular a vortex tend to simulate too few split events. In addition, atmospheric waves from the surface can disrupt the vortex. If the lower stratospheric westerly winds are too strong in a model, some of these waves that favor split events can be filtered out. On the other hand, models with a vortex that is on average centered too far from the pole tends to simulate too few displacement events. Our results are useful for supporting the future development of climate models.

1. Introduction

In the winter hemisphere, high latitudes receive little solar radiation, and in the stratosphere a strong meridional temperature gradient develops, with a concomitant strong vertical zonal wind shear, strong westerly winds forming the edge of the stratospheric polar vortex (SPV), which has a cold core located approximately over the pole. In the northern hemisphere, this vortex is more variable, due to increased troposphere-to-stratosphere planetary wave propagation, arising from increased land-sea temperature contrasts and topographical variations relative to the southern hemisphere (e.g., Waugh & Randel, 1999). If stratospheric winds are westerly and not too strong, planetary waves can propagate into the stratosphere, where they may disturb the polar vortex (Charney & Drazin, 1961; Matsuno, 1970). A feature of the northern hemisphere polar vortex variability is that the vortex can break down over a period of days. Temperatures over the pole can increase by up to 40K in this short time period (e.g., Butler et al., 2017; Scherhag, 1952) and the mean zonal winds can reverse. Such events, known as major sudden stratospheric warmings (SSWs), take place approximately six times per decade in the northern hemisphere, based on a relatively short

observational record beginning in the late 1950s (Charlton & Polvani, 2007), with only a single such event observed in the southern hemisphere (Baldwin et al., 2021; Kruger et al., 2005).

SSWs can trigger extreme weather, particularly in the northern hemisphere, where they are associated with a negative phase of the North Atlantic Oscillation, accompanied by cold snaps over Europe and Eurasia (Kolstad et al., 2010; King et al., 2019). Their surface impacts can be detected up to 60 days after SSW onset in the stratosphere (Baldwin & Dunkerton, 2001). Consequently, a greater understanding of these events is crucial in developing climate forecasts. The realistic capture of SSWs is vital in accurately representing climate variability and potential future weather and climate extremes (Ayarzagüena et al., 2018).

Climate models simulate SSWs, although there is a widespread of SSW frequency of occurrence (hereafter frequency) amongst models (Ayarzaguena et al., 2020; Charlton-Perez et al., 2013; Wu & Reichler, 2020), with most models tending to underestimate the frequency (Charlton et al., 2007; Seviour et al., 2016, hereafter S16). As noted by Polvani et al. (2017), internal variability of the polar vortex means that SSW frequencies can vary across different simulations of the same model, even where the atmosphere model is identical. However, due to the small sample size of SSWs in the observational record, climate models are important for obtaining larger sample sizes, from which more robust inferences about the real world can be made, contingent on faithful representation of the physical processes governing stratospheric variability.

Model lid height and the vertical resolution of the stratosphere are important factors that can affect SSW frequency. "Low-top" models, with an uppermost level below the stratopause, tend to produce fewer SSWs (e.g., Cagnazzo & Manzini, 2009; Charlton-Perez et al., 2013; Osprey et al., 2013), while vertical resolution is important for representation of the stratospheric circulation. Lee and Black (2015) attributed this reduction of SSW occurrence to the variability of both planetary wave activity and vortex strength being too low in low-top models. However, model lid height and stratospheric resolution alone are not sufficient to ensure an accurate representation of SSW frequency. Some low-top models have frequencies consistent with observations, although the seasonal distribution of SSW events does not necessarily match with observations, and the spread is still large amongst stratosphere-resolving models (e.g., Wu & Reichler, 2020). In a study of the fifth Coupled Model Intercomparison Project (CMIP5) high top models, S16 found that the stratospheric vertical resolution was not significantly correlated with the frequency of split or displacement SSWs.

Climatological properties of the model SPV may also be important in determining SSW frequency. There is an inverse correlation between vortex strength and SSW frequency (Jucker et al., 2014) as when zonal winds are stronger, it is harder to achieve wind reversal, and planetary wave propagation at lower wavenumbers is inhibited as wind speeds increase (Charney & Drazin, 1961). Alternatively, a model with high SPV variability may also contain more SSWs (e.g., Kim et al., 2017). In turn, the climatological strength and variability of the SPV zonal winds are influenced by average wave forcing from the troposphere (Taguchi, 2017). The state of the lower stratosphere may be important in modulating this wave forcing, acting as a valve for wave propagation (Martineau et al., 2018).

SSWs have been classified according to changes in vortex morphology. A split event is where the vortex separates into two distinct "child vortices," whereas a displacement event occurs when the vortex shifts equatorward (e.g., Charlton & Polvani, 2007; Mitchell et al., 2011; Seviour et al., 2013, hereafter S13). The more barotropic annular mode time-height profile of split events is consistent with the development of SSWs by internal resonances (e.g., Esler & Scott, 2005; Matthewman & Esler, 2011), with anomalies occurring synchronously throughout the depth of the stratosphere (S16), while displacements may be associated with the descent of a Rossby wave critical layer (Matsuno, 1970, 1971). Recent studies have reached different conclusions concerning whether displacement and split events have different causes and surface impacts. Some studies find little difference between the surface impacts of the two types (e.g., Charlton & Polvani, 2007; Cohen & Jones, 2011; Maycock & Hitchcock, 2015), while Mitchell et al. (2013), S13 and S16 identify a stronger surface signal at shorter time lags for split SSWs, and Hall et al. (2021) find significantly lower temperatures associated with split events over NW Europe. In part, some of these different conclusions can be attributed to different methodologies for classifying split and displacement events (e.g., Maycock & Hitchcock, 2015).

In observational data, the number of split and displacement events are roughly equal, over the short observation period (e.g., Charlton & Polvani, 2007). Several studies have assessed total SSW frequency in climate

Table 1
Summary of Model Parameters

Model	Lid height km and hPa	Total number of levels	$dz1$ 1–100 hPa m	$dz2$ 50–150 hPa m	$dz3$ 1–50 hPa m
CanESM5	48.4 km/1 hPa	49	1,697	1,282	1,826
MIROC6	87 km/0.004 hPa	81	871	769	913
MPI-ESM-1-2-HAM*	80 km/0.01 hPa	47	2,149	1,538	2,282
MPI-ESM1-2-HR	80 km/0.01 hPa	95	750	699	760
MPI-ESM1-2-LR	80 km/0.01 hPa	47	2,149	1,538	2,282
INM CM5-0	59.6 km/0.2 hPa	73	632	592	668
CESM2	40 km/2.26 hPa	32	2,611	1,099	4,564
CESM2 WACCM*	130 km/4.5×10^{-6} hPa	70	1,535	1,099	1,611
UKESM1-0-LL*	85 km/0.005 hPa	85	1,194	641	1,369
HadGEM3-GC31-LL	85 km/0.005 hPa	85	1,194	641	1,369

Note. The vertical resolutions over different sections of the stratosphere ($dz1$, $dz2$, $dz3$) are calculated by dividing the vertical distance by the number of levels. Model levels, where given in pressure or sigma levels, were converted to height according to $z = -H \ln(p/p_o)$, with $H = 7,000$ m and $p_o = 1,013.25$ hPa. Models with interactive chemistry are denoted *.

models, and in CMIP5 and CMIP6, an increasing number of models relative to previous generations are showing a more realistic overall SSW frequency (Wu & Reichler, 2020). However, those studies which have considered splits and displacements separately (S16; Maycock & Hitchcock, 2015) indicate that the simulated frequency of displacement events tends to be more realistic, while splits are underrepresented, leading to a too-high ratio of displacements to splits in most models, compared with the observational record. S16 attributed this discrepancy to biases in models' stratospheric climatological state. The motivation for this study is to better understand the factors behind the differences in split and displacement SSW frequency that is often found in models as well as to assess whether the biases discussed above have persisted or been improved in the latest generation of models.

In this study, we examine a range of metrics representing SPV variability, comparing CMIP6 models with reanalysis data, and develop simple regression models to explain split and displacement SSW frequency. Data and methods for SSW classification are presented in Section 2, our results are discussed in Section 3, and a summary of our findings is provided in Section 4.

2. Data and Methods

We use daily data from the CMIP6 archive at the Centre for Environmental Data Analysis (CEDA). We use the preindustrial control simulation (piControl) of each available model, which has a prescribed or calculated CO_2 concentration, designed for the evaluation of unforced variability (Eyring et al., 2016). This provides a large number of SSW events, which will help to address the issue of internal variability over shorter historical simulations. Models used had at least 500 years of daily data available. For each model, the first 500 years were selected (this does not include the spin-up period). This reduces the data set to the 10 models summarized in Table 1. We recognize that there are many more models in CMIP6, but data were unavailable over the whole time period for other models at the time. However, our models are representative of the range of SSW frequencies simulated in CMIP6 models (See Wu & Reichler, 2020; their Figure 1b). Where spatial data were required, model data were regridded by horizontal linear interpolation to the coarsest horizontal resolution of the models (CanESM5).

Models are compared with the European Centre for Medium-Range Weather Forecasts reanalyzes; ERA-Interim (Dee et al., 2011) which covers 1979–2019 and the ERA5 reanalysis (Hersbach et al., 2020) which at present extends from 1979 to 2020. The eight vertical levels common to CMIP6 and ERA are used (10, 50, 100, 250, 500, 700, 850, and 1,000 hPa). We chose not to use ERA40 (Uppala et al., 2005) to extend the sample size of SSWs in the observational record. This is because in preliminary testing, certain metrics, such
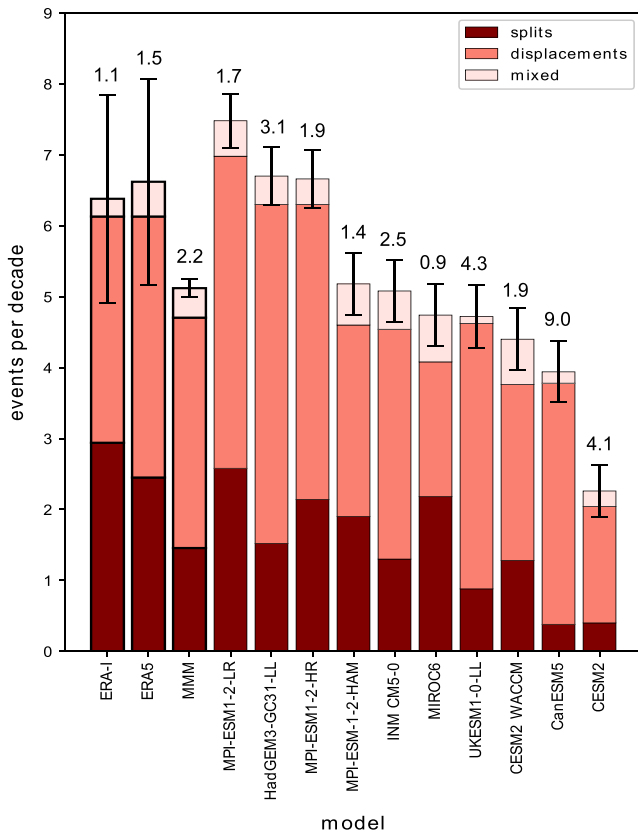
Figure 1. Frequency of split, displacement, and mixed SSW events in each of the CMIP6 models, compared with ERA. Error bars show the 95% confidence intervals for the frequency of all events. Confidence intervals are calculated for the population proportion using the binomial distribution. Reanalyzes and multimodel mean are distinguished by bold borders. Figure above each bar is the ratio of displacements to splits, to 1 decimal place. Models are arranged left to right from highest to lowest SSW frequency. CMIP6, sixth Coupled Model Intercomparison Project; MMM, multimodel mean; SSW, sudden stratospheric warming.

as daily mean zonal-mean zonal wind speed for the winter months, were found to have a very different distribution to those of ERA-I and ERA5.

To test whether the comparisons between CMIP6 piControl and ERA reanalysis are valid, we also repeat some aspects of the analysis using historical simulations from the same models. We use a single ensemble member (usually r1i1p1f1) from each historical simulation. By using the historical simulations (1850–2014), any discrepancy between piControl and reanalysis data that may occur due to anthropogenic or natural forcing can be identified.

SSW onset dates are identified by the reversal of zonal-mean zonal wind direction at 60°N and 10 hPa (Charlton & Polvani, 2007). Only SSWs with an onset date between December and March inclusive are included. Periods of wind reversal must be separated by at least 20 consecutive days of westerly winds to qualify as separate events. To distinguish split and displacement events, we use a modified version of the moment diagnostics method of S13 (Gerber et al., 2021). In this approach, two-dimensional moment diagnostics, M_{nm} of a distribution f(x,y) is used to classify the polar vortex as either a split or a displacement event. They are given in Cartesian coordinates by:

$$M_{nm} = \iint_s x^n y^m f(x,y) dx\, dy \tag{1}$$

where s is the surface over which integration occurs (the northern hemisphere) and n and m denote the order of the moment in the x and y directions, respectively. When applied to the SPV these moments describe the longitude-latitude shape of the vortex in terms of an equivalent ellipse (Waugh, 1997). The two moments used here are the latitude of the vortex centroid (a first-order moment) and the aspect ratio (the ratio of the major to minor axes; a second-order moment). To calculate the moment diagnostics, it is necessary to define the vortex edge. This is done for each model individually, and is the contour of the mean December-January-February-March (DJFM) zonal-mean geopotential height (GPH) at 60°N, 10 hPa. For further details of the calculations of moment diagnostics, see Waugh (1997), Matthewman et al. (2009), and S13.

The latitude of the vortex centroid is used to diagnose displacement events, and the aspect ratio defines split events. The S13 method uses 10 hPa GPH to define the vortex moments. Here, following Gerber et al. (2021), a hybrid approach is used, combining the SSW onset date identification of Charlton and Polvani (2007), and the S13 moment diagnostics based on 10 hPa GPH, to classify SSWs. The aspect ratio and centroid latitude of the SPV are calculated for each day in a window of 10 days either side of the onset date. Thresholds for the centroid latitude and aspect ratio are set and the total number of days when each threshold is exceeded in the window is calculated. The thresholds are defined as when the centroid latitude is equatorward of 66°N for a displacement event, and when the aspect ratio exceeds 2.4 for a split event (S13). The SSW is defined according to which threshold is exceeded with the greatest frequency in the 21-day window (Gerber et al., 2021). In some cases, SSW events cannot be clearly defined, containing equal elements of both splits and displacements, or neither threshold is exceeded, and in these cases the event is categorized as a mixed event.

Compared with the Charlton and Polvani (2007) classification, 10 events in reanalysis are reclassified; 4 change from displacement to split, 5 from split to displacement, and 1 event is a mixed event. S13 identifies a number of different events not included in Charlton and Polvani (2007), but none of the events identified using the hybrid method here change classification if the unmodified S13 method is applied.

ERA5 identifies one additional event (SSW onset 2002-01-17), and reclassifies two events compared with ERA-I. A split event occurring on 2001-12-30 was reclassified as a displacement in ERA5, and an additional split event on 21-02-1989 was identified as a mixed event. These changes may occur with very small changes in field values and changes in the vortex edge contour, if the event is a more marginal example of an SSW type.

We calculate the modal (i.e., most likely) climatological value of the daily DJFM aspect ratio and centroid latitude for each model, over the whole 500-year period, and for reanalyzes. We use the mode rather than the mean as it will not be influenced by the impact of extreme values on the distribution, such as those associated with SSW events. To calculate the mode, we follow Mitchell et al. (2011) and S13 and define the mode as the peak of the probability distribution fitted to the daily distribution of moment diagnostics for each model. Selecting the mode as the maximum value of a histogram can be problematic as it is sensitive to bin size. We therefore fit a Generalised Extreme Value Distribution (GEV) to the aspect ratio, of the form:

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

$$a = 1 + \xi\frac{x-\mu}{\sigma} \tag{3}$$

with parameters μ the location parameter, σ the scale parameter, and ξ the shape parameter, determined by maximum-likelihood estimation (Wilks, 2011). For the centroid latitude, we transform the data to better fit a Gaussian distribution, by cubing the latitudes (e.g., Mitchell et al., 2011). The cube root of the mode is then taken to return to the original value. In each case, the moment value at the peak of the distribution is taken as the mode. Alternative methods such as fitting a kernel density estimate to the distribution provide very similar results, as does using the location parameter of the GEV for aspect ratio.

Unless otherwise stated in the results, statistical significance is judged at the 95% level. This means that for the 10 models here, a significant correlation will be one where the absolute value of the correlation coefficient exceeds 0.63.

3. Results

3.1. Model Intercomparison of SSW Frequencies

The decadal frequencies of split, displacement and mixed events are shown for each CMIP6 model and compared with the frequencies from the ERA reanalyzes and the multimodel mean (Figure 1). In this study, most results refer to the differences between models, but a multimodel mean (MMM) is also calculated from the total number of events over all models as in S16, so overall each SSW event has an equal weighting, irrespective of the model from which it originates. There is a wide range of frequencies amongst models, together with different ratios of displacements to splits (hereafter DS ratios). In reanalyzes, the DS ratios are 1.1 in ERA-I and 1.5 in ERA5. Differences are due to the reclassifying of an event between reanalyzes. However, only MIROC6 has roughly equal numbers of displacements and splits (DS ratio of 0.87); all other models exhibit a larger proportion of displacements, ranging from a DS ratio of 1.42 (MPI-ESM-1-2-HAM) to 8.95 (CanESM5) (Figure 1). Notably, the two models with the lowest lid heights (CanESM5 and CESM2) have the great DS ratios, the lowest split frequency, and the lowest overall frequency of SSWs in agreement with results from CMIP5 (e.g., Charlton-Perez et al., 2013). Overall, 9 out of 10 models have a split frequency below that of reanalysis. Similarly, in all but two of the CMIP5 models in S16 (11 out of 13), split frequency is below that in reanalysis; therefore, this bias is consistent.

70% of the individual model confidence intervals for total SSW frequency overlap with those of the reanalyzes, broadly in line with the analysis of Wu and Reichler (2020), who used ERA40 and historical simulations. The historical simulation SSW frequencies and DS ratios are close to those of piControl (Figure S1); the correlation between piControl and historical simulation total SSW frequency is 0.89 across the 10 models, and the correlation between the model DS ratio of the piControl and historical simulations is 0.93 (not shown).
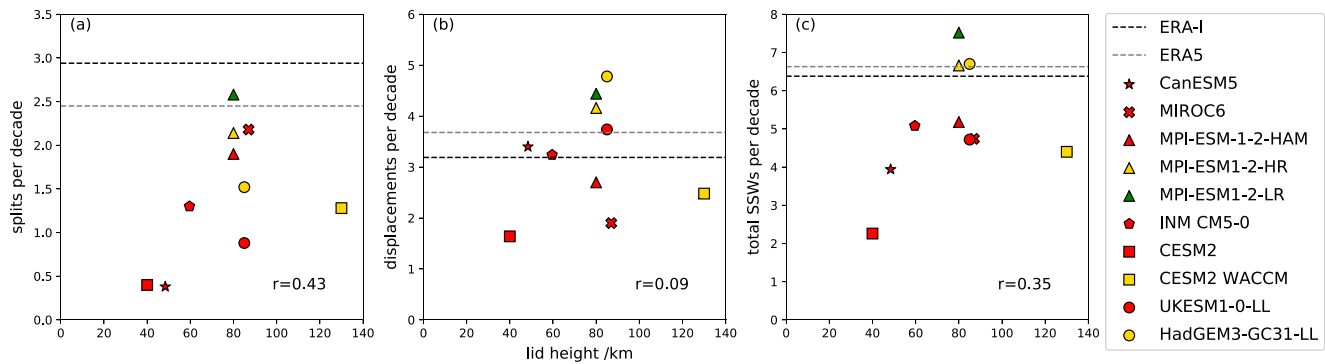
Figure 2. Model lid height and (a) split SSW frequency, (b) displacement SSW frequency, and (c) total SSW frequency. Dashed horizontal lines show frequency for reanalyzes. In this and subsequent scatterplots, models from the same "family" have the same marker shape. Significant correlations are shown in bold. SSW, sudden stratospheric warming.

Models that are structurally similar, for example, with the same dynamical cores, can produce large differences in SSW frequency, for example, HadGEM3 and UKESM, and MPI-ESM1-2-LR and MPI-ESM-1-2-HAM (Figure 1), however in these instances, the interactive atmospheric chemistry may be an important factor. In each pair, the model with interactive chemistry (UKESM and MPI-ESM-1-2-HAM) simulates fewer SSWs, and the confidence intervals of the pair do not overlap (Figure 1). Haase and Matthes (2019) find that a model with interactive chemistry has a stronger and colder SPV and produces fewer SSWs than the equivalent model without interactive chemistry. In agreement with this result, in each of the pairs (HadGEM3 and UKESM; MPI-ESM1-2_LR, MPI-ESM-1-2-HAM), the model with interactive chemistry has stronger stratospheric zonal winds at 60°N, compared with the model without interactive chemistry. Hence, the greater total SSW frequency in models such as HadGEM3 and MPI-ESM1-2-LR may be related to the absence of interactive chemistry in these models. In addition, CESM2 WACCM includes interactive chemistry and has a lower overall SSW frequency, albeit with confidence intervals that overlap those of reanalyzes, although the low lid height of CESM2 confounds any direct comparison based on interactive chemistry alone.

Comparing with CMIP5 models in S16, the CanESM model also had a high DS ratio, but has changed from having one of the highest frequencies of SSWs in CMIP5 (CanESM2) to one of the lowest SSW frequencies in CMPI6 (CanESM5). HadGEM and MPI-ESM-LR models have some of the highest total SSW frequencies in CMIP5 and CMIP6, while total SSW frequency for MIROC is relatively low for both (CMIP5: ~4.5/decade, CMIP6: 4.8/decade). In each of these cases, the DS ratio is broadly similar in CMIP5 and CMIP6.

3.2. Model Lid Height and Stratospheric Resolution and SSW Frequencies

Model lid height has been identified as a contributory factor to underrepresentation of SSWs (e.g., Charlton-Perez et al., 2013; Shaw & Perlwitz, 2010). However, there is no significant correlation of lid height with either the modal aspect ratio or centroid latitude, which could in turn impact on SSW frequency. Figure 2 shows the variation of SSW frequency with lid height for splits, displacements, and total SSWs. Of interest is the large number (six) of models with a similar lid height (80–87 km), which have a wide range in split event frequency (0.88–2.58 per decade, Figure 2a). Similarly, for displacement events, the same six models have a frequency spread of 1.9–4.78 per decade (Figure 2b). Also of note is that the model with the highest lid (CESM2 WACCM), only has a mid-range frequency of splits, displacements, and total SSWs. Although the two low-lid models are distinct as also having a low frequency of splits, CanESM5 actually has a greater number of displacement events than four of the higher lid models. While low-top models (i.e., those with a top below the stratopause) have a lower overall frequency of SSWs (Figure 2c), mainly as a result of the large underrepresentation of split events (Figure 2a), once the model-top height exceeds 70 km, there is no discernible relationship between lid height and frequency.

We have also assessed the relationship between vertical resolution in the stratosphere and SSW frequency. The resolution was calculated for different levels of the stratosphere (Table 1). In the models evaluated here, no significant relationships were identified between stratospheric vertical resolution at the different

Table 2
Pearson's Correlation Coefficients Between Different Atmospheric Parameters, and With Split, Displacement and Total SSW Frequency

Atmos. variables	SSW type Split disp total	Atmospheric variables Centroid latitude	Aspect ratio	U10_60	VS_10	U100_60	VS_100	HF100 total	HF100 W1	HF100 W2
Centroid latitude	−0.34 **−0.97 −0.82**									
Aspect ratio	**0.78** 0.59 **0.86**	**−0.66**								
U10_60	**−0.76** −0.54 **−0.83**	0.57	**−0.87**							
VS_10	**0.80 0.66 0.90**	**−0.64**	**0.85**	**−0.74**						
U100_60	**−0.82** 0.04 −0.47	0.00	−0.46	**0.63**	−0.50					
VS_100	**0.86** 0.18 0.62	0.28	0.55	**−0.64**	**0.72**	**−0.91**				
HF100 total	0.06 **0.77** 0.51	**−0.83**	0.37	−0.15	0.41	0.33	−0.11			
HF100 W1	−0.19 **0.65** 0.29	**−0.69**	0.08	0.12	0.26	0.47	−0.24	**0.92**		
HF100 W2	0.28 0.24 0.30	−0.29	0.51	−0.27	0.22	0.18	−0.11	0.37	0.06	

Note. Significant correlations ($p < 0.05$) are shown in bold.

levels and the numbers of split, displacement, or total SSWs, in agreement with results from CMIP5 (S16). Wu and Reichler (2020) identified a modest negative correlation at a lower (90%) significance threshold between vertical resolution and total SSW frequency (a coarser resolution is associated with fewer SSWs). Here, correlations are lower overall, although the correlation between 1 and 50 hPa resolution and total SSWs is just short of the 90% significance threshold ($r = −0.52$). This difference may be a result of the small sample size and different models in the sample. Considering specific models, INM has the highest resolution for each stratospheric height range (Table 1) and has a very similar total SSW frequency to the MMM (Figure 1), although the DS ratio is quite high (2.5:1). Conversely, the low-resolution MPI-ESM-LR, and the chemistry-climate version (MPI-ESM-HAM) have low stratospheric vertical resolution and yet simulate quite different total SSW frequencies (7.52 and 5.18 per decade), perhaps due to the interactive chemistry in the latter. However, both have more realistic DS ratios (1.7:1 and 1.4:1 respectively) than models with a higher resolution. CESM2 is an outlier, with a reduced number of SSWs of all types (Figure 1), which may in part be influenced by the low lid height (below the stratopause) and coarse vertical resolution. This is in agreement with studies involving CMIP5 models (e.g., Haase et al., 2018). Beyond this, however, it is difficult to distinguish between the stratosphere-resolving high-top models in terms of a clear impact of vertical resolution and/or lid height on SSW frequency. There are clearly more important factors than vertical resolution and model lid height that influence the number and type of SSWs simulated. Hence, a lid height above 70 km and a well-resolved stratosphere may be regarded as necessary but not sufficient for the accurate representation of SSW frequency.

3.3. SSW Frequency and Atmospheric State

We examine aspects of the SPV climatological state that we hypothesize may be linked to the variation in frequency of split and displacement events across the models, as in CMIP5 (S16). Correlations between the various atmospheric parameters and SSW frequency are summarized in Table 2. There are significant correlations between modal centroid latitude and displacement frequency and aspect ratio and split frequency in the models (Figure 3), in agreement with results from CMIP5 (S16). In the case of splits, the correlation is significant and positive ($r = 0.78$) and increases to 0.83 if the outlier CESM2 model is removed; models with a higher climatological aspect ratio show a greater frequency of splits. Conversely, models with a lower climatological centroid latitude show increased frequency of displacements ($r = −0.97$). There is also a significant negative correlation ($r = −0.66$) between modal aspect ratio and modal centroid latitude, so a vortex that is more preconditioned to displacement events will simulate fewer splits, and vice-versa. Using the historical simulations, the equivalent correlations are 0.86 (splits) and −0.91 (displacements) (Figure S2). While for splits, the low top models clearly have small climatological aspect ratio and CESM2 is a clear outlier (Figure 3a), this is not the case for displacements (Figure 3b), where CanESM5 is in the mid-range of modal centroid latitudes. The intermodel climate aspect ratio variability is significantly negatively
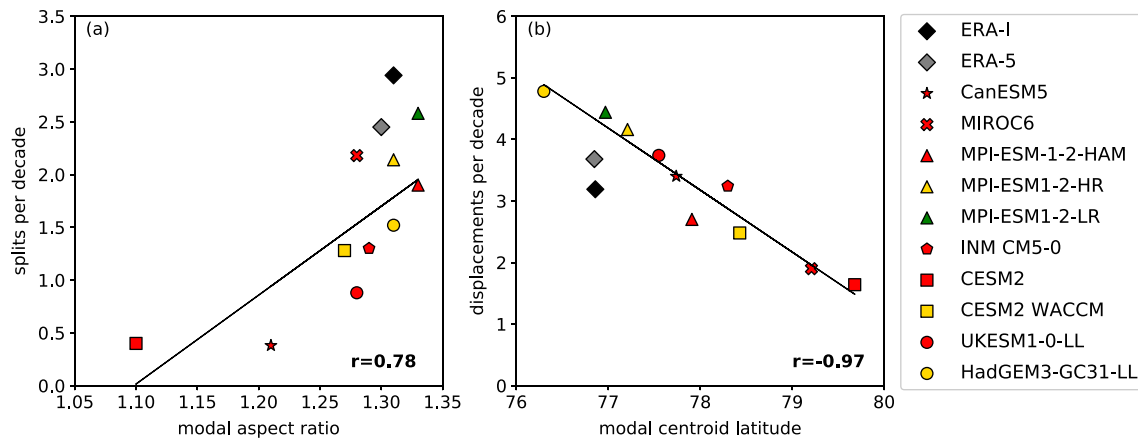
Figure 3. Modal DJFM aspect ratio and split SSWs per decade (a), and modal DJFM centroid latitude and displacement SSWs per decade (b) for CMIP6 models, compared with ERA-I and ERA5. Lines are best-fit regression lines for CMIP6 models only. CMIP6, sixth Coupled Model Intercomparison Project; SSW, sudden stratospheric warming.

correlated ($r = -0.72$, not shown) with stratospheric resolution in the upper stratosphere ($dz3$, Table 2), but there is no strong relationship between stratospheric resolution and centroid latitude. This is in agreement with S16, and may be related to the extent to which waves in the lower stratosphere can reach the mid-stratosphere.

Unlike CMIP5 (S16), the reanalysis values lie a little way off the best fit line of the models. For splits, ERA-I and ERA5 both lie above the line (models show a low SSW frequency bias for a given aspect ratio similar to reanalysis), and for displacements they both lie below the line (i.e., the models are biased high for displacement frequency for an aspect ratio close to that in reanalysis). However, this is likely to be within the range of sampling error, given the small number of events in reanalyzes and if CESM2 is excluded from the split frequency regression, the reanalyzes are closer to the best fit line. The difference in location of ERA-I and ERA5 in Figure 3 is due to ERA5 identifying one additional event and the reclassifying of two events (see Section 2 for details). Six of the 10 models examined here have a vortex that is too axially symmetric, compared to reanalysis, and the modal centroid latitude is higher than the reanalysis for 9 of the 10 models. As with S16, Figure 3 indicates that biases in the climatological vortex state can account for much of the inter-model variability in SSW representation and therefore that a more accurate representation of the models' average vortex state would improve SSW frequency. When comparing with results in S16 which use CMIP5, there has been no clear improvement in the representation of SPV geometry with the development of CMIP6.

Representative examples of the SPV, shown as the DJFM mean of 10 hPa GPH, are initially difficult to distinguish from each other by eye (Figures 4a–4d). However, differences from the ERA-I SPV are distinct for the different models (Figures 4e–4g). HadGEM3 has relatively small differences, significantly lower (up to 130 m) 10 hPa GPH over northern Scandinavia and Russia, and a significantly higher surface over the North Pacific and North America (Figure 4e). By contrast, for MPI-ESM-HAM, the lower heights (up to 300 m) compared with ERA-I occur over the Beaufort-Chukchi Seas, and differences are significant across the whole of the polar cap (Figure 4f). CESM2 shows the most marked difference from ERA-I, with zonally symmetric 10 hPa GPH differences, significantly lower than ERA-I over the Arctic Ocean, but significantly higher in the midlatitudes, indicating a stronger SPV (Figure 4g). The morphological distinctions are also emphasized in Figure 4h, where the contour marking the vortex edge is shown, along with the centroid latitude. HadGEM3 is more clearly displaced from the pole with a center just east of Svalbard, while CESM2 is the most circularly symmetric vortex, with a centroid nearest the North Pole. CESM2 is a clear outlier among the CMIP6 models, and the only one with a lid below 1 hPa.

Next, we examine a possible cause of the low frequency of splits in most models. With a strong SPV, zonal winds are biased high. Through Charney-Drazin filtering, zonal wavenumber 2 planetary waves, which are associated with split vortex events, may not be able to propagate into the stratosphere if lower stratospheric
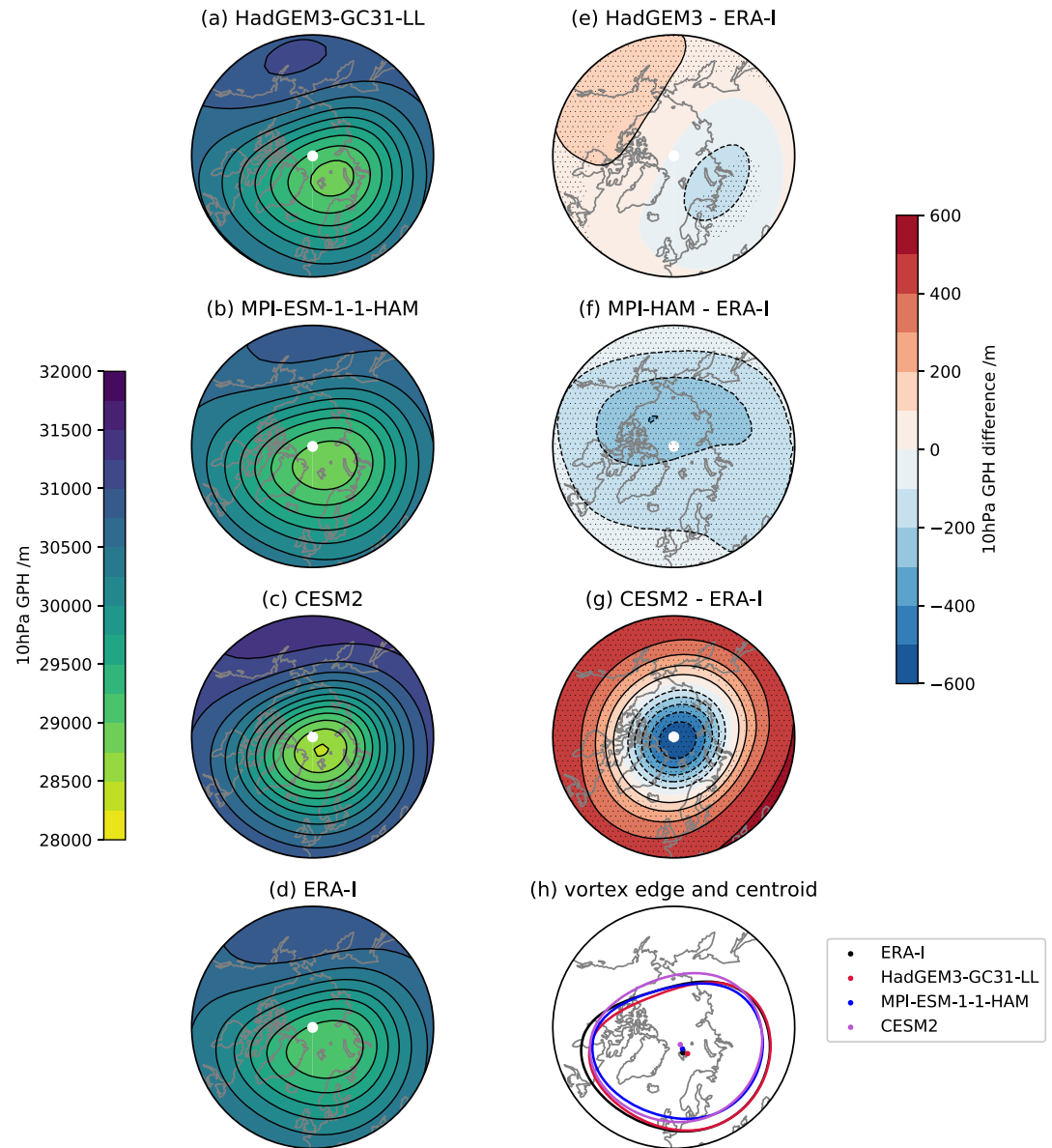
Figure 4. Stratospheric polar vortices from a representative range of models and ERA-I. Three models chosen are from different model families and are: CESM2 (lowest aspect ratio, highest centroid latitude), HadGEM3-GC31-LL (one of the highest aspect ratios, lowest latitude centroid), and MPI-ESM-1-1-HAM (high aspect ratio, mid-range latitude centroid). (a)–(d) 10 hPa GPH, (e)–(g) difference between model 10 hPa GPH and the ERA-I 10 hPa GPH (f), the vortex edges (mean 10 hPa GPH, 60°N DJFM), and centroid latitudes of the vortices. Negative contours are dotted and the zero contour is omitted. Stippling on (e)–(g) shows areas where height difference is significant ($p < 0.01$, estimated using a permutation test with 10,000 resamplings of shuffled seasonal mean data).

zonal winds are too strong. We compare DJFM-mean zonal-mean zonal wind (ZMZW) at 100 hPa, 60°N (u100_60), which we use as a proxy for mean lower stratospheric SPV strength, with the frequency of split and displacement events. As the daily mean ZMZW distribution may be influenced by SSW frequency, we also examine the relationship between the median and the modal u100_60, where the mode is determined by fitting a kernel density estimate (kde), as these measures are robust to the influence of extreme values, however, the results are qualitatively similar. The distributions of daily DJFM u100_60 vary, with skews ranging from slightly positive to slightly negative (Figure S3). Therefore our results agree with Wu and Reichler (2020) that u100_60 contains independent information concerning SSW frequencies and use the mean values of u100_60.
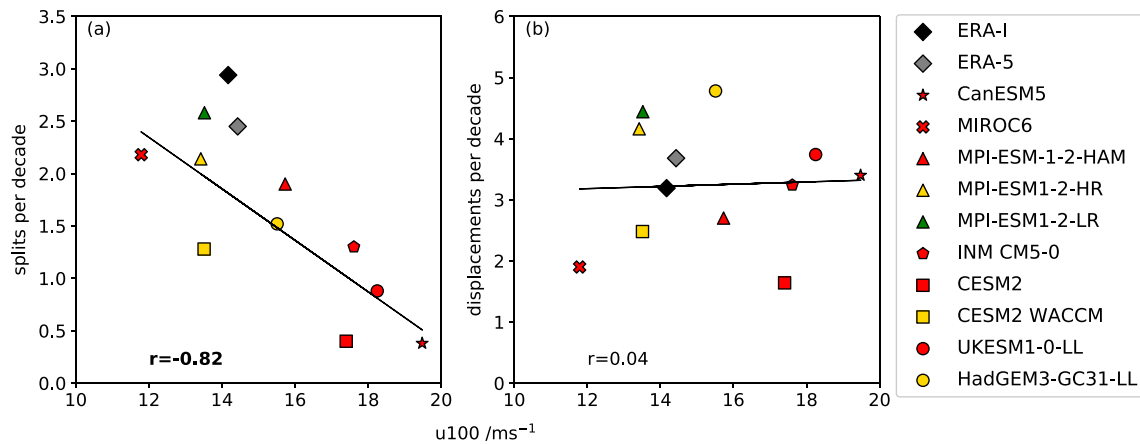
Figure 5. DJFM zonal-mean zonal wind at 100 hPa, 60°N (u100), against frequency of (a) split and (b) displacement SSWs per decade, with linear regression best-fit line for CMIP6 models. Significant correlations are shown in bold. CMIP6, sixth Coupled Model Intercomparison Project; SSW, sudden stratospheric warming.

Models with higher mean u100_60 have fewer splits (Figure 5a; $r = -0.82$). The overall low bias of model split frequency for a given mean u100_60 compared with reanalysis is again evident. However, there is no similar association between displacement frequency and u100_60 ($r = 0.04$, Figure 5b). Results are almost identical for the historical simulations (split frequency, $r = -0.72$, displacement frequency; $r = -0.08$, not shown). For a number of models (CESM2, CanESM5, UKESM, and INM), there is a more pronounced shoulder in the ZMZW around 60°N, with CESM2 even showing a local maximum at this latitude (Figure 6). This is related to a too-strong representation of the SPV, with stronger ZMZW extending downward to lower levels. The same group of four models also have high-biased ZMZW at 10 hPa (Figure S4). The differences in ZMZW at all levels between MPI-ESM-HAM, CanESM5, and ERA-I illustrate this (Figure 7). INM, UKESM, and CESM2 all show a similar difference to that between CanESM5 and ERA-I (Figure 7a), with a too-strong polar vortex throughout the stratosphere. Models with a representation of SPV strength closer to reanalysis tend to show a weaker ZMZW bias in the lower stratosphere in the mid-latitudes (MPI-ESM-LR, HadGEM3, MPI-ESM-HAM), typified by Figure 7b. The high bias in lower SPV strength in some models is not sufficient to prevent the vertical propagation of zonal wavenumber 1 planetary waves into the stratosphere, as there is no association between lower SPV strength and displacement frequency. However, the high-biased models may have sufficiently strong winds in the lower SPV to filter out wavenumber 2 propagation, hence the clear negative association between split frequency and mean u100_60. The
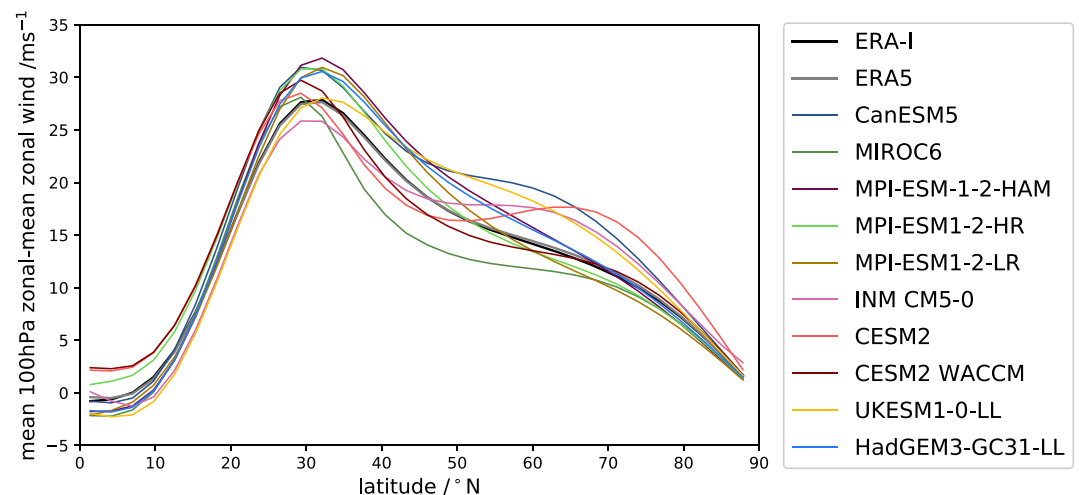

Figure 6. Variation with latitude of DJFM mean zonal-mean zonal wind at 100 hPa for CMIP6 models and ERA reanalyzes. CMIP6, sixth Coupled Model Intercomparison Project.
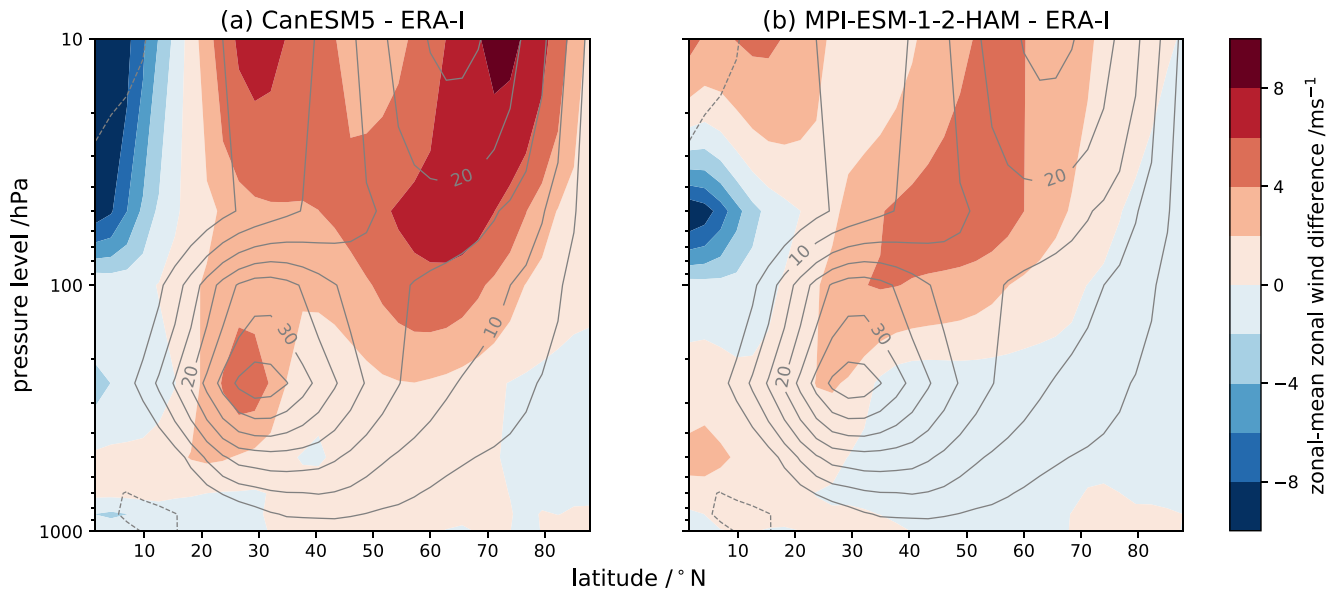
Figure 7. DJFM mean zonal-mean zonal wind difference between (a) CanESM5 and ERA-I, and (b) MPI-ESM-1-2-HAM and ERA-I (shading). Contours show ERA-I DJFM climatology, contours at 5 ms^{-1} intervals, dotted negative, zero contour is omitted.

historical simulations show an almost identical ZMZW as a function of latitude, with the same models biased high at 60°N (Figure S5).

To further investigate the relationship between SPV variability and SSW frequency, we adopt the metric used with idealized models by Wu and Reichler (2019). This is the ratio of standard deviation to the mean of the ZMZW at 60 N, 10 hPa, which can be conceived of as the ratio of vortex variability to strength (VS ratio), and previous studies have shown this to be related to total SSW frequency (Horan & Reichler, 2017). Such vortex variability has been shown to be related to eddy wave-driving (e.g., Shaw et al., 2014). Here, we extend the use of this metric to both 10 hPa (u10_60) and 100 hPa (u100_60) ZMZW. As u10_60 is always more negatively skewed than u100_60 (Figure S6), we also investigated the use of nonparametric indicators of dispersion and location, testing ratios of interquartile range (IQR) to median and IQR to mode. However, the use of the different metrics makes no difference to the significance of the different associations.

There are strong positive associations between the vortex VS ratio and split SSW frequency, at both 10 hPa and 100 hPa, irrespective of the method used to construct the ratio (Figures 8a and 8d). At 10 hPa, the correlation with displacement frequency is only just significant (0.66) compared with that for splits ($r = 0.80$; Figure 8b), and thus the greater part of the correlation for total SSW frequency ($r = 0.90$; Figure 8c) is derived from the association with split events. However, at 100 hPa, there is no significant correlation of the VS ratio with either displacement or total SSW frequency (0.18 and 0.62 respectively; Figures 8e and 8f), while the correlation between split SSW frequency and 100 hPa VS ratio is 0.86.

These results indicate that the variability and strength of the SPV are important factors in determining the frequency of split SSW events. The spread (standard deviation) of the reanalysis daily ZMZW is markedly greater than that of all the models, both at 10 hPa and 100 hPa; however, six (seven) models have greater mean or median ZMZW at 100 (10 hPa) than reanalyzes. This suggests that the models, while capable of simulating mean ZMZW, are underestimating internal variability, which in turn may be due to a deficiency in wave driving, in agreement with Wu and Reichler (2020) (see their Figure 2).

To examine wave driving, we use the DJFM mean daily climatological meridional eddy heat flux, $\overline{v'T'}$ averaged over 45°N –75°N at 100 hPa (Polvani & Waugh, 2004), where v is meridional wind, T is temperature, the primes denote departures from the zonal mean and the overbar denotes the zonal mean. This acts as a measure of the planetary wave activity propagating from the troposphere to the stratosphere (e.g., Newman & Nash, 2001). Meridional eddy heat flux has a significant correlation with displacement frequency ($r = 0.77$)
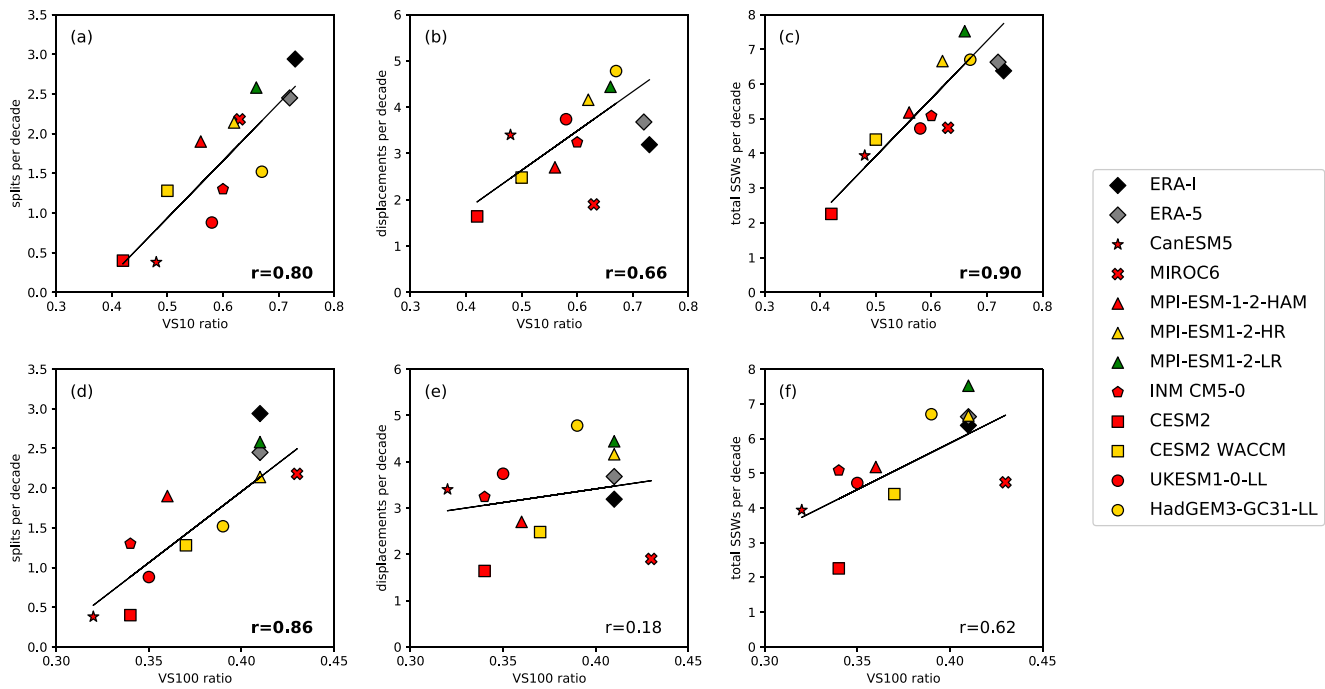
Figure 8. VS ratio (ratio of standard deviation to mean daily ZMZW), plotted against frequencies of splits, displacements, and all events, at 60°N 10 hPa (a)–(c) and 100 hPa (d)–(f).

but no correlation with split frequency ($r = 0.06$) (Figure 9) consistent with Wu and Reichler (2020), who find a significant correlation ($r = 0.6$) between vertical Eliassen-Palm flux and total SSW frequency. Meridional eddy heat flux is also correlated with centroid latitude ($r = -0.83$; Table 2). We also decomposed the meridional eddy heat flux into zonal wavenumber 1 and wavenumber 2 components. The only significant correlation between wavenumbers 1 and 2 components of eddy meridional heat flux and SSW frequency was between zonal wavenumber1 eddy heat flux and displacement frequency ($r = 0.65$; Table 2). This suggests that planetary wave driving may be more significant for the formation of displacement SSWs, whereas split SSWs are strongly associated with vortex morphology and wave filtering in the lower stratosphere, but not planetary wave driving. This result is consistent with Albers and Birner (2014), who find that split SSWs are caused by internal resonance and that precursor wave driving is not sufficient for SSW development. Wavenumber 1 eddy heat flux also showed a significant correlation with the modal centroid latitude
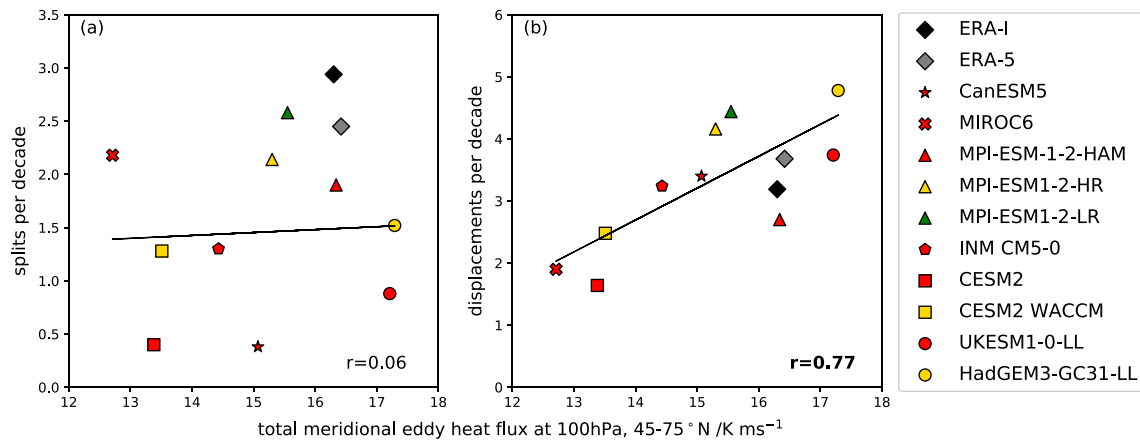

Figure 9. Total eddy meridional heat flux at 100 hPa, 45°N–75°N and (a) split SSW frequency and (b) displacement SSW frequency. Best-fit linear regression lines are shown for the CMIP6 models. Significant correlations are shown in bold. CMIP6, sixth Coupled Model Intercomparison Project; SSW, sudden stratospheric warming.

Table 3
Summary of Regression Models

	Model 1	Model 2	Model 3
	Split frequency	Displacement frequency	Total SSW frequency
y intercept (SSW frequency per decade)	−10.00	85.59	80.00
Input variables	**Model coefficients**		
u100_60/ms^{-1}	−0.15		
Aspect ratio	10.74		
Latitude centroid/°		−1.06	−0.87
u10_60/ms^{-1}			−0.26
Model statistics			
R^2	0.91	0.91	0.82
Adjusted R^2	0.88	0.90	0.76
Mean squared error	0.06	0.09	0.36

($r = -0.69$; Table 2). Increased wavenumber 1 wave driving in models is likely to contribute to a vortex that is displaced further from the pole, and which therefore may be more prone to displacement SSW events.

3.4. Regression Analysis

To combine these results, we use multiple linear regression to construct statistical models of the identified associations. We use a forward selection method (e.g., Wilks, 2011), adding one predictor variable at a time, using the variables identified above that have strong associations with SSW frequency. We ensure that models are not overfitted using the adjusted R^2 value (e.g., Faraway, 2014) which is adjusted according to the number of model terms. This will only increase if an added predictor has some predictive value, whereas using R^2 alone can result in the model with the largest number of terms being chosen. We do not use interaction terms or polynomials, as from visual inspection associations are strongly linear. We omitted the CESM2 model from the regression process as it is a distinct outlier, with significant leverage and influence. This makes the regression models more internally consistent as CESM2 is the only model assessed that has a lid height below 1 hPa. Regression models are summarized in Table 3.

The models use wholly statistical properties of the polar vortex to explain SSW frequency. Split SSW frequency is explained by the strength of the zonal winds at 100 hPa, 60°N (u100_60), and the vortex aspect ratio. As the mean wind strength at 100 hPa increases, so the split frequency is reduced and vice-versa. This is consistent with Charney-Drazin filtering inhibiting the vertical propagation of tropospheric wavenumber 2 waves. In addition, it is harder to elongate and split a vortex that is more axially symmetric.

The association between modal centroid latitude and displacement frequency is so strong that no other terms are included in the model, although modal centroid latitude can be understood as incorporating wave driving components, as modal centroid latitude is associated with wavenumber1 $\overline{v'T'}$ ($r = -0.69$; Table 2) and total $\overline{v'T'}$ ($r = -0.83$; Table 2).

Total SSW frequency is explained by the ZMZW at 10 hPa (u10_60) and modal centroid latitude. u10_60 is selected as it is associated with split frequency ($r = -0.76$; Table 2), while modal centroid latitude is strongly associated with displacement frequency ($r = -0.97$; Table 2). Wu and Reichler (2019) found that total SSW frequency is related to vortex strength at 10 hPa and heat fluxes at 100 hPa. The model for total SSW occurrence frequency can be interpreted in these terms, as heat fluxes are incorporated into the centroid latitude term and vortex strength into the u10_60 term. In addition, however, here we extend the analysis and establish that split and displacement SSWs have different predictors. Applying the regression models above to ERA-I and ERA 5 results in an underprediction of split events (by 33.9% in ERA-I, 26.6% in ERA5),

and an overprediction of displacement events (29% in ERA-I, 12% in ERA5), confirming the bias in the climate models.

4. Summary

In this study, we have assessed SPV metrics for 10 CMIP6 models, along with the frequency of split and displacement SSW events and have compared them with ERA reanalysis data.

CMIP6 models have a large intermodel spread in SSW frequency, as well as some common biases. Specifically, models tend to underrepresent split events compared with reanalyzes. In comparison with CMIP5 models, while in some cases the frequency or displacement-split ratio has changed significantly between model generations, there is no clear overall improvement in the representation of polar vortex variability in CMIP6 compared with CMIP5 (S16).

We recognize that the range of models used is limited, but models are broadly representative of the range of CMIP6 models and several results are consistent with other studies (e.g., Wu & Reichler, 2020). The models used in this study span the range of SSW frequencies from 20 CMIP6 models in Wu and Reichler (2020). Our use of piControl enables us to analyze a greater number of SSW events. Although some models are closely related, having common components, there are often large differences in SSW frequency, DS ratio, and atmospheric parameters in these models. Some of these differences may be attributable to whether or not the models include an interactive chemistry component. Two models from the same family, MPI-ESM1-2-HR and MPI-ESM1-2-LR, produce similar results.

Historical and piControl simulations show similar results in terms of intermodel SSW frequency spread, and associations with vortex morphology. Any trend that may be present in historical simulations is therefore small compared to the intermodel variability.

As with CMIP5, there are strong associations between split SSW occurrence frequency and modal aspect ratio and between displacement SSW occurrence frequency and modal centroid latitude (S16). Models with a vortex that is too axially symmetric simulate fewer split SSW events, while a model SPV that is located further poleward will tend to have fewer displacement events.

The association between SSW frequency and model lid height and stratospheric vertical resolution is unclear. Low-top models (1 hPa and below) have fewer SSWs, but once the lid height is raised, and providing that there are sufficient levels to allow the representation of the stratospheric circulation, there is little gain achieved by raising the lid height or increasing vertical resolution. It is likely that other confounding variables may obscure the impact of vertical resolution, such as the presence of interactive atmospheric chemistry and the choice and tuning of gravity wave parameterizations.

Wu and Reichler (2019) showed that the strength and variability of the SPV are associated with SSW frequency. Here, we show that at 100 hPa this is only true of split events, but at 10 hPa, there is a stronger association with displacement and total SSW frequency. This may relate to Charney-Drazin filtering of higher wavenumbers propagating upwards from the troposphere, and hence those with stronger SPV in the lower stratosphere have lower frequencies of split events, which have been associated with wavenumber 2 disturbances.

The relative frequencies of split and displacement SSWs can be explained using different predictors in linear regression models. Displacement frequency is almost entirely explained by the modal centroid latitude of the model, which is largely influenced by stratospheric wave driving. However, the frequency of split events is explained by both the modal climatological aspect ratio and the strength of the SPV at 100 hPa.

Our results have related split and displacement SSW frequency to climatological aspects of the SPV geometry and lower stratospheric zonal winds. In doing so, we hope to provide a focus for future modeling development aiming to reduce the longstanding biases in simulated SSW frequency. This, in turn, would help to improve the model representation of high-latitude stratospheric, tropospheric, and surface climate variability.

This study can be extended by investigating purely physical factors associated with SSW frequency, which in turn contribute to the statistical descriptors examined here. For example, Wu and Reichler (2020) identified tropopause and stratospheric temperature as significant factors, and the refractive index of the lower stratosphere is significantly linked with upward wave activity flux. However, they conclude that mean vortex strength is the most reliable predictor of SSW frequency, although the reasons for the large intermodel spread are unclear, but may be related to gravity wave parameterization.

Data Availability Statement

ERA-I data are freely available from the ECMWF website (https://www.ecmwf.int/) and ERA5 data are available from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). CMIP6 data were accessed at the Centre for Environmental Analysis (CEDA) archive at JASMIN (https://www.jasmin.ac.uk/) but are also freely available from the Earth System Grid Federation (ESGF) (https://esgf-index1.ceda.ac.uk/projects/esgf-ceda/). Python code is available at https://doi.org/10.5281/zenodo.4817437

References

Albers, J. R., & Birner, T. (2014). Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings. *Journal of the Atmospheric Sciences*, *71*(11), 4028–4054. https://doi.org/10.1175/JAS-D-14-0026.1

Ayarzagüena, B., Charlton-Perez, A. J., Butler, A. H., Hitchcock, P., Simpson, I. R., Polvani, L. M., et al. (2020). Uncertainty in response of sudden stratospheric warmings and stratosphere-troposphere coupling to quadrupled CO_2 concentrations in CMIP6 models. *Journal of Geophysical Research: Atmospheres*, *125*(6), e2019JD032345. https://doi.org/10.1029/2019JD032345

Ayarzagüena, B., Polvani, L. M., Langematz, U., Akiyoshi, H., Bekki, S., Butchart, N., et al. (2018). No robust evidence of future changes in major stratospheric sudden warmings: A multi-model assessment from CCMI. *Atmospheric Chemistry and Physics*, *18*(15), 11277–11287. https://doi.org/10.5194/acp-18-11277-2018

Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., et al. (2021). Sudden stratospheric warmings. *Reviews of Geophysics*, *59*(1), e2020RG000708. https://doi.org/10.1029/2020RG000708

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

Butler, A. H., Sjoberg, J. P., Seidel, D. J., & Rosenlof, K. H. (2017). A sudden stratospheric warming compendium. *Earth System Science Data*, *9*(1), 63–76. https://doi.org/10.5194/essd-9-63-2017

Cagnazzo, C., & Manzini, E. (2009). Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European Region. *Journal of Climate*, *22*(5), 1223–1238. https://doi.org/10.1175/2008JCLI2549.1

Charlton, A. J., & Polvani, L. M. (2007). A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *Journal of Climate*, *20*(3), 449–469. https://doi.org/10.1175/jcli3996.1

Charlton, A. J., Polvani, L. M., Perlwitz, J., Sassi, F., Manzini, E., Shibata, K., et al. (2007). A new look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations. *Journal of Climate*, *20*(3), 470–488. https://doi.org/10.1175/JCLI3994.1

Charlton-Perez, A. J., Baldwin, M. P., Birner, T., Black, R. X., Butler, A. H., Calvo, N., et al. (2013). On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *Journal of Geophysical Research D: Atmospheres*, *118*(6), 2494–2505. https://doi.org/10.1002/jgrd.50125

Charney, J. G., & Drazin, P. G. (1961). Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *Journal of Geophysical Research*, *66*, 83–109. https://doi.org/10.1029/jz066i001p00083

Cohen, J., & Jones, J. (2011). Tropospheric precursors and stratospheric warmings. *Journal of Climate*, *24*(24), 6562–6572. https://doi.org/10.1175/2011JCLI4160.1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Esler, J. G., & Scott, R. K. (2005). Excitation of transient Rossby waves on the stratospheric polar vortex and the barotropic sudden warming. *Journal of the Atmospheric Sciences*, *62*(10), 3661–3682. https://doi.org/10.1175/JAS3557.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., et al. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Faraway, J. J. (2014). *Linear models with R* (2nd ed., p. 274). CRC Press.

Gerber, E. P., Martineau, P., Ayarzaguena, B., Barriopedro, D., Bracegirdle, T. J., Butler, A. H., et al. (2021). Extratropical stratosphere-troposphere coupling. In M. Fujiwara, G. L. Manney, L. Gray, & J. S. Wright (Eds.), *Stratosphere-troposphere processes and their role in climate (SPARC) reanalysis intercomparison project (S-RIP)*. SPARC. www.sparc-climate.org/publications/sparc-reports (in press)

Haase, S., & Matthes, K. (2019). The importance of interactive chemistry for stratosphere-troposphere coupling. *Atmospheric Chemistry and Physics*, *19*(5), 3417–3432. https://doi.org/10.5194/acp-193417-2019

Haase, S., Matthes, K., Latif, M., & Omrani, N.-E. (2018). The importance of a properly represented stratosphere for northern hemisphere surface variability in the atmosphere and the ocean. *Journal of Climate*, *31*(20), 8481–8497. https://doi.org/10.1175/JCLI-D-17-0520.1

Hall, R. J., Mitchell, D. M., Seviour, W. J. M., & Wright, C. J. (2021). Tracking the stratosphere-to-surface impact of sudden stratospheric warmings. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD03381. https://doi.org/10.1029/2020JD033881

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *The Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Horan, M. F., & Reichler, T. (2017). Modeling seasonal sudden stratospheric warming climatology based on polar vortex statistics. *Journal of Climate*, *30*(24), 10101–10116. https://doi.org/10.1175/JCLI-d-17-0257.1

Jucker, M., Fueglistaler, S., & Vallis, G. K. (2014). Stratospheric sudden warmings in an idealized GCM. *Journal of Geophysical Research: Atmospheres*, *119*(19), 11054–11064. https://doi.org/10.1002/2014JD022170

Kim, J., Son, S.-W., Gerber, E. P., & Park, H.-S. (2017). Defining sudden stratospheric warming in climate models: Accounting for bias in model climatologies. *Journal of Climate*, *30*(14), 5529–5546. https://doi.org/10.1175/JCLI-D-16-0465.1

King, A. D., Butler, A. H., Jucker, M., Earl, N. O., & Rudeva, I. (2019). Observed relationships between sudden stratospheric warmings and European climate extremes. *Journal of Geophysical Research: Atmospheres*, *124*(24), 13943–13961. https://doi.org/10.1029/2019JD030480

Kolstad, E. W., Breiteig, T., & Scaife, A. A. (2010). The association between stratospheric weak polar vortex events and cold air outbreaks in the northern hemisphere. *Quarterly Journal of the Royal Meteorological Society*, *136*(649), 886–893. https://doi.org/10.1002/qj.620

Kruger, K., Naujokat, B., & Labitzke, K. (2005). The unusual midwinter warming in the southern hemisphere stratosphere 2002. *Journal of the Atmospheric Sciences*, *62*(3), 603–613. https://doi.org/10.1175/JAS-3316.1

Lee, Y.-Y., & Black, R. X. (2015). The structure and dynamics of the stratospheric northern annular mode in CMIP5 simulations. *Journal of Climate*, *28*(1), 86–107. https://doi.org/10.1175/JCLI-D-13-00570.1

Martineau, P., Chen, G., Son, S.-K., & Kim, J. (2018). Lower-stratospheric control of the frequency of sudden stratospheric warming events. *Journal of Geophysical Research: Atmospheres*, *123*(6), 3051–3070. https://doi.org/10.1002/2017JD027648

Matsuno, T. (1970). Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *Journal of the Atmospheric Sciences*, *27*(6), 871–883. https://doi.org/10.1175/1520-0469(1970)027<0871:vpospw>2.0.co;2

Matsuno, T. (1971). A dynamical model of the stratospheric sudden warming. *Journal of the Atmospheric Sciences*, *28*(8), 1479–1494. https://doi.org/10.1175/1520-0469(1971)028<1479:admots>2.0.co;2

Matthewman, N. J., & Esler, J. G. (2011). Stratospheric sudden warmings as self-tuning resonances. Part I: Vortex splitting events. *Journal of the Atmospheric Sciences*, *68*(11), 2481–2504. https://doi.org/10.1175/JAS-D-11-07.1

Matthewman, N. J., Esler, J. G., Charlton-Perez, A. J., & Polvani, L. M. (2009). A new look at stratospheric sudden warmings. Part III: Polar vortex evolution and vertical structure. *Journal of Climate*, *22*(6), 1566–1585. https://doi.org/10.1175/2008JCLI2365.1

Maycock, A. C., & Hitchcock, P. (2015). Do split and displacement sudden stratospheric warmings have different annular mode signatures? *Geophysical Research Letters*, *42*(24), 10943–10951. https://doi.org/10.1002/2015GL066754

Mitchell, D. M., Charlton-Perez, A. J., & Gray, L. J. (2011). Characterizing the variability and extremes of the stratospheric polar vortices using 2D moment analysis. *Journal of the Atmospheric Sciences*, *68*(6), 1194–1213. https://doi.org/10.1175/2010JAS3555.1

Mitchell, D. M., Gray, L. J., Anstey, J., Baldwin, M. P., & Charlton-Perez, A. J. (2013). The influence of stratospheric vortex displacements and splits on surface climate. *Journal of Climate*, *26*(8), 2668–2682. https://doi.org/10.1175/JCLI-D-12-00030.1

Newman, P. A., & Nash, E. R. (2001). Quantifying the wave driving of the stratosphere. *Journal of Geophysical Research*, *105*(D10), 12485–12497.

Osprey, S. M., Gray, L. J., Hardiman, S. C., Butchart, N., & Hinton, T. J. (2013). Stratospheric variability in twentieth-century CMIP5 simulations of the met office climate model: High top versus low top. *Journal of Climate*, *26*(5), 1595–1606. https://doi.org/10.1175/JCLI-D-12-00147.1

Polvani, L. M., Sun, L., Butler, A. H., Richter, J. H., & Deser, C. (2017). Distinguishing stratospheric sudden warmings from ENSO as key drivers of wintertime climate variability over the North Atlantic and Eurasia. *Journal of Climate*, *30*(6), 1959–1969. https://doi.org/10.1175/JCLI-d-16-0277.1

Polvani, L. M., & Waugh, D. W. (2004). Upward wave activity flux as a precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *Journal of Climate*, *17*(18), 3548–3554. https://doi.org/10.1175/1520-0442(2004)017<3548:UWAFAA>2.0.C0;2

Scherhag, R. (1952). Die explosionsartige Stratosphären-Erwärmung des Spätwinters 1951/52. *Berichte des DWD der US-Zone*, *6*, 51–63.

Seviour, W. J. M., Gray, L. J., & Mitchell, D. M. (2016). Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models. *Journal of Geophysical Research: Atmospheres*, *121*(4), 1400–1413. https://doi.org/10.1002/2015JD024178

Seviour, W. J. M., Mitchell, D. M., & Gray, L. J. (2013). A practical method to identify displaced and split stratospheric polar vortex events. *Geophysical Research Letters*, *40*(19), 5268–5273. https://doi.org/10.1002/grl.50927

Shaw, T. A., & Perlwitz, J. (2010). The impact of stratospheric model configuration on planetary-scale waves in northern hemisphere winter. *Journal of Climate*, *23*(12), 3369–3389. https://doi.org/10.1175/2010JCLI3438.1

Shaw, T. A., Perlwitz, J., & Weiner, O. (2014). Troposphere–stratosphere coupling: Links to North Atlantic weather and climate, including their representation in CMIP5 models. *Journal of Geophysical Research: Atmospheres*, *119*(10), 5864–5880. https://doi.org/10.1002/2013JD021191

Taguchi, M. (2017). A study of different frequencies of major stratospheric sudden warmings in CMIP5 historical simulations. *Journal of Geophysical Research: Atmospheres*, *122*(10), 5144–5156. https://doi.org/10.1002/2016JD025826

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, *131*(612), 2961–3012. https://doi.org/10.1256/qj.04.176

Waugh, D. W. (1997). Elliptical diagnostics of stratospheric polar vortices. *Quarterly Journal of the Royal Meteorological Society*, *123*(542), 1725–1748. https://doi.org/10.1002/qj.49712354213

Waugh, D. W., & Randel, W. J. (1999). Climatology of Arctic and Antarctic polar vortices using elliptical diagnostics. *Journal of the Atmospheric Sciences*, *56*(11), 1594–1613. https://doi.org/10.1175/1520-0469(1999)056<1594:coaaap>2.0.co;2

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed., p. 676). Academic Press.

Wu, Z., & Reichler, T. (2019). Surface control of the frequency of stratospheric sudden warming events. *Journal of Climate*, *32*(15), 4753–4766. https://doi.org/10.1175/JCLI-D-18-0801.1

Wu, Z., & Reichler, T. (2020). Variations in the frequency of stratospheric sudden warmings in CMIP5 and CMIP6 and possible causes. *Journal of Climate*, *33*(23), 10305–10320. https://doi.org/10.1175/JCLI-D-20-0104.1